

**ÉCOLE DOCTORALE (SANTÉ, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT)**

Institut de Chimie Organique et Analytique, UMR 7311

**THÈSE** présentée par :  
**Laure GUILLOTIN**

soutenue le : **12 Novembre 2015**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : CHIMIE

**Développement d'outils enzymatiques pour  
la synthèse de protéines S-glycosylées**

**THÈSE dirigée par :**

**Richard DANIELLOU**  
**Pierre LAFITE**

Professeur, Université d'Orléans – Directeur de Thèse  
Maitre de conférences, Université d'Orléans -  
Co- encadrant de thèse

**RAPPORTEURS :**

**Magali REMAUD-SIMEON**  
**Samir MESSAOUDI**

Professeur, INSA Toulouse  
Chargé de recherche – Université Paris Sud XI

**JURY :**

**Arnaud TATIBOUET**  
**Magali REMAUD-SIMEON**  
**Samir MESSAOUDI**  
**Ludovic LANDEMARRE**  
**Pierre LAFITE**

Professeur, Université d'Orléans – Président du jury  
Professeur, INSA Toulouse  
Chargé de recherche – Université Paris Sud XI  
Directeur - Entreprise GLYcoDiag  
Maitre de conférences, Université d'Orléans -  
Co- encadrant de thèse  
Professeur, Université d'Orléans – Directeur de Thèse

**Richard DANIELLOU**











## Remerciements

Ce travail de doctorat a été réalisé au sein de l'Institut de Chimie Organique et Analytique (ICOA – UMR 7311) de l'université d'Orléans sous la direction du Professeur Richard Daniellou et du Docteur Pierre Lafite. Je tiens à remercier le Professeur Olivier Martin pour m'avoir accueillie au sein de son laboratoire.

Je remercie sincèrement Madame Magali Remaud-Simeon, Professeure de l'INSA de Toulouse, ainsi que Monsieur Samir Messaoudi, Chargé de Recherche de l'institut BioCIS de l'université Paris Sud XI, pour avoir accepté d'être rapporteurs et membres de la Commission d'Examen de cette thèse.

Je souhaite également remercier Monsieur Arnaud Tatibouët, Professeur de l'université d'Orléans, et Monsieur Ludovic Landemarre, Directeur de l'entreprise GLYcoDiag, d'avoir accepté de faire partie de cette Commission d'Examen.

Ces trois années ont été une expérience scientifique intense et enrichissante, ponctuée de périodes de doutes, d'espoirs et de joies mais elles ont également été le lieu d'une aventure humaine extraordinaire. Pour cela je tiens à remercier en premier lieu Richard, mon directeur de thèse, et Pierre, mon co-encadrant, pour m'avoir accueillie au sein de l'équipe « Glycobio » et sans qui la réalisation de ce projet n'aurait pas été possible. Richard je te remercie de m'avoir fait découvrir le monde des enzymes, ton pari est réussi la chimiste que je suis est devenue une vraie mordue de la biocatalyse ! Merci aussi pour m'avoir accordé ta confiance pour la conduite de ce projet mais également pour ton soutien et tes compétences qui m'ont permis de surmonter les difficultés du sujet. Pierre, je ne te remercierai jamais assez pour ta présence, tes conseils précieux et ton optimisme infailible qui a su me remonter le moral quand moi je n'y croyais plus. Je n'oublie pas non plus ce cher Kiwi, et sa moustache, sans qui nos expériences de cristallogenèse n'auraient pas été possibles ! ☺

Je tiens à exprimer mes remerciements chaleureux à mon voisin de bureau, Philippe, toujours prompt à rendre service et avec qui j'ai pu partager de nombreuses discussions. Merci à toi pour ton sourire et tes anecdotes pour détendre l'atmosphère.

Ce projet a également fait l'objet d'un partenariat avec l'entreprise GLYcoDiag pour la réalisation des analyses de microbiologie. J'adresse donc mes remerciements à ses membres, Ludovic, son directeur, pour ses conseils avisés mais également Perrine avec qui j'ai partagé d'agréables moments.

Je tiens à adresser mes plus vifs remerciements au Professeur Sabine Berteina-Raboin et au Docteur Marie-Aude Hiebel pour m'avoir permis de réaliser l'ensemble de mes manipulations de synthèse organique au sein de leur équipe. Je remercie chaleureusement Laëtitia Fougère pour sa grande disponibilité et pour m'avoir fait bénéficier de son expertise en LC-MS. Merci pour ton aide et ta bonne humeur, cela a été agréable de travailler avec toi.

Merci à toute la « clique » de l'ICOA qui, malgré l'éloignement géographique (et oui une rue à travers ce n'est pas la porte à côté...), m'a permis d'apprécier ces trois années passées à Orléans. Merci à Geoffrey, Romain, Jo, Ben, Cédric, Matthieu, Max, Audrey, Jihen, ... pour tous ces bons moments passés ensemble. Je n'oublierais pas nos soirées du mercredi soir,

les sessions de volley intenses ou encore la valse des blagues plus ou moins douteuses qui nous ont valu plus d'un fou rire. En cette période de coupe du monde, impossible d'oublier les discussions autour du rugby avec les représentants officiels de la région Auvergne, je pense bien évidemment aux fidèles jaunards MP (futur Président de Région), Geof' et Audrey.

J'adresse une mention spéciale à mon ami Geoffrey (ou Goffray), mon complice d'Edifice, ainsi qu'à mon petit rayon de soleil Caennais, Elo, avec qui j'ai pu tisser des liens très forts durant cette thèse. Je garderai en mémoire nos moments de délires et de franches rigolades pour décompresser quand on en avait besoin.

Je témoigne mon amitié sincère à mes deux camarades qui ont rejoint les rangs de la « Glycobio », Jiji (petit Padawan) et Cédric. Merci pour votre fraîcheur et votre joie de vivre communicative, j'ai beaucoup apprécié travailler (enfin rédiger) à vos côtés. Avec vous, je crois que la relève est assurée ! Je pense également à l'ensemble des stagiaires qui sont passés au laboratoire, Tibo, GuiJou, Ana, Nikolina, Maya, Kévin, Donia et Mélanie.

Je remercie tous mes amis, d'ici ou d'ailleurs pour leurs encouragements et leur présence essentielle. Votre amitié m'est très précieuse.

Enfin je veux remercier ma famille qui m'a toujours apporté un soutien inconditionnel tout au long de mes études. Ces dernières lignes sont adressées à mes parents et à mon frère Yann, pour tout leur amour et pour être toujours présents à mes côtés.

## Abbreviations

**[M +Na]<sup>+</sup>** = Espèce cationique correspondant à l'adduit « molécule + sodium »

**[M +H]<sup>+</sup>** = Espèce cationique correspondant à l'adduit « molécule + proton »

**[M -H]<sup>-</sup>** = Espèce anionique correspondant à l'adduit « molécule - proton »

**AA** : Acide aminé

**AA<sub>n</sub>** : Acide aminé non-naturel

**AARS** : Aminoacyl-ARNt-synthétase

**ACN** : Acétonitrile

**A<sub>H</sub>** : Activité d'hydrolyse

**AR** : Analytical Reagent

**ARN<sub>m</sub>** : ARN messenger

**ARN<sub>t</sub>** : ARN transférase

**a<sub>ROH</sub>** : Activité de l'alcool

**A<sub>T</sub>** : Activité de transglycosylation

**a<sub>w</sub>** : Activité de l'eau

**αRha** : α-rhamnosidase

**βGlc** : β-glucosidase

**βMan** : β-mannosidase

**βME** : β-Mercaptoéthanol

**BSA** : Bovine Serum Albumine

**BuOH** : Butanol

**Cam** : Chloramphénicol

**CAZy** : Carbohydrate Active enZymes database

**CAZymes** : Carbohydrate Active enZymes

**CCM** : Chromatographie sur couche mince

**Coll.** : Collaborateurs

**C-term** : Extrémité C-terminale

**CTP** : 4-Chlorothiophénol

**DIPEA** : *N,N*-Diisopropylethylamine

**DCM** : Dichlorométhane

**DMF** : Diméthylformamide

**DMSO** : Diméthylsulfoxyde

**DNP** : Dinitrophényle

**DO** : Densité optique

**DtGly** : *Dictyoglomos thermophilum* Glucosidase

**DtMan** : *Dictyoglomos thermophilum* Mannosidase

**DtRha** : *Dictyoglomos thermophilum* Rhamnosidase

**DTT** : Dithiothréitol

**EC** : Enzyme Classification

**ESI** : Electrospray ionisation

**EtOAc** : Acétate d'éthyle

**Et<sub>3</sub>N** : Triéthylamine

**Fuc** : Fucopyranoside

**Gal** : Galactopyranoside

**GalNac** : *N*-acétylgalactosamine

**gCOSY** : gradient COrelation SpectroscopY

**GDP** : Guanosine diphosphate

**GG** : Glycoside de glycérol

**GH** : Glycoside Hydrolase

**gHSQC** : gradient Heteronuclear Single Quantum Coherence

**Glc** : Glucopyranoside

**GlcNac** : *N*-acétylglucosamine

**GT** : Glycoside Transferase

**His•Tag** : Séquence *N*-terminale et/ou C-terminale de 6 Histidines

**HRMS** : Spectrométrie de masse haute résolution

**HMBC** : Heteronuclear Multiple Bond Coherence

**Kan** : Kanamycine

**k<sub>cat</sub>** : Turn-over number

**K<sub>M</sub>** : Constante de Michaelis-Menten

**LB** : Lysogeny Broth

**LC-MS** : Liquid Chromatography-Mass Spectrometry

**LCN** : Ligation Chimique Native

**LPE** : Ligation de Protéine Exprimée

**IMAC** : Immobilized Metal Affinity Chromatography

**IPTG** : Isopropyl β-D-1-thiogalactopyranose

**iPrOH** : Isopropanol

**Man** : Mannopyranoside

**MPD** : 2-méthyl-2,4-pentandiol  
**MS** : Mass Spectrometry  
**M** : Poids moléculaire (MW : Molecular Weight)  
**N-term** : Extrémité N-terminale  
**pb** : Paire de bases  
**PCR** : Polymerase Chain Reaction  
**PDB** : Protein Data Bank  
**PE** : Petroleum ether  
**pH** : Hydrogen potential  
**pHOpt** : pH Optimal  
**pKa** : Potentiel de la constante d'acidité  
**pNP** : *para*-nitrophényle  
**pNPFuc** : pNP-β-D-fucopyranoside  
**pNPGlc** : pNP-β-D-glucopyranoside  
**pNPGal** : pNP-β-D-galactopyranoside  
**pNPMan** : pNP-β-D-mannopyranoside  
**pNPOH** : *para*-nitrophénol  
**pNPRha** : pNP-α-L-rhamnopyranoside  
**pNPXyl** : pNP-β-D-xylopyranoside  
**ppm** : partie par million  
**R** : Rendement  
**R<sub>f</sub>** : Reference front

**r<sub>H</sub>** : Vitesse d'hydrolyse  
**r<sub>S</sub>** : Vitesse de synthèse  
**r.<sub>S</sub>** : Vitesse de déalkylation  
**Rha** : Rhamnose  
**RMN** : Résonance magnétique nucléaire  
**S, S<sub>c</sub>** : Facteurs de sélectivité  
**SDS-PAGE** : Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis  
**SEC** : Chromatographie d'exclusion stérique  
**SPPS** : Synthèse Peptidique en Phase Solide  
**SR** : Site de restriction  
**tBoc** : *Tert*-butoxycarbonyl  
**TP** : Thiophénol  
**T<sub>R</sub>** : temps de rétention  
**Tris** : Tris(hydrosyméthyl)aminométhane  
**Ttyr** : Thiotyrosine  
**TtyrME** : Thiotyrosine méthyl ester  
**UDP** : Uridine diphosphate  
**Wt** : Wild-type  
**Xyl** : Xylopyranoside

**Abbreviations des acides aminés :**

Nom	Code à 1 lettre	Code à 3 lettres	Nom	Code à 1 lettre	Code à 3 lettres
Alanine	A	Ala	Leucine	L	Leu
Arginine	R	Arg	Lysine	K	Lys
Asparagine	N	Asn	Méthionine	M	Met
Aspartate	D	Asp	Phénylalanine	F	Phe
Cystéine	C	Cys	Proline	P	Pro
Glutamate	E	Glu	Sérine	S	Ser
Glutamine	Q	Gln	Thréonine	T	Thr
Glycine	G	Gly	Tryptophane	W	Trp
Histidine	H	His	Tyrosine	Y	Tyr
Isoleucine	I	Ile	Valine	V	Val



# SOMMAIRE

<b>Introduction générale</b>	<b>17</b>
<b>I. Les glycosidases, outils pour la synthèse de molécules complexes</b>	<b>20</b>
1. Généralités et aspects mécanistiques	20
1.2. Mécanismes des glycosidases	21
1.2.a. Mécanismes usuels	21
1.2.b. Mécanismes non-usuels	23
2. Application à la synthèse de O-glycoconjugués	25
2.1. Les glycoside hydrolases natives	25
2.1.a. L'hydrolyse inverse, approche thermodynamique	25
2.1.b. La transglycosylation, approche cinétique	27
2.2. Les glycosynthèses	29
2.3. Ingénierie combinatoire des glycoside hydrolases	33
<b>II. Les thioglycosides, briques élémentaires pour la synthèse d'oligosaccharides complexes</b>	<b>35</b>
1. Les thioglycosides naturels	35
1. Applications en synthèse organique	37
2. Synthèse de thioglycosides par voie enzymatique	40
2.1. Les glycoside hydrolases natives	40
2.2. Les thioglycoligases et thioglycosynthèses	41
<b>II. La liaison thioglycosidique, forme rare chez les protéines glycosylées</b>	<b>46</b>
1. Les sucres et les protéines, partenaires clés du monde du vivant	46
1.1. Les formes rares de protéines glycosylées	47
1.2. Biosynthèse des thioglycoprotéines, les S-glycosyltransférases naturelles	49
2. Outils de synthèse pour l'accès aux néothioglycopeptides et protéines	51
2.1. Glycosylation par voie chimique	51
2.2. Les outils enzymatiques, nouvelles perspectives ?	56
<b>III. Objectifs de la thèse</b>	<b>58</b>
<b>Bibliographie</b>	<b>63</b>
<b>Chapitre 1 - Développement d'une banque de glycosides hydrolases natives : caractérisation et valorisation</b>	<b>69</b>
<b>I. Stratégie de production : Du choix d'un microorganisme vers la conception d'une banque de glycosidases natives</b>	<b>71</b>
1. <i>Dictyoglomus thermophilum</i>	72
2. De la sélection des gènes à l'utilisation des outils de biologie moléculaire	74
2.1. Plasmide de clonage pET28-a(+)	75
2.2. Vecteur d'expression : <i>E.coli</i> Rosetta-(DE3)	77
<b>II. Caractérisation de nouvelles glycosidases recombinantes</b>	<b>81</b>
1. DtMan, $\beta$ -D-mannosidase de <i>D. thermophilum</i>	81
1.1. Conditions optimales de réaction et paramètres cinétiques	85

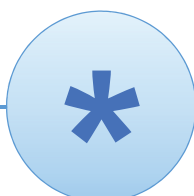


1.2. Influence du co-solvant	88
1.3. Cristallogenèse de DtMan	89
2. DtRha, $\alpha$ -L-rhamnosidase de <i>D.thermophilum</i>	94
2.1. Conditions optimales de réaction et paramètres cinétiques	96
2.2. Influence du co-solvant	98
2.3. Cristallogenèse de DtRha	99
3. DtGly, $\beta$ -D-glucosidase de <i>D. thermophilum</i>	102
3.1. Conditions optimales de réaction et paramètres cinétiques	105
3.2. Influence du co-solvant	107
3.3. Cristallogenèse de DtGly	108
<b>III. Valorisation de biocatalyseurs pour la synthèse de glycosides d'alkyle, applications en cosmétique</b>	<b>108</b>
1. Criblage d'activité transglycosylase	109
2. Application à la synthèse de glycosides de glycérols	112
2.1. Voie de synthèse chimio-enzymatique	114
2.2. Recherche de propriétés antimicrobiennes des glycoconjugués synthétisés	115
<b>IV. Conclusion</b>	<b>118</b>
<b>Bibliographie</b>	<b>121</b>
<b>Chapitre 2 - Conception et évaluation de mutants thioglycoligases pour la synthèse de S-glycoconjugués</b>	<b>125</b>
<b>I. Thioglycoligases : rappels et généralités</b>	<b>127</b>
1. Rappels	127
2. Création de points de mutation sélectifs par mutagenèse dirigée	130
<b>II. Production des mutants thioglycoligases de DtGly : Application à la synthèse de thioglycosides</b>	<b>133</b>
1. DtGly E159A et E159Q	133
1.1. Conception et surexpression des mutants	133
1.2. Caractérisation	135
2. Synthèse de S-glycoconjugués catalysée par DtGly E159Q	136
2.1. Etude préliminaire	136
2.2. Synthèse de thioglycoconjugués, relation structure/activité	140
<b>III. Etude de l'effet de mutations sur l'activité de synthèse de DtMan</b>	<b>145</b>
1. Etude de variants thioglycoligases de DtMan	145
1.1. Conception, et production de la banque de mutant DtMan E425X	146
1.2. Caractérisation	149
1.3. Essais de thioglycosylation	154
2. Etude cristallographique DtMan E425C	157
<b>IV. Conclusion</b>	<b>160</b>
<b>Bibliographie</b>	<b>163</b>

<b>Chapitre 3 - Validation du concept de thioglycosylation enzymatique d'un acide aminé</b>	<b>165</b>
<b>I. Introduction à la glycosylation enzymatique de protéines</b>	<b>167</b>
<b>II. Synthèse de l'accepteur de glycoside, la Thiotyrosine</b>	<b>172</b>
1. Synthèse de l'acide aminé non-naturel	172
2. Etude de la Thiotyrosine en solution	177
<b>III. Couplage glycosidique de la Thiotyrosine catalysé par DtGly E159Q</b>	<b>180</b>
1. Optimisation des conditions et de suivi de la réaction enzymatique	180
2. Thioglucosylation enzymatique de TtyrME catalysée par DtGly E159Q	188
<b>IV. Conclusion</b>	<b>197</b>
<b>Bibliographie</b>	<b>199</b>
<b>Conclusion Générale &amp; Perspectives</b>	<b>201</b>
<b>Bibliographie</b>	<b>215</b>
<b>Experimental Part</b>	<b>217</b>
<b>Bibliographie</b>	<b>267</b>
<b>ANNEXES</b>	<b>269</b>



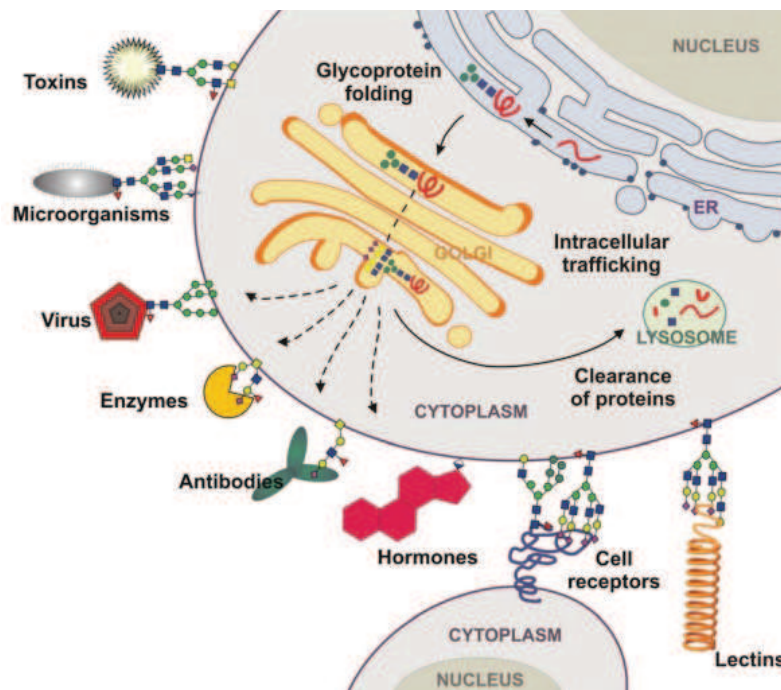
## Introduction générale





## Introduction générale

Dans le monde du vivant, les glycosides font partie des familles de molécules les plus fréquemment rencontrées. Ces glycanes peuvent se trouver sous la forme de monosaccharides (lorsqu'ils sont issus du métabolisme énergétique par exemple), d'oligo- ou polysaccharides (comme éléments structuraux de la paroi cellulaire des plantes) ou de glycoconjugués (tels que les glycoprotéines et glycolipides). (Campo, **2010**; Defaus, **2014**) Principalement présents à la surface des cellules, ces glycoconjugués sont impliqués dans des processus biologiques variés tels que les mécanismes de reconnaissance ou d'adhésion cellulaire, les modifications co- et post-traductionnelles des protéines, l'infection par des agents pathogènes ou dans la réponse immunitaire (Figure 1). (Varki, **1993**) En conséquence, les sucres jouent un rôle important dans l'apparition de nombreuses maladies et sont devenus d'un intérêt majeur pour la communauté scientifique comme nouvelles cibles thérapeutiques ou de diagnostique. (Ernst, **2009**; Fernandez-Tejada, **2015**) Depuis les 30 dernières années, les domaines de la glycochimie et de la glycobiologie se sont développés pour favoriser l'accès à ces composés et améliorer nos connaissances, notamment en terme de relation structure/activité, nécessaires au développement de nouveaux candidats médicaments, tels que des vaccins anti-cancéreux ou de lutte contre les maladies infectieuses. (Fernandez-Tejada, **2015**) Cependant les applications des glycoconjugués ne se limitent pas à l'industrie pharmaceutique, ils sont par exemple exploités pour leurs propriétés tensio-actives dans les domaines de la cosmétique, de la détergence ou l'industrie agro-alimentaire. (Mladenoska, **2008**)



**Figure 1** : Rôles des glycanes dans les mécanismes cellulaires. (Defaus, **2014**)

D'un point de vue structural, les charpentes osidiques font partie des molécules les plus complexes, de par i) le type de motif sucre (pyranose, furanose, série D ou L), ii) le type de liaison pouvant être générée (1,1)-(1,2)-(1,3)-(1,4)-(1,5)-(1,6), iii) la configuration de la position anomérique ( $\alpha/\beta$ ) et iv) le type de liaison osidique (S, N, O,...).(Lafite, **2012**) Pour cela, la synthèse de formes homogènes de glycanes par voie chimique est un véritable challenge. Elle requiert la plupart du temps la préparation de donneurs activés, le passage par de nombreuses étapes de protection/déprotection pour contrôler la régiosélectivité de la réaction ainsi que la stéréosélectivité de la liaison créée et la séparation des différents produits formés qui conduit à de faibles rendements finaux. Parallèlement, en s'inspirant des outils mis à disposition par la nature, les chercheurs ont développé des méthodes de production par voie enzymatique en exploitant le potentiel de synthèse des glycosyltransférases (GT) et des glycoside hydrolases (GH). Si les GT, responsables *in-vivo* de la biosynthèse des glycosides, présentent une efficacité évidente pour le transfert d'unités glycosidiques, leur application en synthèse reste notamment limitée par l'utilisation de nucléotides sucres et de co-facteurs, tels que l'uridine diphosphate (de coût élevé) ainsi que par leur disponibilité sous forme recombinante.(Gantt, **2011**; Toone, **1989**; van Rantwijk, **1999**) En revanche, le développement des connaissances mécanistiques et structurelles des glycoside hydrolases, responsables du métabolisme des glycosides, a permis la conception de biocatalyseurs originaux présentant des propriétés optimisées pour la synthèse de dérivés glycosylés.

## I. Les glycosidases, outils pour la synthèse de molécules complexes

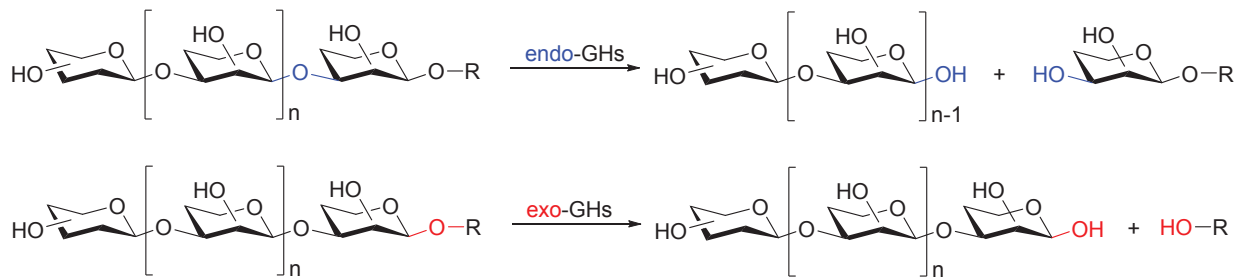
### 1. Généralités et aspects mécanistiques

#### 1.1. Classification

Les glycoside hydrolases sont des enzymes ubiquitaires spécifiques des sucres impliquées dans de nombreux processus biologiques. Chez les mammifères elles se retrouvent principalement sous forme extracellulaire, comme dans la salive ou le tube digestif, et jouent un rôle important dans la dégradation de glycanes à des fins nutritives ou de lutte anti-bactérienne. D'autres GH sont quant à elle localisées dans le réticulum endoplasmique ou l'appareil de Golgi et participent aux processus de modifications post-traductionnelles.(Fernandez Mayoralas, **1997**; Rosnoblet, **2013**) A l'inverse des glycosyltransférases, les GH (EC 3.2.1.xx) catalysent l'hydrolyse de la liaison glycosidique entre deux unités sucres ou une unité sucre et un aglycone (entité non glycosidique). Elles peuvent être classifiées en deux groupes distincts selon leur capacité à cliver les liaisons internes ou externes du substrat (Figure 2) :

- Les « endo »-glycosidases hydrolysent la liaison interne d'une chaîne oligosaccharidique

- Les « exo »-glycosidases catalysent la rupture de la liaison glycosidique à l'extrémité non réductrice d'un oligosaccharide.



**Figure 2 :** Les endo/exo-Glycoside Hydrolases.

Du fait de la diversité moléculaire des sucres, les glycosidases présentent un large spectre d'activité qui a nécessité l'élaboration d'une classification plus spécifique et plus précise que la nomenclature de l'Union Internationale de Biochimie et de Biologie Moléculaire (IUBMB) des enzymes. En 1998, la base de données regroupant l'ensemble des Carbohydrate-Active enZymes (CAZy) fut créée, rassemblant l'ensemble des GH en différentes familles en fonction des similarités dans leur séquence en acides aminés et leur structure tridimensionnelle. (Cantarel, **2009**; Henrissat, **1991b**; Lombard, **2014**) A l'heure actuelle plus de 130 familles de glycosidases ont été identifiées permettant ainsi d'étendre les connaissances sur les mécanismes et le mode de repliement de ces enzymes (<http://www.cazy.org/> 05/09/2015). (Henrissat, **1991a**) Pour le glycobiologiste, cette classification CAZy est devenue un outil précieux et indispensable pour répertorier au sein d'une famille une GH nouvellement identifiée.

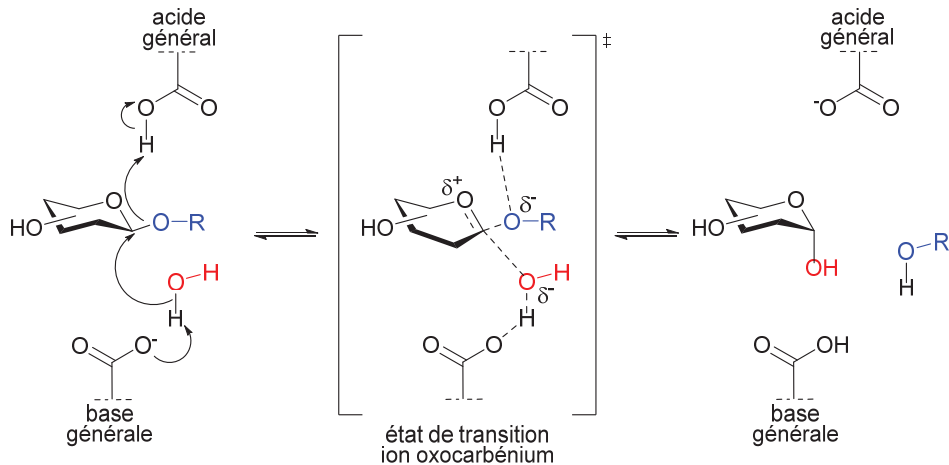
## 1.2. Mécanismes des glycosidases

### 1.2.a. Mécanismes usuels

Depuis de nombreuses années, l'intérêt porté aux GH n'a cessé de s'accroître et a permis de mieux comprendre le mode d'action de ces catalyseurs. Deux grands mécanismes ont alors été mis en évidence par le Prof. Koshland et font toujours l'objet de nombreuses études grâce notamment à la résolution des structures tridimensionnelles d'enzymes en type sauvage, mutées ou co-cristallisées avec des ligands à partir d'expérience de cristallographie aux rayons X (RX). L'hydrolyse de la liaison glycosidique peut être réalisée avec inversion ou rétention de la configuration de la position anomérique. (Koshland, **1953**)

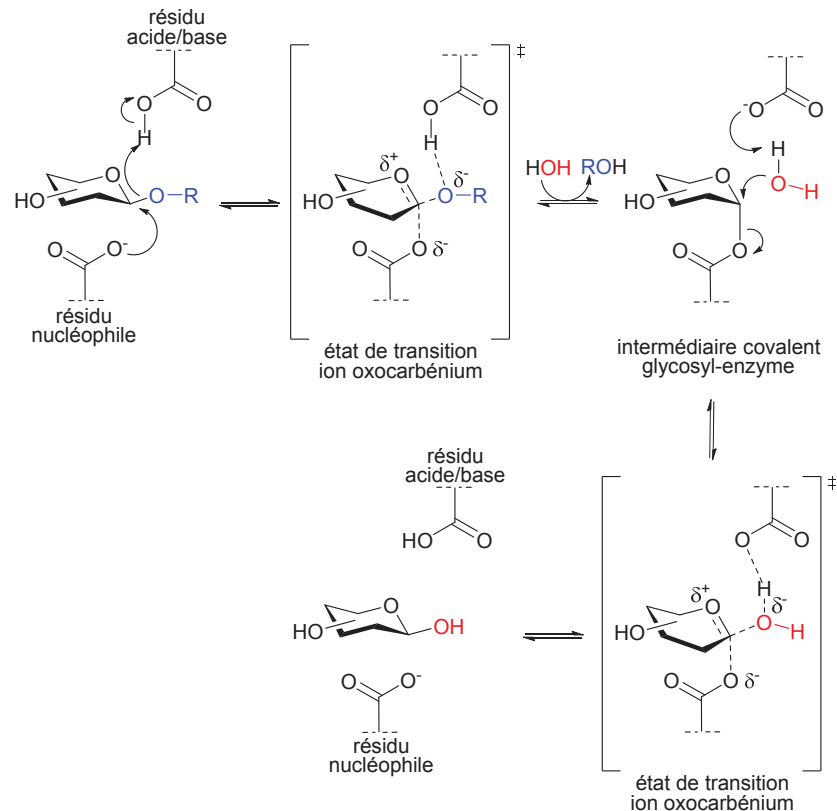
Chez les glycosidases avec inversion de configuration, l'hydrolyse suit un mécanisme de déplacement en une étape réalisé entre 2 résidus catalytiques, jouant le rôle d'acide/base général, et une molécule d'eau. Dans le site actif, les 2 acides carboxyliques clés sont approximativement séparés l'un de l'autre de 10 Å favorisant l'attaque nucléophile sur le côté opposé du lien osidique. (Rye, **2000**) Pendant l'acte catalytique, le résidu jouant le rôle de base générale active une molécule d'eau entrant dans le site actif tandis que le second acide carboxylique déclenche, par catalyse acide, le départ du groupe partant en position anomérique (Schéma 1).





**Schéma 1** : Mécanisme des GH à inversion de configuration.

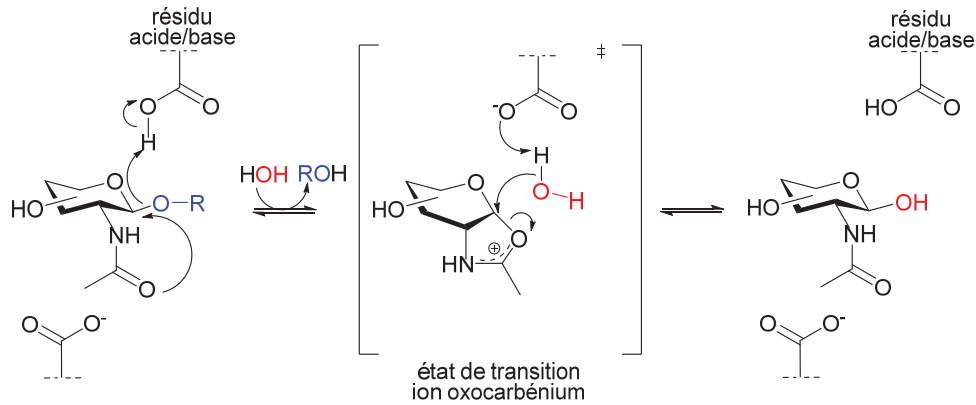
Le second mécanisme majoritaire rencontré chez les glycosidases est le mécanisme avec rétention de configuration. Il s'opère selon un double déplacement de liaisons favorisant ainsi une double inversion de la configuration de la position anomérique. Si les mécanismes d'inversion et de rétention de configuration partagent des caractéristiques communes telles que : i) le passage par un état de transition de type oxocarbenium et ii) la participation de 2 acides carboxyliques catalytiques, ils se différencient sur certains points. En effet chez les GHs à rétention de configuration, les 2 résidus catalytiques sont à une distance de 5,5 Å l'un de l'autre. (Rye, **2000**) La proximité de ces 2 acides aminés ne permet pas l'action simultanée du nucléophile entrant (molécule d'eau) et des résidus catalytiques comme décrit dans le précédent mécanisme. Ainsi dans un premier temps, le départ du groupe partant est favorisé par catalyse acide (assuré par l'un des résidus jouant le rôle d'acide général), tandis que le résidu nucléophile (acide aminé sous forme de carboxylate) attaque le centre anomérique pour former un intermédiaire covalent aussi appelé intermédiaire glycosyl-enzyme. Dans un second temps, le résidu acide/base déprotone une molécule d'eau pénétrant dans le site actif. Ce nucléophile activé peut alors attaquer l'intermédiaire covalent par le côté opposé à la liaison anomérique et générer le produit d'hydrolyse correspondant (Schéma 2). (Bojarova, **2009**; Fernandez Mayoralas, **1997**; Hancock, **2006**; Rye, **2000**)



**Schéma 2** : Mécanisme des GH à rétention de configuration.

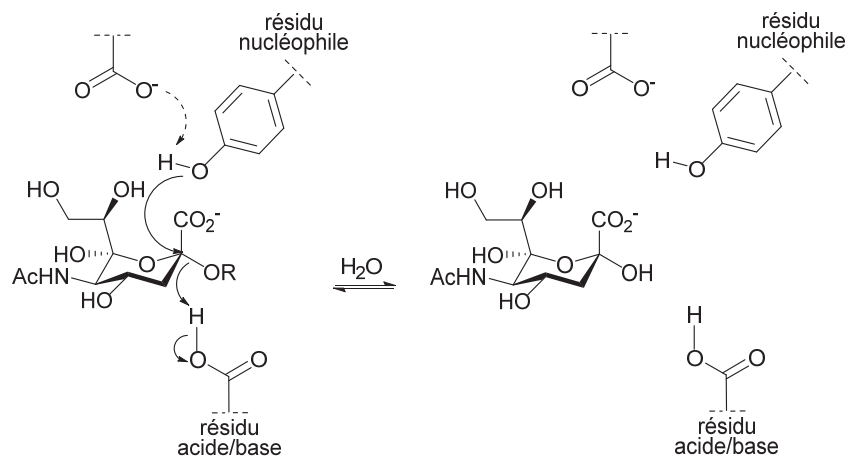
### 1.2.b. Mécanismes non-usuels

Si les 2 mécanismes d'inversion et de rétention de configuration sont les plus rencontrés chez les glycosidases, différentes études décrivent néanmoins l'existence de mécanismes plus atypiques chez ces enzymes. Parmi ces dernières sont rapportées les  $\beta$ -*N*-acétylhexoaminidases, glycosidases extrêmement spécifiques des sucres portant un groupement déoxy *N*-acétyl en position 2. A la fin du XX<sup>ème</sup> siècle ces enzymes ont fait l'objet de nombreuses études, notamment pour des applications anti-fongiques. (Horsch, **1997**) A l'heure actuelle ces biocatalyseurs sont répertoriés dans les familles 3, 18, 20, 84 et 85 des GHs ([www.cazy.org/](http://www.cazy.org/) 07/07/2015). (Abbott, **2009**; Rao, **2006**; Rao, **1999**; Vocadlo, **2000**) Des recherches réalisées sur le développement d'inhibiteurs d'hexoaminidases conjuguées à des études cristallographiques ont mis en évidence que les enzymes appartenant aux familles 18 et 20 des GH clivaient la liaison osidique *via* un mécanisme dit « substrat-assisté » (Schéma 3). Le groupement *N*-acétyl en position C-2 joue ici le rôle de nucléophile pour attaquer le centre anomérique et générer un intermédiaire oxazolonium. L'intervention d'une molécule d'eau activée par le résidu servant de base générale conduit au produit d'hydrolyse correspondant. Tout comme les GH « classiques » les  $\beta$ -*N*-acétylhexoaminidases possèdent dans leur site actif 2 acides carboxyliques catalytiques mais ne jouant pas les mêmes rôles. En effet, si le rôle du résidu acide/base est conservé dans la réaction catalytique, des expériences de mutagenèse et de cristallographie menées par Rao *et coll.* ont démontré que le second carboxylate n'avait aucun rôle catalytique mais participait en revanche à la stabilisation de l'intermédiaire oxazolonium. (Rao, **1999**)



**Schéma 3 :** Mécanisme des  $\beta$ -N-acétylhexoaminidases.

Les sialidases et trans-sialidases, regroupées dans les familles GH33 et GH34, ont quant à elles été décrites pour catalyser l'hydrolyse ou le transfert d'acides sialiques en utilisant un résidu nucléophile alternatif au glutamate (Glu) ou aspartate (Asp) classiquement rencontrés. En effet, les acides sialiques (substrats de ces enzymes) possèdent en position anomérique une fonction acide carboxylique, ainsi pour limiter l'effet de répulsion électrostatique avec un résidu Glu ou Asp ces enzymes utilisent une tyrosine, activée par un résidu basique adjacent, pour jouer le rôle de nucléophile et former l'intermédiaire glycosyl-enzyme (Schéma 4). (Watts, **2003**) Enfin un mécanisme atypique a été mis en évidence chez les représentantes de GH4 qui nécessitent l'assistance unique du  $\text{NAD}^+$  et d'un ion divalent ( $\text{Mn}^{2+}$ ) lors de l'acte catalytique. (Lodge, **2003**; Rajan, **2004**; Thompson, **1998**; Varrot, **2005**) Ces enzymes ne suivent pas un mécanisme à double déplacement mais agissent selon un processus catalytique multi-étapes impliquant des réactions d'oxydation, élimination, addition et de réduction.



**Schéma 4 :** Mode d'action des sialidases.

## 2. Application à la synthèse de O-glycoconjugués

### 2.1. Les glycoside hydrolases natives

De nos jours, les phénomènes de société tels que le réchauffement climatique, l'épuisement des ressources fossiles ou la gestion des déchets ont propulsé le remplacement des procédés chimiques industriels coûteux en énergie et qui nécessitent l'utilisation de solvants organiques. L'utilisation de systèmes éco-compatibles répondants aux principes de la chimie verte est devenue un challenge d'intérêt. (Menon, **2012**) Dans cette optique, les systèmes enzymatiques sont parmi les meilleures alternatives puisqu'ils sont :

- i) actifs en conditions douces (dans l'eau à température ambiante et à pression atmosphérique), (Arnold, **2001**)
- ii) peu polluants, avec un faible rejet de déchets toxiques, (Arnold, **2001**)
- ii) recyclables sur plusieurs cycles catalytiques (par greffage sur support solide) (Ansari, **2012**)
- iv) *etc...*

Les sucres sont au cœur de nombreux procédés industriels et les glycosidases constituent un pool enzymatique extrêmement diversifié tout d'abord d'un point de vue mécanistique mais aussi compte tenu de la grande variété de substrats pouvant être reconnus. Les GH rassemblent, par conséquent, tous les critères pour devenir des outils prometteurs et originaux dans de nombreuses applications. En dégradant les glycanes, elles trouvent une utilité dans l'industrie du textile, papetière ou dans les bioraffineries. (Bergquist, **2014**; Choi, **2010**) Par exemple, dans un souci de proposer des alternatives aux énergies fossiles, l'intérêt porté aux xylanases et cellulases s'est fortement accru afin de proposer des systèmes biocatalytiques efficaces pour la dégradation de biomasse lignocellulosique et la production de bioéthanol de 2<sup>de</sup> génération. (Bhalla, **2013**; Juturu, **2014**; Menon, **2012**)

Dans les années 70 les études réalisées sur les glycosidases ont prouvé leur capacité à synthétiser de nouvelles liaisons glycosidiques, par hydrolyse inverse ou réaction de transglycosylation. Depuis ces découvertes le développement des outils de bio-ingénierie ont permis d'améliorer les activités de synthèse de ces enzymes qui trouvent des applications pour la production de O-glycoconjugués dans les domaines de la pharmaceutique, de la cosmétique ou encore pour l'agroalimentaire. (Farran, **2015**)

#### 2.1.a. *L'hydrolyse inverse, approche thermodynamique*

Au début du XX<sup>ème</sup> les équipes des Prof. Bourquelot et Viebel ont initié la synthèse de glycosides d'alkyle par voie enzymatique en synthétisant pour la première fois des dérivés du glucose en présence d'émulsine (mélange de glycoside hydrolases extrait d'amande). Leurs travaux ont alors mis en évidence la possibilité d'inverser la réaction d'hydrolyse des glycosidases en faveur de la création d'une nouvelle liaison osidique. (Bourquelot, **1913**; Viebel, **1936**) La réaction d'hydrolyse inverse s'effectue en présence d'une glycosidase, du monosaccharide et d'un nucléophile tel qu'un alcool. Sous contrôle thermodynamique, elle conduit cependant à la formation majoritaire du produit d'hydrolyse. En effet lors de l'acte

catalytique, le dérivé glycosylé nouvellement formé peut à son tour devenir substrat de l'enzyme et donc conduire au produit d'hydrolyse, produit thermodynamique. Par conséquent ces réactions d'hydrolyse inverse sont la plupart du temps limitées par de faibles rendements (Figure 3).



**Figure 3 :** L'hydrolyse inverse réaction sous contrôle thermodynamique.

De nombreuses études ont été réalisées depuis les travaux pionniers de Bourquelot et Viebel mettant en évidence l'influence de différents paramètres thermodynamiques pour déplacer l'équilibre dans le sens de la synthèse et non de l'hydrolyse :

-i) La concentration en réactifs (donneur/accepteur).

Plusieurs travaux font référence à l'observation d'un meilleur rendement de glycosylation en présence d'une forte concentration en réactifs de départ. (Ajisaka, **1987**; Chahid, **1992**; Monsan, **1995**) Si de meilleurs rendements d'hydrolyse inverse sont obtenus en augmentant la concentration en donneur de glycoside, elle donne en revanche lieu à la formation de sous-produits indésirables. (Ajisaka, **1987**) Enfin, si les glycosidases acceptent un large panel de substrats, la glycosylation directe d'alcool primaire s'avère être 3 à 5 fois plus rapide que pour les alcools secondaires. (van Rantwijk, **1999**; Vic, **1992**)

-ii) La concentration en eau.

Etymologiquement, le terme hydrolyse provient du grec *hudôr* + *lusi* qui signifie dissolution à l'eau. Il serait donc facile de penser que remplacer l'eau par un système non-aqueux permettrait de déplacer l'équilibre dans le sens unique de la glycosylation directe. Cependant, la présence d'eau dans le milieu réactionnel est essentielle au bon fonctionnement des glycosidases mais également pour la solubilisation des glycosides (non-protégés), de par leur nature hydrophile. Il a alors été établi qu'une concentration minimale en eau  $a_w$  (activité de l'eau) devait être assurée. (Basso, **2002**; van Rantwijk, **1999**) Différentes études réalisées sur des systèmes biphasiques ont relevé que l'addition de co-solvant suivant leur nature (polarité, solvant protique, aprotique) augmentait les rendements de condensation par comparaison aux réactions réalisées en milieu aqueux. (Panintrarux, **1995**; Vic, **1992**)

-iii) Éliminer le produit de la réaction.

L'équilibre réactionnel peut être déplacé en éliminant le produit de condensation du milieu réactionnel par précipitation, extraction ou transformation en un autre produit. Un exemple de cette approche est la synthèse du raffinose en flux continu *via* l'immobilisation d'une  $\alpha$ -D-galactosidase de *Mycena vinacea*. Le procédé développé par Ajisaka *et coll.* consiste à faire passer une solution contenant 10% de galactose ( $^w/v$ ) et 50% de sucrose ( $^w/v$ ) à travers une série de colonnes d'enzymes immobilisées et de charbon actif. Les trisaccharides ayant une meilleure affinité pour le charbon actif que les composés mono- ou di-saccharidiques, l'équilibre de la réaction peut être continuellement déplacé dans le sens de la condensation

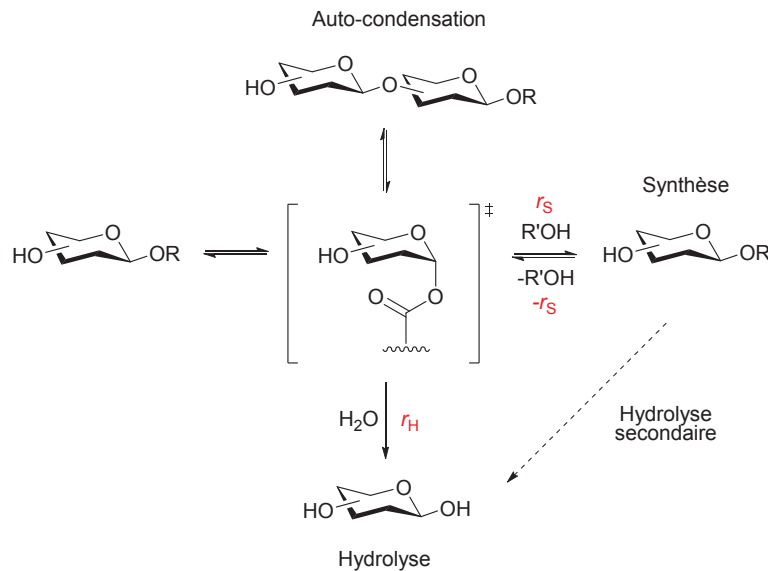
par extraction du trisaccharide accumulé dans la colonne de charbon actif. Cette méthode nécessite de déterminer le temps de circulation minimum nécessaire à la formation et l'accumulation du raffinose avant son élution (sol. EtOH 50%). (Ajisaka, **1989**)

L'amélioration des performances des glycosidases en hydrolyse inverse repose sur un savant dosage de ces différents paramètres thermodynamiques. Malgré tout, la synthèse d'oligosaccharides et de dérivés glycosylés par hydrolyse inverse reste limitée par de faibles rendements et la formation d'isomères non désirables. Les GH sont également capables de créer des liaisons glycosidiques par réaction de transglycosylation qui s'effectue sous contrôle cinétique.

### 2.1.b. La transglycosylation, approche cinétique

Chez les glycosidases à rétention de configuration l'intermédiaire covalent formé peut être intercepté par un nucléophile (autre qu'une molécule d'eau) pour conduire à un nouveau glycoside de manière totalement stéréospécifique. La glycoside hydrolase devient alors une transglycosylase. (Toone, **1989**) Malgré la stéréospécificité de cette réaction, la transglycosylation présente l'inconvénient de former de nombreux sous-produits de réaction. En effet dès la formation de l'intermédiaire covalent 3 candidats potentiels peuvent attaquer le centre anomérique (Schéma 5) :

- i) un accepteur de glycoside nucléophile, conduisant au produit de transglycosylation recherché,
- ii) une molécule d'eau, pour former le produit d'hydrolyse,
- iii) le donneur de sucre de départ formant ainsi un produit d'auto-condensation.



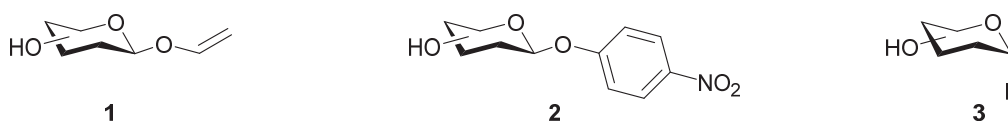
**Schéma 5** : Mécanisme de transglycosylation.

A la vue de ce schéma, force est de constater que le produit de transglycosylation peut dès sa formation devenir substrat de l'enzyme pour subir une nouvelle transglycosylation ou une hydrolyse (secondaire). La réaction de transglycosylation est donc une réaction sous contrôle cinétique puisqu'en fonction des conditions de réaction le produit qui se forme le plus

rapidement sera favorisé. Durant l'acte catalytique, à mesure que l'accepteur de glycoside est consommé, la concentration en produit augmente jusqu'à atteindre son maximum lorsque l'équilibre entre les vitesses  $r_s$  (vitesse de synthèse) et  $r_{-s}$  (vitesse de déalkylation) est atteint. Lors de cet équilibre, la réaction n'est plus sous contrôle cinétique mais thermodynamique puisque le produit de transglycosylation (produit cinétique) devient substrat pour l'hydrolyse (produit thermodynamique). Ainsi pour assurer une transglycosylation efficace il est nécessaire de régler la balance entre l'hydrolyse du donneur, la consommation de l'accepteur et l'hydrolyse du produit formé. Les principes énoncés précédemment pour l'hydrolyse inverse peuvent également s'appliquer aux transglycosylases, augmenter la concentration en donneur / accepteur, (Mitsuo, **1984**) diminuer la concentration en eau, éliminer le produit à mesure de sa formation. Par ailleurs, 2 autres paramètres clés sont à prendre en compte pour favoriser de bons rendements de transglycosylation :

- i) Le donneur de glycosyle.

Pour assurer un contrôle cinétique le donneur de sucre doit réagir rapidement avec l'enzyme pour diminuer le temps de réaction et donc limiter l'hydrolyse du produit de transglycosylation. On parle alors de donneur activé. De la même manière le donneur doit avoir une meilleure affinité (faible  $K_M$ ) pour l'enzyme que le produit synthétisé et ainsi favoriser son accumulation dans le milieu réactionnel. Le donneur de glycosyle activé devra donc être caractérisé par un ratio  $k_{cat}/K_M$  supérieur à celui du produit de transglycosylation. De nombreuses études ont démontré que les donneurs de glycosyles présentant des groupes partant de type vinylique (**1**, Figure 4) (Chiffolleau-Giraud, **1997**) ou benzylique (**2**) (Woudenberg-van Oosterom, **1998**) (tels de que les *ortho/para*-nitrophénol ou le dinitrophénol) ou encore les donneurs fluorés (**3**) (Kobayashi, **1991**; Williams, **2000**) permettaient d'obtenir de meilleurs rendements de transglycosylation que des donneurs non-activés (mono ou disaccharides). Les donneurs activés par la présence d'un groupement fluoro, présentent l'avantage d'être peu sensibles à l'hydrolyse spontanée en milieu aqueux et de ne générer qu'un ion fluorure qui peut être facilement séparé du produit de transglycosylation.



**Figure 4** : Structures de donneurs activés.

- ii) La sélectivité enzymatique.

Le rendement de transglycosylation repose sur la compétition entre les deux entités nucléophiles (alcool,  $H_2O$ ) pouvant agir sur l'intermédiaire covalent, mais aussi sur les propriétés intrinsèques de l'enzyme vis-à-vis de ces deux accepteurs. En conséquence, si la sélectivité de l'enzyme pour l'alcool est bonne (déplacement de l'équilibre vers la formation du produit), le taux de transglycosylation sera supérieur à celui de l'hydrolyse. D'autre part, comme énoncé précédemment, quand le temps de réaction devient trop long le produit de transglycosylation formé peut subir une hydrolyse dite secondaire qui peut être négative pour le rendement global. En 1999, Rantwijk *et coll.* ont proposé de quantifier la sélectivité



enzymatique en introduisant les facteurs de sélectivité ( $S$  et  $S_c$ ) comme fonctions du ratio des vitesses synthèse/hydrolyse ( $r_S/r_H$ ). (van Rantwijk, **1999**) Ces deux paramètres thermodynamiques peuvent être mesurés expérimentalement en fonction des constantes d'activités de l'alcool  $a_{ROH}$  et de l'eau  $a_w$  ( $S$ ) ou de leurs concentrations ( $S_c$ ) (Figure 5) :

$$\frac{r_S}{r_H} = S \frac{a_{ROH}}{a_w} \quad \frac{r_S}{r_H} = S_c \frac{[alcool]}{[eau]}$$

**Figure 5** : Facteurs de sélectivité enzymatique.

En prenant en compte ces deux équations il serait une fois de plus tentant de penser que l'activité de transglycosylation pourrait être exacerbée en utilisant un milieu réactionnel de faible concentration en eau. Or des études font référence à l'influence de la constante  $a_w$  sur l'activité de transglycosylation d'un panel de glycosidases. Les mesures réalisées avec des alcools miscibles ou non (éthanol, hexanol) à l'eau ont révélé l'augmentation globale de la sélectivité de l'enzyme avec l'augmentation de la concentration en eau. (Hansson, **2001**; van Rantwijk, **1999**) En 2008 les travaux effectués par Mladenoska *et coll.* sur 3  $\beta$ -galactosidases d'origines différentes ont confirmé l'influence de la constante  $a_w$  sur l'efficacité de la glycosylation. Cette sélectivité, fonction de la concentration en eau, peut alors être rapportée à la flexibilité de l'enzyme dans le milieu réactionnel. Plus le milieu est riche en eau, plus l'enzyme est « flexible » (maximum atteint pour  $a_w = 1$ ) et donc agira en faveur de la réaction de transglycosylation. A l'inverse, plus le milieu sera appauvri en eau, plus le catalyseur sera rigide et favorisera la réaction de petits nucléophiles (tels qu'une molécule d'eau) pour cliver l'intermédiaire covalent. (Mladenoska, **2008**) En suivant ce concept de « flexibilité » de l'enzyme, le même constat a été établi selon la structure des accepteurs de nucléophile : l'activité de transglycosylation est favorisée dans le cas d'alcools de faible longueur de chaîne, que celle-ci soit linéaire ou branchée. (Chahid, **1992**; Mladenoska, **2008**)

Les études réalisées sur les glycosidases ont permis de mettre en évidence leur potentiel de synthèse, comme une alternative aux procédés de chimie classique, pour la production de glycosides simples ou oligosaccharides par hydrolyse inverse ou transglycosylation. Néanmoins, leur activité de synthèse reste limitée par deux facteurs : un faible rendement de transglycosylation et la génération du produit d'hydrolyse. Les nombreuses stratégies explorées pour tenter de déplacer l'équilibre de la réaction vers la formation du dérivé glycosylé ont montré quelques limites et ont poussé les scientifiques à orienter leurs recherches vers d'autres alternatives. A la fin du XX<sup>ème</sup> siècle, l'avènement des outils de biologie moléculaire a favorisé le développement de glycosidases « à façon » présentant des activités de transglycosylation améliorées sur la base de l'ingénierie rationnelle ou combinatoire des protéines.

## 2.2. Les glycosynthases

Pendant de nombreuses années, l'usage des enzymes en synthèse organique a été peu exploité du fait du coût élevé de leur production mais également des faibles performances démontrées par les enzymes de type sauvage ou les mixtures enzymatiques provenant

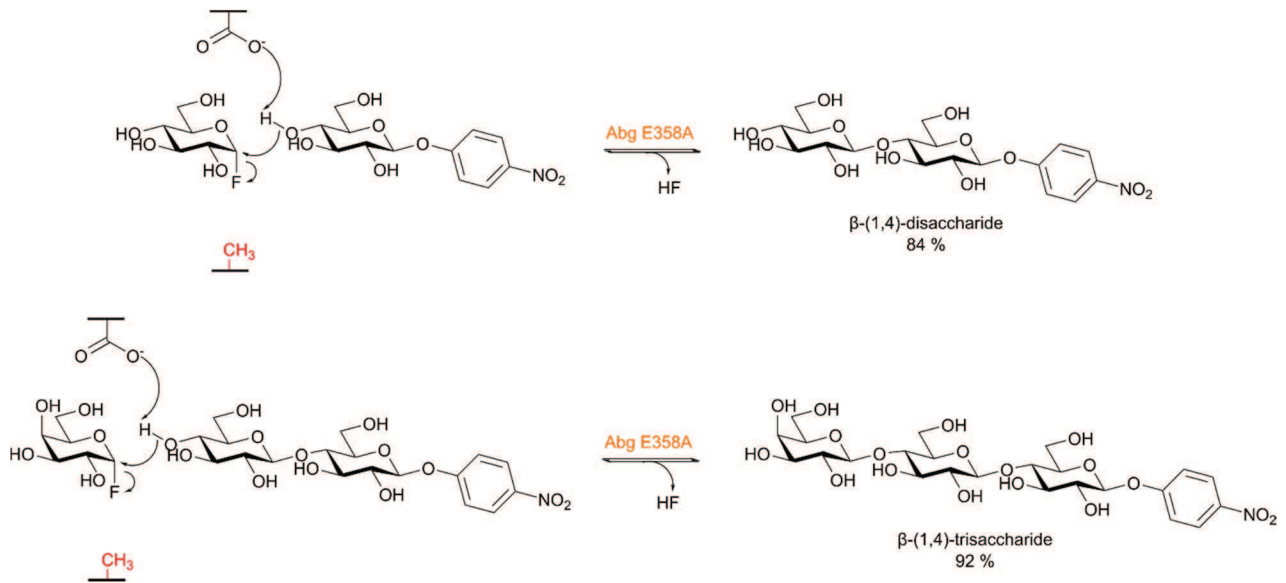


d'extraits de plantes ou micro-organismes (où la protéine active était dans la plupart des cas non identifiée).(Strohmeier, **2011**) Dans les années 70 l'apparition de la technologie de l'ADN recombinant a permis de reconsidérer la place des biocatalyseurs en chimie de synthèse. La manipulation des séquences d'ADN est devenue plus aisée et les enzymes naturelles peuvent désormais être modifiées spécifiquement pour s'approcher de l'activité désirée. L'utilisation de la machinerie cellulaire d'organismes vivants appelés hôtes d'expression (tels que les bactéries, levures, champignons, cellules de mammifères,...) permet la production de ces protéines recombinantes en grande quantité et à faible coût.(Arnold, **2001**)

Par la suite, le couplage de cette méthodologie aux outils de génie génétique a permis d'améliorer les connaissances sur les gènes et les protéines. A partir de micro-organismes contenant l'activité enzymatique recherchée, les chercheurs se sont appliqués à la production de banques de gènes mutés par exposition aux radiations UV, utilisation d'agents mutagènes ou introduction d'erreurs par les ADN polymérases. Le criblage d'activité de ces organismes mutants, permet alors d'identifier les gènes modifiés et leur séquençage permet de déterminer précisément la nature de la modification.(Strohmeier, **2011**) En 1979 le Prof. Smith développa la technique de mutagenèse dirigée, permettant l'introduction *in-vitro* de points de mutation spécifiques sur une séquence d'ADN.(Gillam, **1979**) Cette importante découverte a rendu alors possible la production de protéines présentant de nouvelles propriétés par mutation de résidus spécifiques sur la séquence d'ADN correspondante.

A l'heure actuelle l'amélioration ou la modification de l'activité d'une protéine par ingénierie rationnelle repose sur l'étude approfondie du mécanisme catalytique et de la structure tridimensionnelle de cette dernière.(Strohmeier, **2011**) Cette méthodologie appliquée aux glycosidases a conduit de manière générale à la modification des résidus catalytiques ou des résidus de la seconde sphère de la protéine pour réduire l'activité d'hydrolyse en faveur de l'auto-condensation ou de la transglycosylation.

C'est en 1998, à partir de ces principes d'ingénierie rationnelle, que l'équipe du Prof. Withers a fait une importante découverte en proposant la première transglycosylase totalement dépourvue d'activité d'hydrolyse ce qui donna lieu au concept de **glycosynthase**.(Mackenzie, **1998**) Ils ont alors démontré que le remplacement du résidu catalytique carboxylate de l'exo- $\beta$ -glucosidase d'*Agrobacterium sp* (Abg) par un acide aminé non nucléophile conduit à la génération d'un mutant inactif Abg E358A. En revanche, en présence de donneurs activés, en général un  $\alpha$ -glycoside fluoré, l'enzyme Abg E358A est capable du transfert de l'unité glycosidique vers une variété d'accepteurs sans hydrolyse de la liaison nouvellement formée. Les expériences réalisées sur Abg E358A à partir d' $\alpha$ -glucoside ou  $\alpha$ -galactoside fluoré sur une variété d'accepteurs, tels que les glycosides d'aryle, ont conduit à la synthèse de di- et tri-saccharides avec d'excellents rendements (jusqu'à 80 % et 90 % pour la glucosynthase et galactosynthase respectivement, Schéma 6) tandis que des résultats encourageants ont été obtenus *via* l'activité glucosynthase d'Abg E358A pour la synthèse d'oligosaccharides de plus haut degré de polymérisation (60 %).(Mackenzie, **1998**)



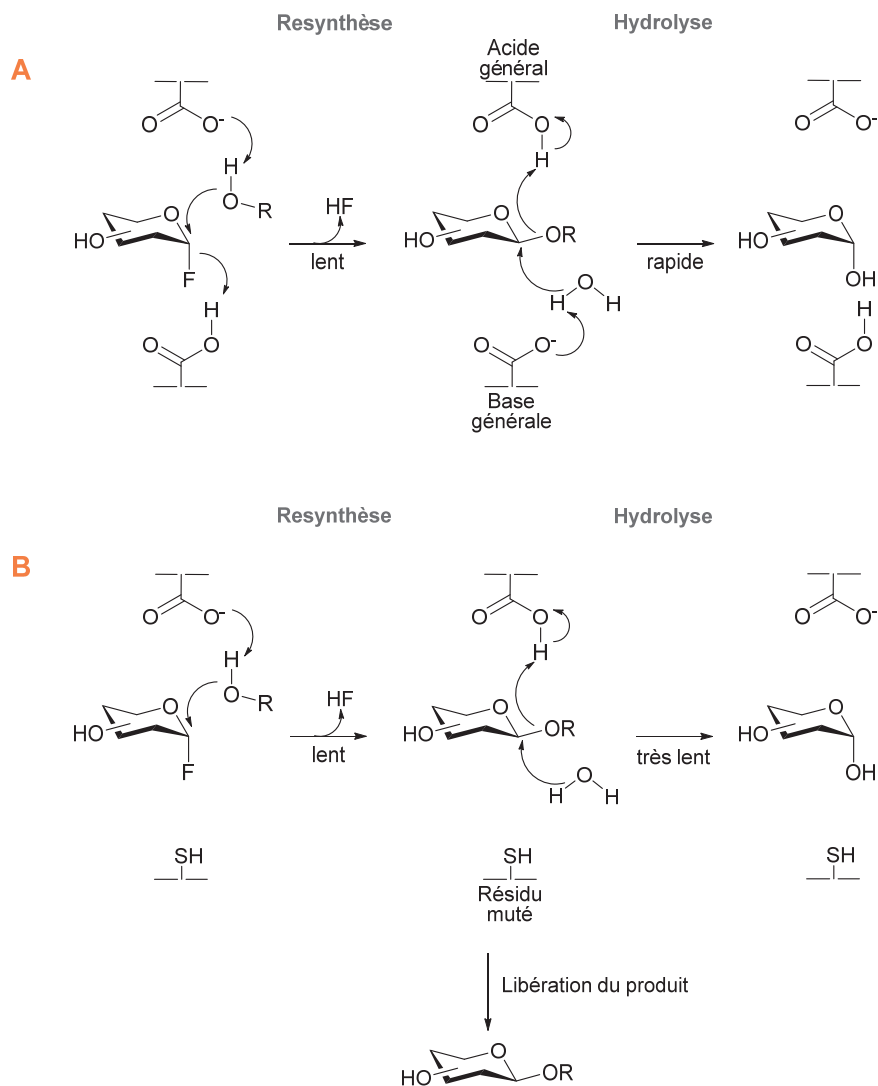
**Schéma 6 :** Synthèses d'oligosaccharides catalysées par la glycosynthase Abg E358A.

En cherchant à améliorer les propriétés de transglycosylase d'Abg, il a été démontré que le remplacement du Glutamate 358 par un résidu polaire, telle que la Sérine, permettait d'augmenter les taux de réaction (jusqu'à 24 fois supérieur par rapport au mutant E358A). L'amélioration de l'activité catalytique du mutant Abg E358S observée serait alors due à la présence de la chaîne latérale polaire de la Sérine, permettant une meilleure interaction avec l'atome de fluor anomérique et donc favorisant son départ. (Mayer, **2000**) Parallèlement Malet *et coll.* ont utilisé ce même concept pour la génération de la première endo-glycosynthase à partir de la β-(1,3)-(1,4)-glucanase de *Bacillus licheniformis*. (Malet, **1998**) Depuis ces travaux pionniers de nombreux exemples de glycosynthases ont été décrits et font l'objet de nombreuses revues. (Armstrong, **2013**; Cobucci-Ponzano, **2012b**; Hancock, **2006**; Kittl, **2010**; Perugino, **2005**; Perugino, **2004**)

Si les glycosynthases décrites précédemment proviennent toutes de glycosidases à rétention de configuration, la littérature recense quelques exemples de la génération de glycosynthase à inversion de configuration. En 2006 Honda *et coll.* ont produit la première endo-xylosynthase par mutation du résidu base générale de l'exo-oligoxylanase (Rex) de *Bacillus halodurans*. Le mutant Rex E263C est alors capable de la synthèse de xylo-oligosaccharides avec de très bons rendements. (Honda, **2006**) Deux autres exemples ont par la suite été décrits à partir d'endo-glycosidases, l'α-(1,2)-fucosynthase de *Bifidobacterium bifidum* et la glycosynthase dérivée de la chitinase de *Bryum coronatum*. (Ohnuma, **2012**; Wada, **2008**)

Chez les glycosidases à inversion de configuration deux acides aminés jouant le rôle d'acide et de base générale sont impliqués dans le mécanisme catalytique qui ne passe pas par un intermédiaire covalent glycosyl-enzyme. Pour développer les mutants glycosynthases correspondants, les chercheurs se sont basés sur les travaux du Prof. Hehre qui a démontré la capacité des glycosidases à inversion de configuration d'hydrolyser des donneurs de glycosyle fluorés présentant une configuration anomérique identique à celle du produit

d'hydrolyse final. Le mécanisme, appelé mécanisme de resynthèse-hydrolyse de Hehre, (Hehre, **1979**) met en jeu une première étape de transglycosylation (resynthèse) du donneur fluoré sur un accepteur de glycoside avec inversion de la configuration anomérique. Ensuite, la liaison glycosidique nouvellement formée subit rapidement une réaction d'hydrolyse avant de pouvoir quitter le site actif de l'enzyme (**A**, Schéma 7). La génération du mutant glycosynthase repose alors sur la mutation du résidu base générale (qui joue le rôle du donneur de proton lors de l'étape de resynthèse) et sur l'utilisation d'un donneur fluoré activé qui ne nécessite pas cette assistance acide pour faciliter le départ du groupe partant. L'activité hydrolytique de l'enzyme étant considérablement ralentie par la perte du résidu base générale, le produit de transglycosylation alors formé peut quitter le site actif avant de subir l'hydrolyse (**B**, Schéma 7).

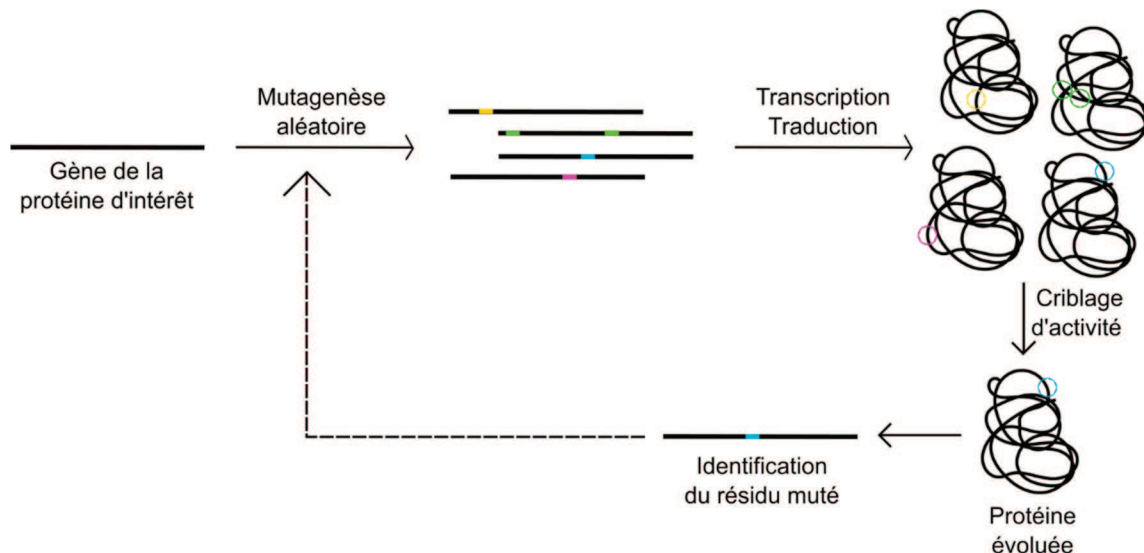


**Schéma 7 : (A) Mécanisme de « resynthèse-hydrolyse » de Hehre ; (B) mécanisme des glycosynthases issues de glycosidases à rétention de configuration.**

A l'heure actuelle l'ingénierie rationnelle des glycosidases a permis la génération d'une large gamme de glycosynthases (parmi 18 familles de la classification CAZy) pouvant être utilisées dans de nombreuses applications, dont la synthèse d'éliciteurs de défense chez les plantes, de glycosaminoglycanes ou d'oligosaccharides immunomodulateurs entre autre. (Cobucci-Ponzano, **2012a**) Néanmoins nos connaissances incomplètes sur la structure des protéines et l'incapacité de prédire les résidus clés ayant un impact sur la spécificité ou l'activité catalytique limite l'utilisation de l'ingénierie rationnelle pour le design de glycosynthases. (Ben-David, **2008**; Kwan, **2011**) Une seconde approche basée sur l'évolution naturelle des protéines peut être appliquée, sans connaissances mécanistiques ou structurelles préalables, et permet la génération de nouveaux biocatalyseurs aux propriétés enzymatiques améliorées.

### 2.3. Ingénierie combinatoire des glycoside hydrolases

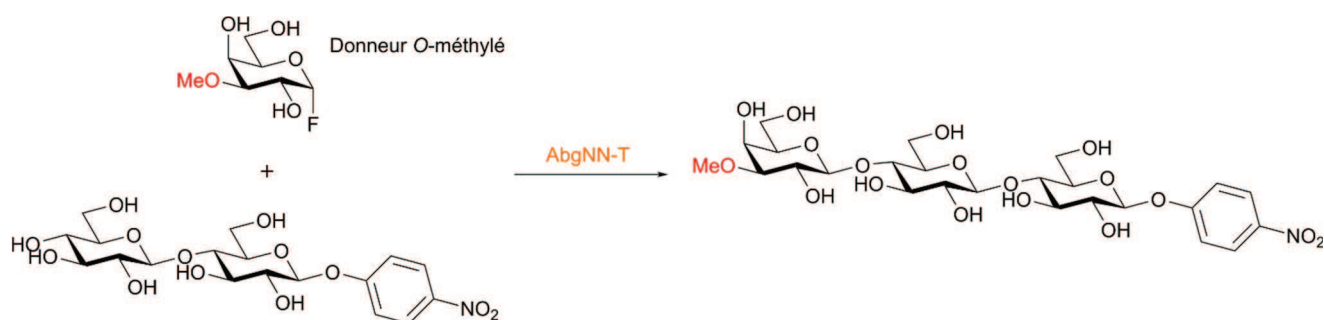
L'évolution dirigée (ou ingénierie combinatoire) permet de reproduire *in-vitro* et de manière accélérée l'évolution naturelle des protéines et d'adaptation à leur environnement en suivant un processus répétitif de mutagenèse / criblage / sélection. (Arnold, **1999**) Ainsi à partir d'un gène d'intérêt une banque de mutants est générée par mutagenèse aléatoire, des méthodes de criblages d'activité ou de sélection sont ensuite utilisées pour isoler les hits positifs de la librairie. La caractérisation des variants sélectionnés permet alors d'identifier les résidus clés qui améliorent les propriétés de la protéine. De manière générale, cette première génération de mutants est le point de départ d'une succession de cycles de mutagenèse / criblage ou sélection qui permet à terme d'affiner les performances de la protéine initiale (Schéma 8). La diversité génétique apportée par évolution dirigée implique la formation de banque comportant un grand nombre de variants. Elle nécessite donc l'utilisation de méthodes de criblage et de sélection adaptées qui permettront d'identifier de manière rapide et fiable le(s) mutant(s) présentant l'activité recherchée. (Cheng, **2015**; Lehmann, **2001**)



**Schéma 8 :** Stratégie d'évolution dirigée des protéines.

Les glycosynthases sont capables de catalyser la réaction de transglycosylation de manière efficace mais présentent généralement une activité plus lente que leurs analogues en type sauvage ce qui nécessite donc l'utilisation d'une grande quantité de mutant et/ou

l'augmentation du temps de réaction.(Mayer, **2000**) La stratégie de l'évolution dirigée appliquée aux glycosynthases a permis d'augmenter l'activité catalytique mais également d'améliorer la régiosélectivité de la réaction ainsi que leur versatilité vis-à-vis du substrat.(Feng, **2005**) Les premiers travaux réalisés sur la glycosynthase Abg E358G ont permis, après 2 cycles de mutagenèse, d'identifier un nouveau mutant glycosynthase Abg 2F6 présentant une activité catalytique 27 fois supérieure au variant parent Abg E358G. L'enzyme ainsi conçue présente également une meilleure spécificité vis-à-vis du substrat puisque des taux de réaction élevés ont été observés sur une large gamme de donneurs et d'accepteurs.(Kim, **2004**) La même stratégie appliquée au variant XynB2 (E335G) de la  $\beta$ -xylosidase de *B. stearothermophilus* a permis d'identifier un meilleur mutant glycosynthase dont le  $k_{cat}$  a été multiplié par 35 par rapport au variant initial.(Ben-David, **2008**) Plus récemment les travaux de Shim *et coll.* sur Abg de type sauvage et ses variants glycosynthases ont conduit à la génération du premier mutant glycosynthase capable de reconnaître les donneurs de glycosides O-méthylé en position C-3 du sucre. Ce mutant AbgNN-T est alors capable de synthétiser des oligosaccharides contenant une unité 3-méthyl-O-glucoside et galactoside par le transfert d'une seule unité sucre (Schéma 9) avec une activité glycosynthase 39 fois supérieure au mutant Abg 2F6 décrit précédemment. La conception de ce nouveau biocatalyseur s'apparente comme un outil précieux pour la préparation de glycoconjugués naturels (par exemple en reproduisant la réaction de méthylation post-traductionnelle) ou pour la synthèse de glycanes méthylés.(Shim, **2012**)



**Schéma 9** : Synthèse d'un oligosaccharide méthylé par AbgNN-T.

La méthodologie d'évolution dirigée a également été appliquée aux GH natives pour augmenter leur activité de transglycosylation. En effet parmi ces enzymes, toutes ne sont pas aptes à agir comme glycosynthases, plusieurs exemples de glycosidases mutées sur le résidu nucléophile ont été décrits dans la littérature sans présenter d'activité glycosynthase, telles que la  $\beta$ -xylosidase de *Thermosporum saccharolyticum*, la  $\beta$ -galactosidase de *Bacillus circulans*(Ducros, **2003**) ou encore la  $\beta$ -xyloxydase de *Sulfolobus solfataricus*.(Cobucci-Ponzano, **2003**) Les données collectées à l'heure actuelle ne permettent pas de comprendre l'origine de ce phénomène et quels éléments structuraux seraient requis pour la génération « en routine » de glycosynthases.(Ben-David, **2008**) Les glycosidases sont caractérisées par un ratio hydrolyse/transglycosylation ( $r_H/r_T$ ) qui dans la plupart du temps est en faveur de l'activité hydrolytique.(Bissaro, **2015**) Les chercheurs se sont donc appliqués à la génération de mutants glycosidases par évolution dirigée afin d'obtenir un ratio  $r_H/r_T$  favorable à la réaction

de transglycosylation. Plusieurs glycosidases présentant une forte activité transglycosidase ont été obtenues par ingénierie combinatoire, par exemple le variant Tt $\beta$ gly (A205T/F401S/H365Y) de la  $\beta$ -glucosidase de *T. thermophilus* présente une activité de transglycosylation 100 fois supérieure à celle d'hydrolyse.(Feng, **2005**) Plus récemment une approche semi-rationnelle (combinaison de l'ingénierie rationnelle et combinatoire) est utilisée pour générer des transglycosylases efficaces et flexibles vis-à-vis du substrat. Par exemple, les mutants  $\alpha$ -transglucosidases de l'amylosucrase de *Neisseria polysaccharea* et de la dextranase de *Leuconostoc mesenteroides* capables de reconnaître des accepteurs glycosylés non-naturels ont été appliqués à la synthèse chimio-enzymatique d'oligosaccharides anti-microbiens comme potentiels candidats vaccins.(Champion, **2009**; Champion, **2012**; Salamone, **2015**)

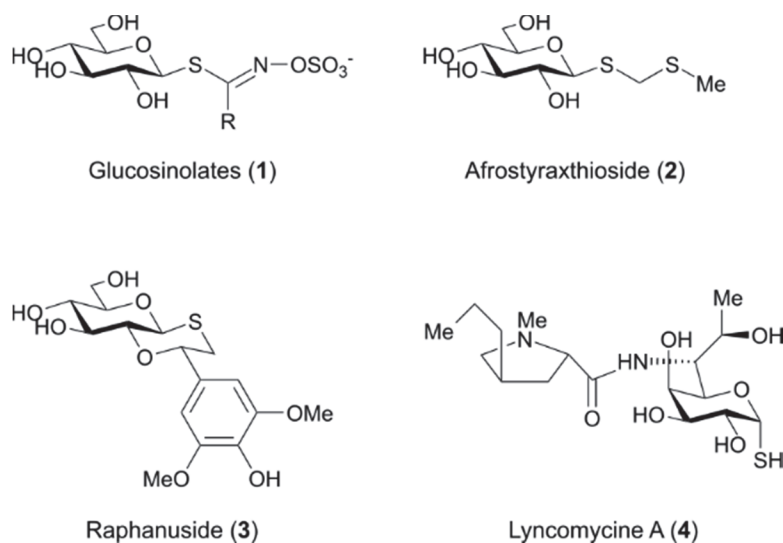
Ainsi les techniques de bio-ingénierie des glycosidases se révèlent donc être des outils puissants de glycodiversification en améliorant les capacités catalytiques des enzymes mais également en augmentant le panel de substrats reconnus.(Kwan, **2011**) Ces biocatalyseurs produits à façon sont capables de synthétiser efficacement de nombreux dérivés glycosylés, du glycoside simple au néoglycoconjugué, ou pour la synthèse d'oligosaccharides de faible à haut degré de polymérisation. Elles trouvent alors des applications dans de nombreux domaines, promesses de leur future utilisation à l'échelle industrielle en remplacement des procédés chimiques classiques. Outre les liaisons O-glycosidiques, les GH sont également capables de créer des liaisons S-glycosidiques pour la synthèse de thioglycosides, composés largement étudiés dans les glycosciences.

## **II. Les thioglycosides, briques élémentaires pour la synthèse d'oligosaccharides complexes**

### **1. Les thioglycosides naturels**

Dans la Nature les 1-thioglycosides, dont l'atome d'oxygène anomérique est remplacé par un atome de soufre, sont peu abondants en comparaison à leurs analogues en série O-. Ils appartiennent principalement à la famille des glucosinolates,  $\beta$ -thioglycosides *N*-hydroxysulfates (**1**, Figure 6), retrouvés largement chez les plantes de la famille des *Brassicacées* (ou crucifères), telles que les espèces du genre *Brassica* (la moutarde, le colza, divers choux, ...) ou encore dans l'organisme modèle *Arabidopsis thaliana*. A ce jour plus de 130 structures différentes ont été caractérisées chez les plantes.(Agerbirk, **2012**) Les autres formes naturelles de 1-thioglycosides comprennent l'afrostyraxthioside (**2**) d'*Afrotyrax lepidophyllus*,(Ngane, **2006**) le raphanuside (**3**) extraits des graines de *Raphanus sativus* L.(Yang, **2010**) et la lyncomycine A de *Streptomyces lincolnensis* reconnue et utilisée pour ses propriétés antibiotiques (**4**).(Herr, **1962**; Hoeksema, **1964**) Enfin, la découverte de protéines S-glycosylées naturelles a récemment été rapportée dans la littérature, ces dernières seront cependant décrites plus en détail dans la suite de cette introduction.



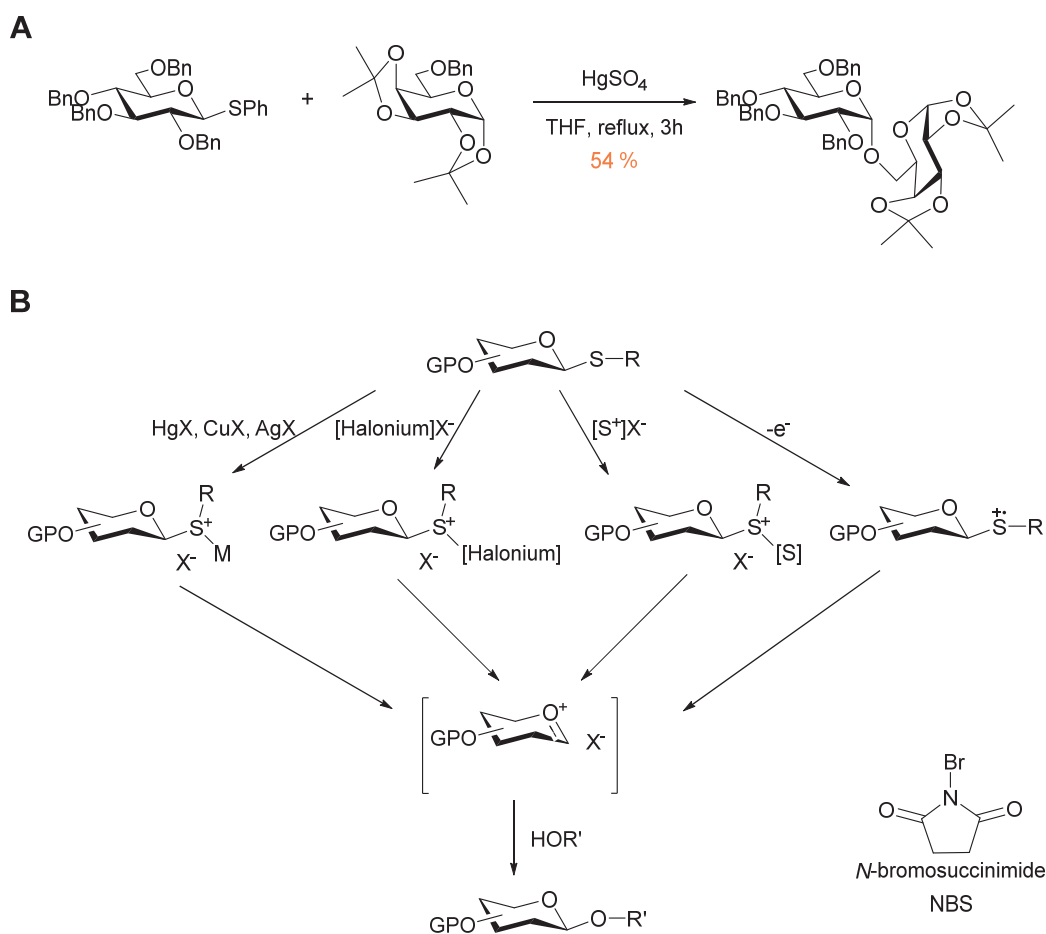


**Figure 6 :** Structures de thioglycosides naturels.

L'étude de ces composés naturels a inspiré les chercheurs et fait désormais l'objet de nombreux travaux. Solubles dans l'eau et tolérés par la plupart des systèmes biologiques, leurs avantages reposent sur l'adoption de conformations similaires aux analogues O-glycosylés mais aussi et surtout sur leur relative résistance vis-à-vis de l'hydrolyse chimique et enzymatique. (Witczak, **2005**) Grâce à leurs propriétés remarquables, ces composés sont donc devenus de précieux outils dans de nombreuses applications, : i) en synthèse organique comme donneurs de glycosides dans diverses réactions de glycosylation, ii) comme intermédiaires stables pour la résolution de structures de protéines en cristallographie (Driguez, **2001**) ou iii) en tant qu'agents thérapeutiques par la génération d'antibiotiques glycosylés (tels que les dérivés de la lyncomycine A) (Witczak, **1995**) ou d'inhibiteurs de glycosidases, impliquées dans de nombreuses maladies. (Wardrop, **2010**) De plus les thioglycosides sont également utilisés en synthèse chimique ou enzymatique comme précurseurs clés pour la génération de thiooligosaccharides complexes et de thioglycoconjugués. Ces glycomimétiques stables dans les milieux physiologiques sont alors utilisés à des fins thérapeutiques ou pour améliorer nos connaissances sur les fonctions biologiques des O-glycoconjugués (glycolipides et glycoprotéines). (Mandal, **2014**)

## 1. Applications en synthèse organique

Les thioglycosides sont largement utilisés en glycochimie dans les réactions de glycosylation comme donneur ou accepteur de glycosides. En 1973 Ferrier *et coll.* ont reporté la première synthèse d'un disaccharide à partir du tetrabenzyl -1- thioglucopyranose, activé par du sulfate de Mercure (**A**, Schéma 10).(Ferrier, **1973**) Depuis cette date de nombreux travaux ont été publiés dans la littérature et décrivent l'utilisation d'une variété de promoteurs permettant l'activation efficace de donneurs S-glycosylés pour la synthèse d'oligosaccharides. Ces promoteurs sont regroupés sous 4 grandes catégories, les sels de métaux, les réactifs de type halonium (*N*-bromosuccinimide par exemple), les réactifs soufrés ou encore les méthodes avec transfert d'électrons (**B**, Schéma 10).(Lian, **2015**)

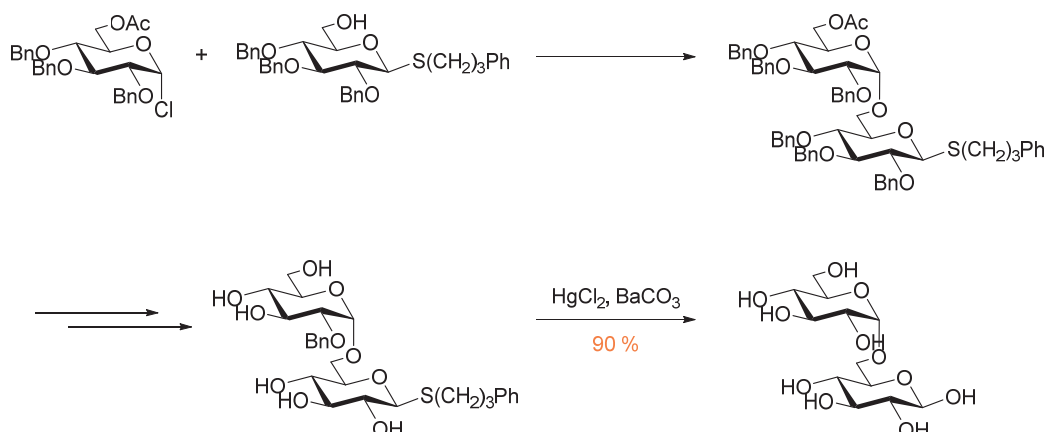


**Schéma 10 : (A) Synthèse d'un disaccharide à partir d'un donneur thioglycosylé ; (B) Classes principales de promoteurs pour la glycosylation de thioglycosides.**

Les thioglycosides sont également caractérisés par une stabilité vis-à-vis des conditions d'introduction et de clivage de groupements protecteurs inhérents à la chimie des sucres. Par exemple, pour décrire la synthèse de l'isomaltose Pfaffli *et coll.* ont utilisé une unité glucose protégée sur sa position anomérique par un groupement thioarylique (Schéma 11). Après formation du disaccharide protégé, le clivage des fonctions acyle et benzyle a pu être réalisé sans déprotéger l'hydroxyle anomérique, alors libéré dans un second temps.(Pfaffli, **1972**)



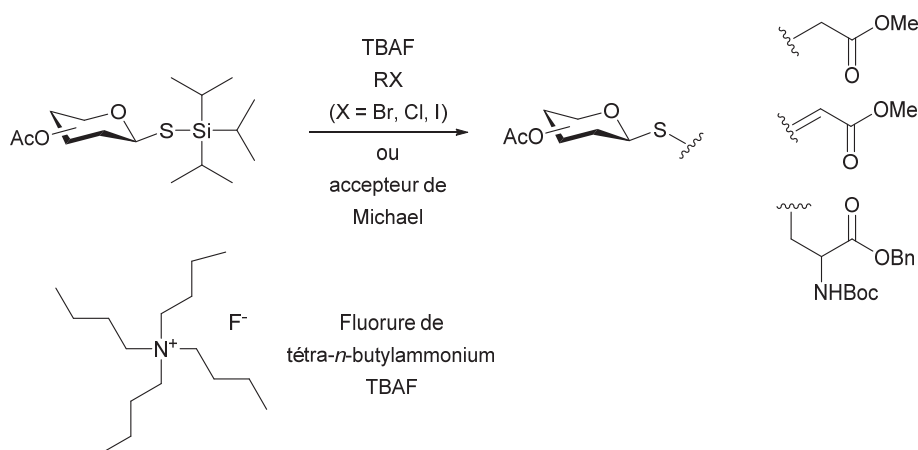
Cette synthèse est également le 1<sup>er</sup> exemple décrivant l'utilisation d'un thioglycoside comme accepteur dans la réaction de glycosylation.



**Schéma 11** : Synthèse de l'isomaltose à partir d'un accepteur thioglycosylé.

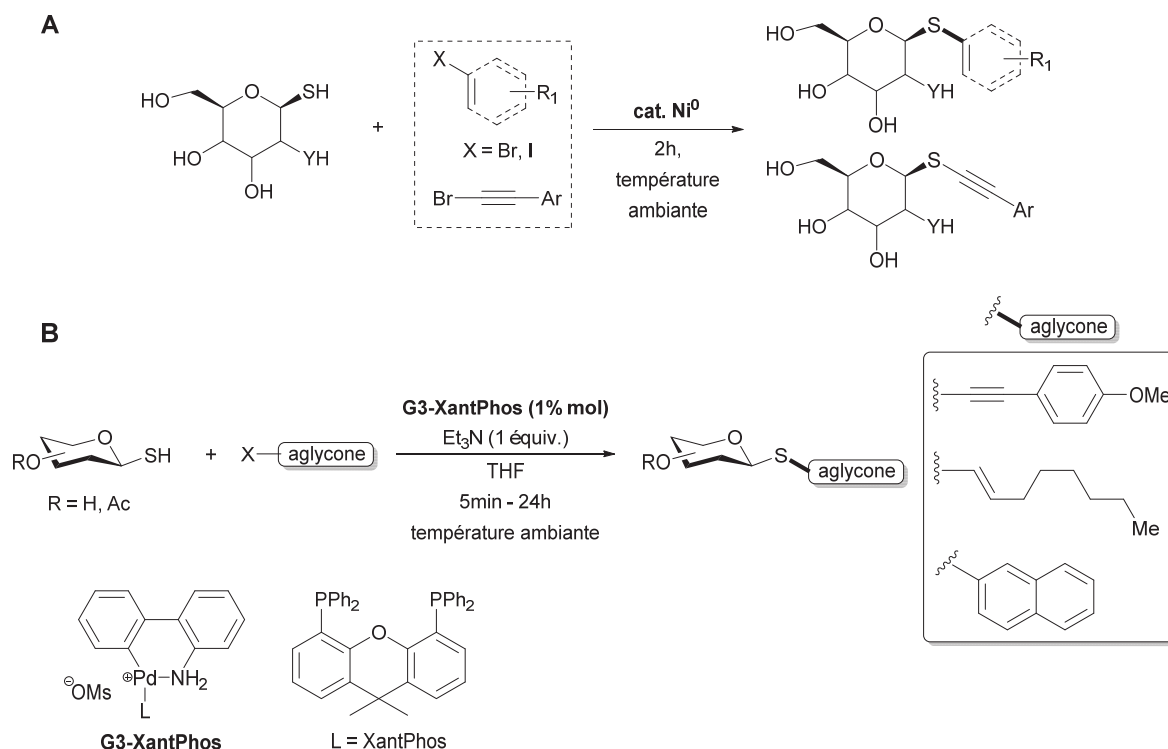
Les thioglycosides possèdent également l'avantage majeur d'être facile d'accès, classiquement par réaction d'un thiol soit sur un sucre per-acétylé *via* une activation par un acide de Lewis ( $\text{BF}_3 \cdot \text{Et}_2\text{O}$  par exemple) (Ferrier, **1976**) soit sur un glycoside halogéné par un déplacement de type  $\text{S}_\text{N}2$  (Pachamuthu, **2006**). De nombreuses voies de synthèses ont été décrites dans la littérature et utilisent les thioglycosides précurseurs élémentaires, *via* des réactions de type  $\text{S}_\text{N}2$ , ou d'addition type Michael, et permettent l'accès aux (1,3) / (1,4) et (1,6)-thiooligosaccharides de configuration  $\alpha$  ou  $\beta$  (Pachamuthu, **2006**). Ces composés sont des outils précieux en glycobiologie comme mimes d'oligosaccharides naturels pour l'étude des interactions glycanes / protéines (Driguez, **1997**).

Ces méthodes permettent également la fonctionnalisation des thioglycosides pour la synthèse de thioglycoconjugués. A l'heure actuelle les efforts se concentrent cependant sur la génération de systèmes favorisant l'accès à un large panel de thioglycosides à partir d'un même donneur. En 2014 Mandal *et coll.* ont décrit la synthèse rapide et efficace de thioglycosides fonctionnalisés en une étape à partir de tri-isopropylsilylthio-glycosides (TIPS thio-glycosides) (Mandal, **2014**). La liaison S-Si facilement clivable en présence d'une source de fluor (TBAF) permet l'addition nucléophile du thiolate sur des accepteurs électrophiles (halogénures d'alkyle, bromures de glycosyle ou accepteurs de type Michael) conduisant aux thioglycosides correspondant avec d'excellents rendements (> 90 %) (Schéma 12). Des résultats similaires ont été obtenus par extension de leur concept à la synthèse de disaccharides ou à la glycosylation d'un acide aminé, la 4-iodo-L-Alanine.



**Schéma 12 :** Synthèse one-pot de thioglycosides à partir de TIPS-thioglycosides.

Plus récemment l'équipe du Dr. Messaoudi a exploité la réactivité en catalyse des métaux de transition, largement utilisés en chimie organique pour la génération de molécules fonctionnalisées, pour produire des thioglycosides présentant une grande diversité moléculaire.(Brachet, **2013a**) L'originalité de leur méthode repose sur le couplage direct d'accepteurs de type alkynyle, alkenyle et aryle sur des thioglycosides non protégés. La stratégie suivie repose alors sur l'utilisation de systèmes catalytiques accessibles (peu chers et faciles à produire) permettant l'introduction rapide de fonctions variées sous des conditions douces de réaction et facilitant la purification des produits attendus. Ainsi en 2013, l'équipe a rapporté pour la première fois la synthèse d'une banque de thioglycosides (de configuration  $\alpha$  ou  $\beta$ ) fonctionnalisés *via* une catalyse au Nickel en utilisant le système  $[\text{Ni}^0(\text{Bipy})_2(\text{Py})]$  dans le méthanol à température ambiante (**A**, Schéma 13).(Brachet, **2013b**) Si la méthode est efficace et stéréosélective elle nécessite néanmoins l'ajout d'une quantité importante de sels de Nickel (30 % mol), toxiques, qui limite son application à plus grande échelle. Pour contourner cet inconvénient l'équipe a par la suite décrit une méthodologie similaire *via* un couplage Pallado-catalysé. Les réactions sont alors réalisées à partir d'un mélange équimolaire donneur / accepteur en présence du complexe G3/XantPhos (1 % mol), généré *in-situ* à partir du précatalyseur Pd-G3, à température ambiante (**B**, Schéma 13).(Bruneau, **2015**) La réactivité catalytique du Palladium, même en très faible quantité, a permis la synthèse stéréosélective de monoglycosides thioalkynylé, thioalkenylé et thioarylé, et d'oligosides thioarylés avec de très bon rendements (40 -90 %). Ce concept innovant basé sur le couplage catalysé par les métaux de transition se présente donc comme une voie d'accès élégante et rapide à une variété de thioglycosides fonctionnalisés à partir de sucres non protégés.



**Schéma 13** : Fonctionnalisation de thioglycosides non-protégés par couplage catalytique *via* les métaux de transition.

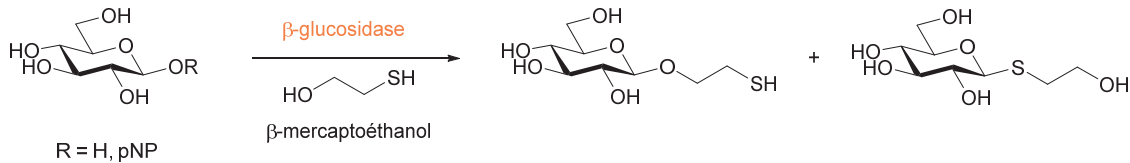
A l'heure actuelle de nombreuses méthodologies de synthèse de thioglycosides sont répertoriées ainsi que leurs applications pour la génération de thiooligosaccharides. (Driguez, **1997**; Pachamuthu, **2006**) Le potentiel des glycosidases a également été exploré dans ce domaine et a conduit à la génération de nouveaux biocatalyseurs spécifiques de la création de liaisons S-glycosidiques, les thioglycoligases et thioglycosynthases.

## 2. Synthèse de thioglycosides par voie enzymatique

### 2.1. Les glycoside hydrolases natives

Alors que la littérature recense de nombreux exemples décrivant la synthèse d'O-oligosaccharides et d'O-alkylglycosides par transglycosylation et hydrolyse inverse, (Crout, **1998**; Vocadlo, **2008**) seulement quelques travaux font référence à la synthèse de dérivés thioglycosylés suivant ces 2 méthodes. En 1994, Dintinger et Vic ont proposé pour la première fois l'accès par voie enzymatique aux mercaptoglycosides. Dans le but de produire de manière simple et rapide des glycoconjugués, les 2 groupes ont focalisé leurs travaux sur le couplage de bras espaceurs sur un motif sucre par transglycosylation et glycosylation directe respectivement. Vic *et coll.* ont orienté leurs études sur les capacités d'hydrolyse inverse d'une  $\beta$ -glucosidase extraite d'amande en utilisant différents accepteurs de glycosides, dont le 1,3-propanedithiol. (Vic, **1994**) La réaction enzymatique a permis de mettre en évidence la formation d'un dérivé 1-thioglycosidique avec un rendement faible mais comparable aux dérivés O-glycosylés (17%, 7-22% respectivement). Avec ce résultat prometteur Vic *et coll.* ont confirmé la possibilité d'obtenir des S-glycosides par voie enzymatique, au même titre que

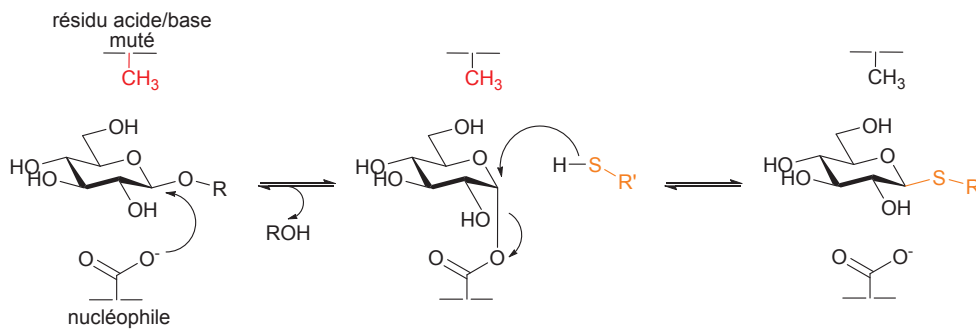
leurs analogues O-glycosylés. Dintinger *et coll.* ont quant à eux étudié le potentiel de O- et S-glycosylation d'une variété de glycosidases (une  $\beta$ -D-glucosidase extrait d'amande, une  $\alpha$ -D-mannosidase de *Jack bean* et une  $\beta$ -D-galactosidase de *Aspergillus oryzae*) avec le  $\beta$ -Mercaptoéthanol comme entité acceptrice de sucre (Schéma 14). (Dintinger, **1994**) Parmi les résultats obtenus, seule la glucosidase d'amande a présenté la double compétence de catalyser la synthèse de O- et S-glycosides. D'autre part les études cinétiques réalisées sur les 2 modes de glycosylation ont démontré que la liaison O-glycosidique était toujours favorisée par rapport à son analogue en série S- (glycosylation ou hydrolyse), confirmant ainsi le caractère inhibiteur des thioglycosides chez les glycosidases. (Debruyne, **1977**; Driguez, **2001**) Ces travaux pionniers ont permis de découvrir que le transfert de motif thiol simple sur des charpentes osidiques pouvait être réalisé par voie enzymatique.



**Schéma 14** : Synthèse du  $\beta$ -Mercaptoglucoside par une  $\beta$ -glucosidase native.

## 2.2. Les thioglycoligases et thioglycosynthèses

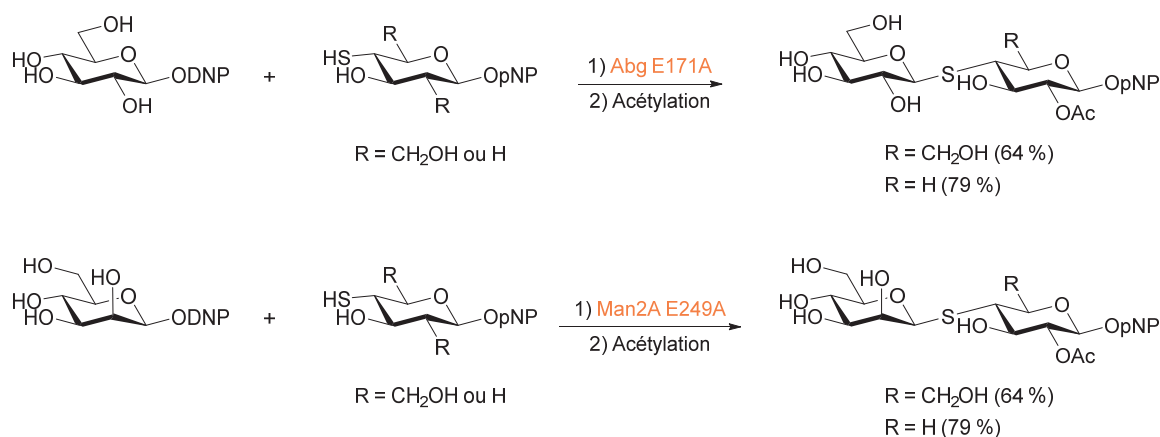
En 2003 l'équipe du Prof. Withers a été la première à explorer le potentiel de glycosidases mutées pour générer des liaisons S-glycosidiques, en introduisant le concept de **Thioglycoligase** (Schéma 15). (Jahn, **2003**)



**Schéma 15** : Mécanisme général des thioglycoligases.

En suivant la même stratégie que celle appliquée pour les glycosynthèses à rétention de configuration, ils ont décrit la production de deux  $\beta$ -glycosidases mutées, la  $\beta$ -glucosidase d'*A. sp* Abg E171A et la  $\beta$ -mannosidase de *C. fimi* Man2A E429A présentant un résidu Alanine non-ionique à la place du carboxylate catalytique jouant le rôle d'acide/base. Par conséquent, l'ingénierie de telles enzymes nécessite l'utilisation de donneurs et accepteurs appropriés pour assurer la réalisation de l'acte catalytique. Ainsi la formation de l'intermédiaire glycosyl-enzyme, 1<sup>ère</sup> étape du mécanisme des glycosidases à rétention de configuration, n'est assurée qu'en présence de donneurs activés, tels que les glycosides fluorés ou portant un groupement dinitrophényle (DNP), qui ne nécessitent pas de catalyse acide pour favoriser le

départ du groupe partant. Dans un second temps, l'absence du résidu acide/base implique la perte de la réactivité du nucléophile rentrant dans le site actif (tel que l'eau ou un alcool) au profit d'une espèce de nucléophilie plus forte (tel qu'un thiol) ne nécessitant pas d'activation basique pour cliver l'intermédiaire covalent. Les expériences de thioglycosylation réalisées avec Abg E171A et Man2A E429A ont permis la synthèse régio- et stéréosélective de  $\beta$ -(1,4)-thiodisaccharides avec des résultats prometteurs (35 – 82 % de rendement) à partir du DNP-glucose et DNP-mannose respectivement (Schéma 16).

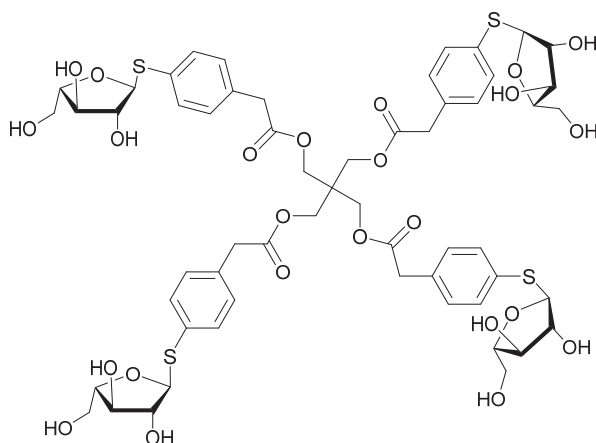


**Schéma 16** : 1<sup>er</sup> exemple de synthèse de thiodisaccharides par les thioglycoligases Abg E171A et Man2A E429A.

Pour améliorer les propriétés de thioglycoligases d'Abg l'équipe a par la suite étudié une banque de mutants obtenue par mutagenèse à saturation de site du résidu E171. Le screening de la librairie a permis d'isoler 5 variants positifs en activité thioligase, parmi lesquels Abg E171Q présentant une activité catalytique jusqu'à 100 fois supérieure à celle des autres mutants à partir du donneur N<sub>3</sub>Glc (1-azido- $\beta$ -D-glucose). (Mulleger, **2005**) Les expériences réalisées sur les mutants thioglycoligase de la  $\beta$ -glucuronidase thermostable de *T. maritima* TMGUA E383Q et E383A ont conduit aux mêmes observations. Le mutant Glutamine catalyse plus efficacement la réaction de thioglycosylation que le mutant Alanine en présence d'un donneur portant un groupe Nitrophénol. La génération de liaisons hydrogène entre le cycle aromatique du donneur de glycoside et la fonction amide de la Glutamine, impossibles dans le cas du mutant Alanine, permettraient alors d'assister le départ du groupe partant et donc d'accélérer la réaction de glycosylation. (Mulleger, **2006a**)

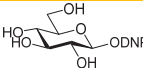
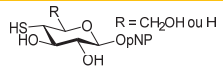
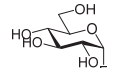
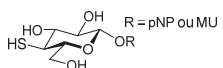
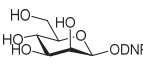
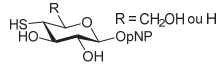
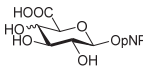
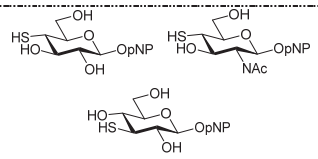
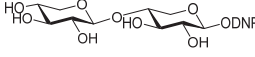
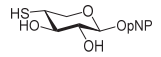
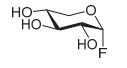
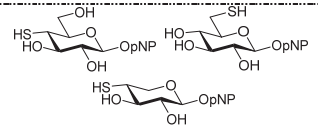
Depuis ces travaux pionniers, d'autres exemples de thioglycoligases ont été décrits dans la littérature, qui en compte à ce jour 8 (Tableau 1, p. 36 - 37), telles que la  $\beta$ -galactosidase BgaX de *Xanthomonas manihotis* capable de créer stéréosélectivement des liaisons  $\beta$ -(1,3) et  $\beta$ -(1,4)-thioglycosidiques. (Kim, **2006a**) Le criblage de la banque de thiodisaccharides synthétisés a permis par la suite l'identification d'un inhibiteur efficace de  $\beta$ -glucosidases. (Kim, **2007**) Récemment, Kim *et coll.* ont étendu la stratégie des thioligases aux  $\alpha$ -glycosidases et synthétisé pour la première fois des  $\alpha$ -thiodisaccharides à partir des mutants  $\alpha$ -xylosidases YicI D482A d'*E. coli* et MalA D416A de *S. solfataricus*. (Kim, **2006b**) En 2011, l'équipe du Prof.

Ferrières a exploré le potentiel des furanosidases en thioglycosylation. L'activité thioligase de l' $\alpha$ -L-arabinofuranosidase Araf51 E173A de *Clostridium thermocellum* a été testée à partir du 1-thioimidylarabinofuranose comme donneur de glycoside en présence d'une variété d'accepteur (thioarabinofuranoside, thioalkyl, thioaryl). Si l'enzyme ne permet pas la synthèse de thiodisaccharides, son potentiel en biocatalyse a été confirmé par la génération de thiofuranosides conjugués avec des rendements modestes (5 %) à plus prometteurs (50 - 98 %). (Almendros, **2011**) Plus particulièrement l'équipe a rapporté la synthèse d'un dendrimère thiofuranosidique par glycosylation simultanée des 4 fonctions thiol d'un dérivé du pentaérythritol catalysée par l'Araf51 E173A (Figure 7). Cet exemple unique à ce jour met en évidence le potentiel des thiofuranosidases pour la synthèse de mimes thérapeutiques de furanoglycosides. (Almendros, **2013**)

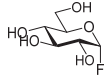
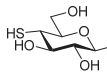
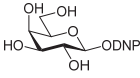
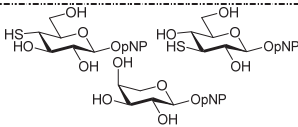
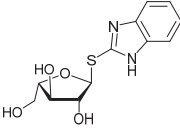
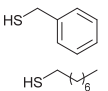
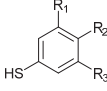


**Figure 7 :** Dendrimère du L-arabinofuranosyl-pentaérythritol.

Introduction générale

Enzyme	Famille GHs	Mutant/Activité	Donneur	Accepteur	Type de liaison	Réfs
<i>Agrobacterium sp</i> β-D-glucosidase (Abg)	GH1	E171 Thioglycoligase			β-S-(1,4)	(Jahn, <b>2003</b> )
		E171/E358 Thioglycosynthase			β-S-(1,4)	(Jahn, <b>2004</b> )
<i>Cellulomonas fimi</i> β-D-mannosidase (Man2A)	GH2	E249 Thioglycoligase			β-S-(1,4)	(Jahn, <b>2003</b> )
<i>Thermotoga maritima</i> β-D-glucuronidase (TMGUA)	GH3	E388 Thioglycoligase			β-S-(1,3) β-S-(1,4)	(Mullegger, <b>2006a</b> )
<i>Bacillus circulans</i> β-D-xylanase (Bcx)	GH11	E172 Thioglycoligase			β-S-(1,4)	(Armstrong, <b>2010</b> )
<i>E. coli</i> α-D-xylosidase (YicI)	GH31	D482 Thioglycoligase			α-S-(1,4) α-S-(1,6) (maj)	(Kim, <b>2006b</b> )

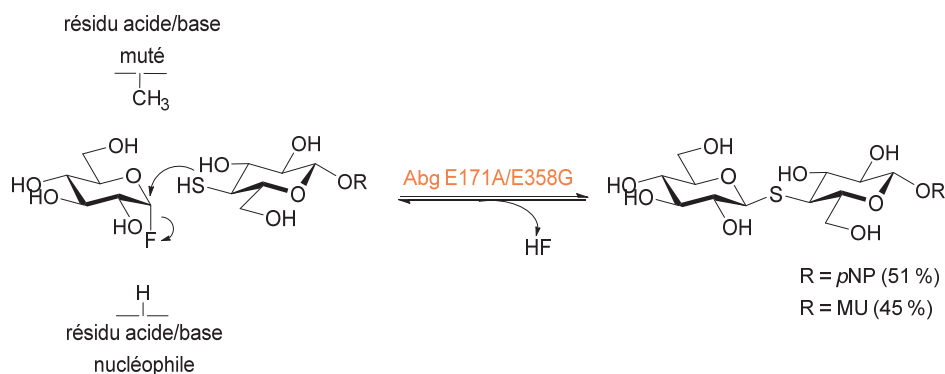
Introduction générale

Enzyme	Famille GHs	Mutant/Activité	Donneur	Accepteur	Type de liaison	Réfs
<i>Sulfolobus solfataricus</i> $\alpha$ -D-glucosidase (MalA)	GH31	D416 Thioglycoligase			$\alpha$ -S-(1,4)	(Kim, <b>2006b</b> )
<i>Xanthomonas manihotis</i> $\beta$ -D-galactosidase (BgAx)	GH35	E184 Thioglycoligase			$\beta$ -S-(1,3) $\beta$ -S-(1,4)	(Kim, <b>2007</b> ; Kim, <b>2006a</b> )
<i>Clostridium thermocellum</i> $\alpha$ -L-arabinofuranosidase (Araf51)	GH51	E173 Thioglycoligase		  $R_1 = H, OH$ $R_2 = H, OMe, Cl$ $R_3 = H, NH_2, Me, OMe, Cl, CO_2H$	liaison $\alpha$	(Almendros, <b>2011</b> ; Almendros, <b>2013</b> )

**Tableau 1** : Thioglycoligases et thioglycosynthases décrites dans la littérature à ce jour.



Enfin, l'équipe du Prof. Withers pionnière dans le design rationnel des glycosidases a mis au point une nouvelle classe de mutants, en combinant les deux stratégies de glycosynthase et thioglycoligase, alors appelée **Thioglycosynthase**. (Jahn, **2004**) Pour valider ce nouveau concept, un variant de la  $\beta$ -glucosidase modèle Abg d'*A.sp* a été généré en substituant les résidus nucléophile et acide/base par 2 acides aminés catalytiquement inactifs (Schéma 17). Le double mutant Abg E171A/E358G est alors capable de catalyser la réaction de thioglycosylation en présence du donneur activé 1-fluoro- $\alpha$ -D-glucopyranoside et des 2 accepteurs  $\alpha$ -thio- $\beta$ -D-glucopyranoside de pNitrophényle et Méthylumbelliféryle (MU) 4-déoxy-4-thio- $\beta$ -D-glucopyranoside avec des rendements n'excédant pas 50 %. D'autre part les analyses cinétiques ont révélé que la thioglycosynthase catalyse plus lentement le transfert de l'unité glycosyle que les mono-variants d'Abg, avec un  $k_{cat}$  100 à 400 fois inférieur par rapport aux constantes de vitesse des mutants glycosynthase E358G et thioglycoligase E171A respectivement.

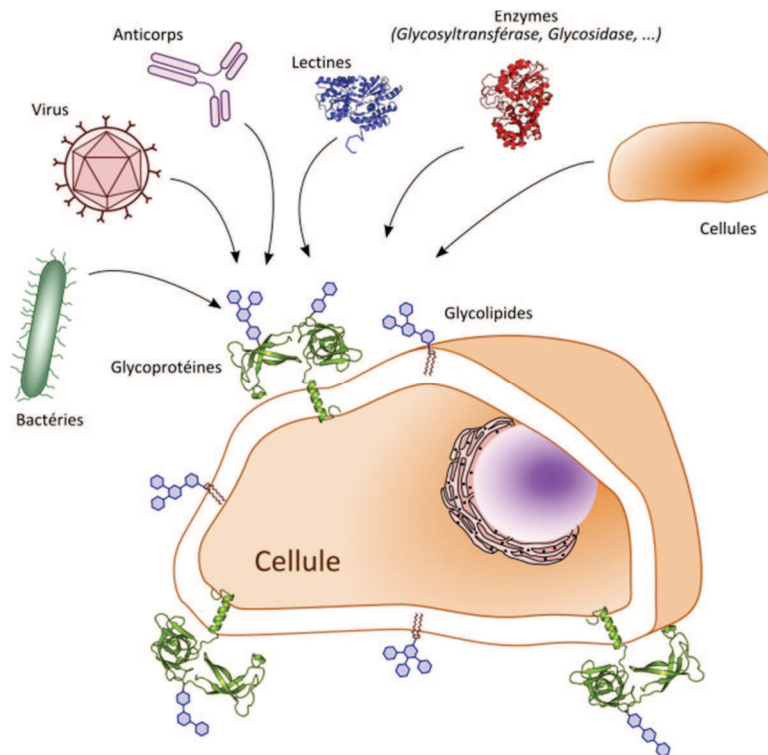


**Schéma 17 :** Synthèse de dithiosaccharides par le double mutant thioglycosynthase Abg E171A/E358G d'*A.sp*.

## II. La liaison thioglycosidique, forme rare chez les protéines glycosylées

### 1. Les sucres et les protéines, partenaires clés du monde du vivant

La glycosylation est un processus ubiquitaire considérée comme la modification co- et post-traductionnelle la plus complexe rencontrée chez les protéines et joue un rôle important dans de nombreux processus biologiques. (Varki, **1993**) Les fonctions de ces oligosaccharides qui recouvrent la surface des protéines sont multiples, ils assurent par exemple l'expression et le repliement correct des protéines mais également leur solubilité, ou encore favorise leur stabilité thermique et vis-à-vis de la protéolyse. Par conséquent, des erreurs de glycosylation peuvent entraîner des modifications importantes de ces protéines et conduire à l'apparition de pathologies, diabète, cancer, maladies neurodégénératives ou auto-immunes,... D'autre part les protéines membranaires sont les médiateurs de la communication cellulaire, les oligosaccharides, qui servent alors de récepteurs, sont déterminants dans les processus de virulence microbienne, d'inflammation ou encore de réponse immunitaire (Figure 8). (Varki, **1993**)



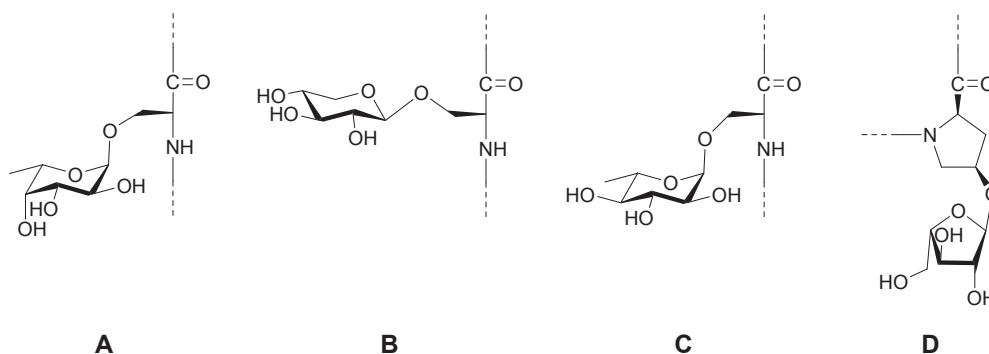
**Figure 8 :** Les glycanes, médiateurs de la communication cellulaire. (Lafite, **2012**)

La glycosylation est le résultat de l'activité d'une variété de GT et GH au sein de l'appareil de Golgi ou du Reticulum Endoplasmique, travaillant ensemble pour la biosynthèse de glycanes complexes reliés à un acide aminé (AA) spécifique. (Mrazek, **2013**) Les oligosaccharides sont majoritairement rattachés via l'oxygène  $O_Y$  d'une Sérine (Ser) ou Thréonine (Thr) (O-glycanes) ou via l'azote  $N_\delta$  d'une Asparagine (Asn) (N-glycanes). Parmi la variété de motifs sucres disponibles, 5 entités sont fréquemment retrouvées dans les séquences oligosaccharidiques, le glucose (Glc), le galactose (Gal), le mannose (Man), le N-acétylglucosamine (GlcNac) ou le N-acétylgalactosamine (GalNac). (Lafite, **2012**; Spiro, **2002**) Dans la Nature, les glycoprotéines sont biosynthétisées sous la forme de mixtures de protéines, ces dernières partageant le même squelette peptidique différent par la nature, le nombre et le site de glycosylation, aussi appelé glycoforme. Cependant, d'autres formes de glycosylation ont été répertoriées dans la Nature, mettant en jeu des motifs sucres ou lien glycosidiques exotiques qui suscitent de plus en plus d'intérêt parmi la communauté scientifique.

### 1.1. Les formes rares de protéines glycosylées

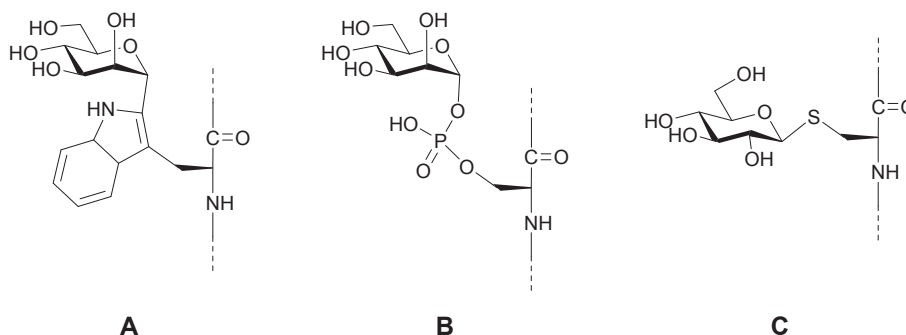
Comme développé ci-dessus, la glycosylation des protéines fait intervenir 2 types de liaisons principales, O-glycosidique et N-glycosidique, à partir des 5 motifs sucre récurrents Glc, Gal, Man, GlcNac et GalNac. D'autres types de glycosylation ont par la suite été identifiées dans la Nature, et varient dans l'identité de l'unité sucre couplée à l'AA de la chaîne peptidique ou dans la nature du lien sucre – AA. (Spiro, **2002**) Parmi les hexoses, le fucose a été le premier exemple de sucre non-usuel reporté sur une protéine de l'urine humaine (pro-urokinase) (A, Figure 9). Le xylose et le rhamnose, très abondants dans la Nature, sont

rarement impliqués dans le phénomène de glycosylation. Les O-xylosylations (**B**) et O-rhamnosylations (**C**) ont cependant été décrites sur les protéoglycanes (protéines de surface) de cellules mammifères et bactériennes respectivement. Des documents font également référence à la présence de motifs furanoses, tel que l' $\alpha$ -L-arabinofuranose (**D**) glycosylé aux protéoglycanes ou lectines de plantes (son de riz et pomme de terre respectivement). (Lafite, 2012)



**Figure 9** : Motifs sucre non-usuel des glycoprotéines. (**A**)  $\alpha$ -L-fucosylation, (**B**)  $\beta$ -D-xylosylation, (**C**)  $\alpha$ -L-rhamnosylation et (**D**)  $\alpha$ -L-arabinofuranosylation.

Parmi les liaisons glycosidiques exotiques, on retrouve les C-glycosylations qui sont le résultat d'une liaison covalente C-C entre une unité  $\alpha$ -mannosyle et le carbone indolique C-2 d'un résidu Tryptophane (Trp) (**A**, Figure 10). Les liaisons phosphoglycosidiques (**B**) sont également représentées et ont été identifiées majoritairement chez les parasites protozoaires (tels que les Leshmanies ou Tripanosomes). Enfin les liaisons S-glycosidiques font partie des modifications de protéines les moins connues (**C**). (Lafite, 2012) A l'heure actuelle la littérature recense 3 exemples de peptides naturels S-glycosylés. La création de telles liaisons glycosidiques est assurée par les S-glycosyltransférases (S-GT) correspondantes qui pourraient alors devenir des outils innovants et originaux pour la synthèse de protéines S-glycosylées. Malgré tout, peu d'études ont été réalisées sur ces biocatalyseurs naturels ou artificiels à ce jour.

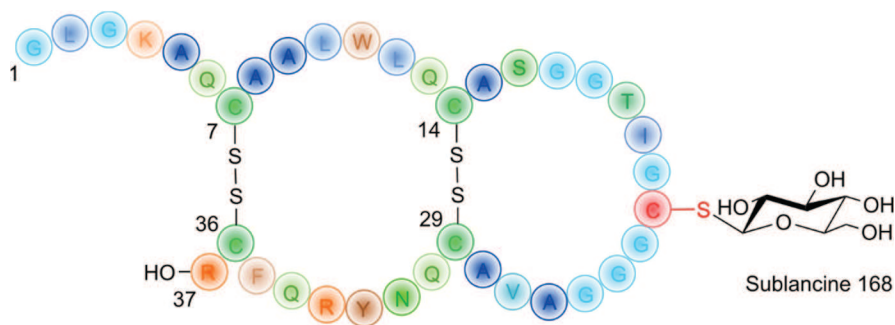


**Figure 10** : Liaisons sucre-acide aminé rares. (**A**) Trp C-mannosylé, (**B**) Ser phosphomannosylée et (**C**) Cys S-glucosylée.

### 1.2. Biosynthèse des thioglycoprotéines, les S-glycosyltransférases naturelles

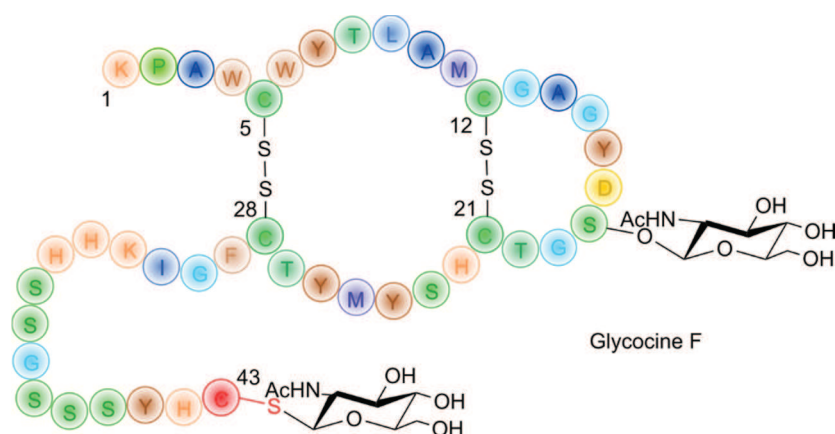
D'un point de vue historique, Lote *et coll.* ont reporté pour la première fois la glycosylation d'un résidu Cystéine de 2 peptides à partir d'extraits humains. Ainsi dans l'urine ils ont identifié un octapeptide S-glycosylé *via* un motif digalactose, (Lote, **1971a**; Lote, **1971b**) puis un décapeptide de motif similaire portant un triglucoside a été caractérisé à partir de membranes d'érythrocytes. (Weiss, **1971**) Malgré l'intérêt de ces études, elles ont été pendant de nombreuses années sujettes à controverses faute de nouveaux exemples de peptides S-glycosylés identifiés. (Sardzik, **2011**) C'est en 2011, et plus récemment en 2014, que la présence de S-glycosylations dans la Nature a été confirmée grâce à 3 études indépendantes faisant référence à la découverte de nouveaux thioglycopeptides aux propriétés antimicrobiennes, la sublancine, le glycocine F, et le thurandacine.

Le glycopeptide sublancine produit par *Bacillus subtilis* 168 est constitué d'un enchaînement de 37 acides aminés dont un résidu Cystéine est S-glucosylé (Figure 11). (Oman, **2011**) La S-glycosyltransférase SunS (appartenant à la famille des GT2) a été identifiée, exprimée chez *E.coli* et caractérisée comme catalyseur efficace du transfert du motif glucose de l'UDP- $\alpha$ -D-glucose sur le résidu Cystéine. Si les GT sont reconnues pour leur étroite spécificité vis-à-vis du substrat, SunS présente quant à elle une relative flexibilité dans la reconnaissance du nucléotide-sucre. En effet Wang *et coll.* ont mis en évidence la synthèse de divers analogues glycosylés à partir d'UDP- $\alpha$ -D-N-acétylglucosamine, UDP- $\alpha$ -D-galactose, GDP- $\alpha$ -D-mannose et d'UDP- $\alpha$ -D-xylose comme donneurs activés. Suite à ces résultats prometteurs, l'exploitation de SunS pour la biosynthèse d'analogues peptidiques de la sublancine a été approfondie en s'appuyant sur des études de relation structure-activité de SunS en fonction du peptide substrat. (Wang, **2011**)



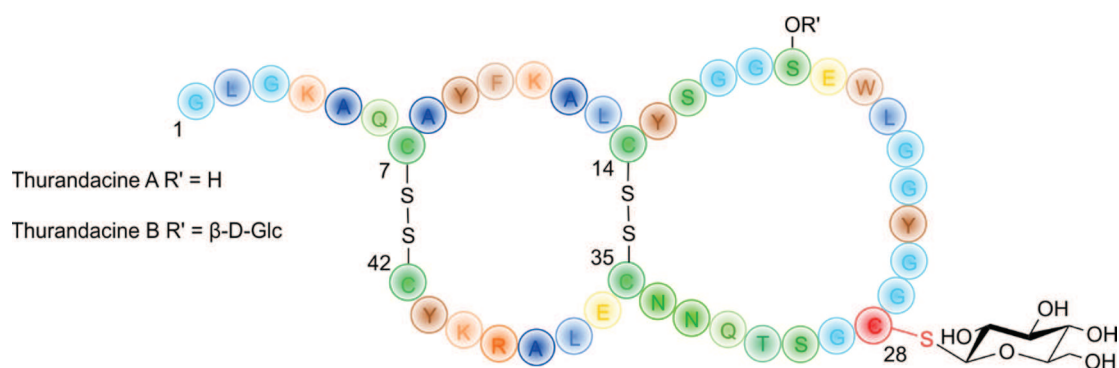
**Figure 11 :** Structure de la sublancine.

La caractérisation du bactériocine glycocin F, sécrété chez *Lactobacillus plantarum* KW30, a permis l'identification d'une liaison thioglycosidique non usuelle entre un motif N-acétylglucosamine et un résidu Cystéine de la chaîne peptidique (Figure 12). (Stepper, **2011**) Des études structurales réalisées par Résonance Magnétique Nucléaire (RMN) ont confirmé l'existence de 2 unités N-acétylglucosamine, l'une fixée sur l'atome d'oxygène d'une Sérine, la seconde sur l'atome de soufre d'une Cystéine. (Venugopal, **2011**)



**Figure 12 :** Structure du glycicine F.

En 2014, Wang *et coll.* ont mis en évidence la voie de biosynthèse des thurandacines A et B produits chez *Bacillus thuringiensis* serovar *andalousiensis* BGSC 4AW1. A partir d'un cluster de gènes présentant de grandes similitudes avec ceux impliqués dans les voies de biosynthèse de la sublancine et du glycicine F, ils ont ainsi pu identifier le gène de la glycosyltransférase ThuS. Les glycopeptides thurandacine sont constitués d'une séquence de 42 acides aminés, la forme A est monoglucosylée sur le résidu Cystéine 28 tandis que la forme B présente 2 sites de glucosylation, sur le résidu Cystéine 28 et le résidu Sérine19 (Figure 13). La transférase ThuS représente ainsi le premier catalyseur de O- et S-glycosylation de peptides *in vitro*.(Wang, 2014)



**Figure 13 :** Structure de la thurandacine A et B.

L'étude des glycoprotéines est devenue d'un intérêt majeur dans le domaine de la glycobiologie pour tenter de comprendre les relations de structure/activité inhérentes à ces molécules et permettre le développement de nouveaux traitements thérapeutiques.(Gamblin, 2009) L'inconvénient majeur de tels travaux repose sur l'obtention de formes homogènes de glycoprotéines. A partir d'extraits naturels leur isolement reste difficile, le challenge à l'heure actuelle porte donc sur la mise au point de méthodologies de synthèse permettant la glycosylation spécifique de protéines à partir de glycanes bien définis.(Pachamuthu, 2006) Dans cette optique les efforts se sont tournés vers la génération de thioglycoprotéines synthétiques comme modèles d'étude. En effet, la relative faible abondance de l'atome de

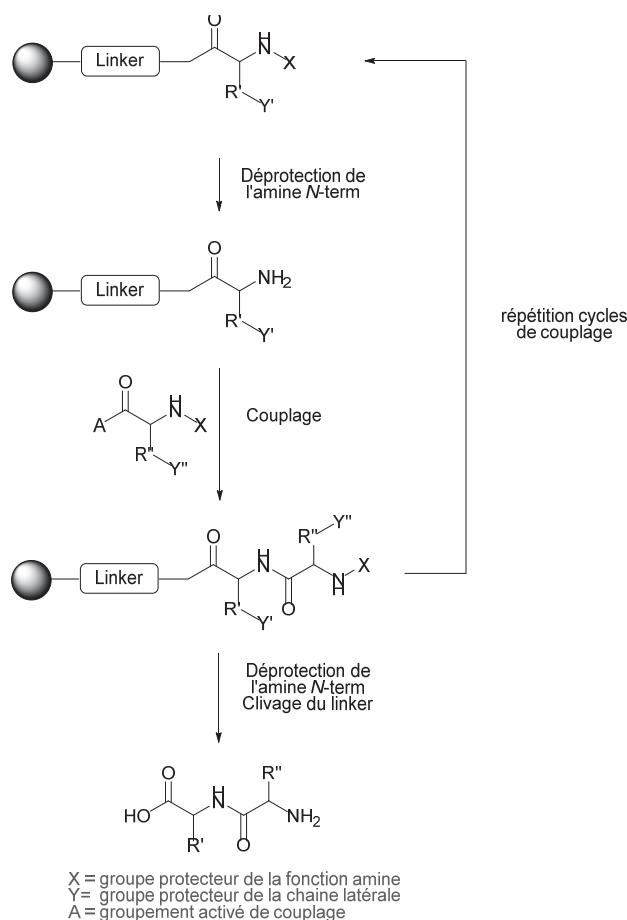
soufre dans les structures de protéines ainsi que sa forte nucléophilie facilite le couplage spécifique de glycanes sur la chaîne peptidique. D'autre part, comme décrit précédemment les liaisons thioglycosidiques sont tolérées par la plupart des systèmes biologiques et restent peu sensibles à l'hydrolyse enzymatique, elles sont donc compatibles avec leur étude *in-vivo*. Enfin, l'obtention de formes homogènes de protéines S-glycosylées permettrait d'accroître nos connaissances sur la S-glycosylation des protéines, phénomène rare et récemment identifié dans le monde du vivant.

## **2. Outils de synthèse pour l'accès aux néothioglycopeptides et protéines**

### **2.1. Glycosylation par voie chimique**

A l'image de la synthèse des thioglycosides et oligothiosaccharides, de nombreuses méthodologies ont été décrites pour permettre l'accès aux S-glycopeptides et S-glycoprotéines.

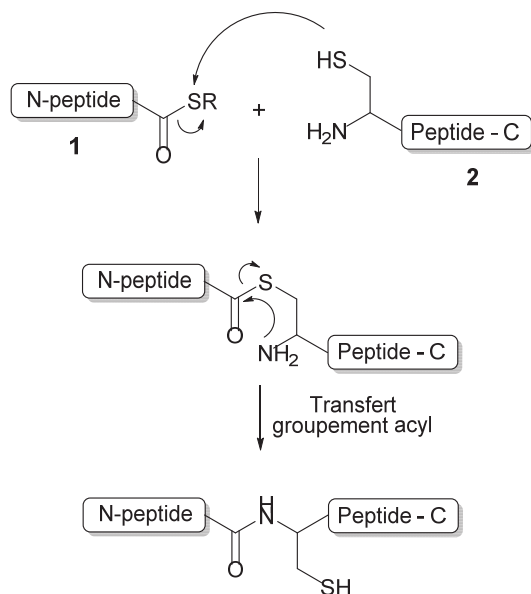
La synthèse de peptides est le résultat de la réaction de couplage entre le groupe carboxyle (extrémité C-terminale) d'un acide aminé et de la fonction amine (extrémité N-terminale) d'un autre acide aminé. A l'heure actuelle les peptides sont synthétisés en phase solide suivant la méthode développée par le Prof. Merrifield en 1963. Cette technique repose sur la fixation sur support solide d'un acide aminé *via* la position C-term, l'extrémité N-term peut alors réagir avec un second acide aminé protégé. Après la réaction de couplage un lavage est effectué pour éliminer l'excédent de réactifs. L'élongation de la séquence peptidique est donc assurée par la succession d'étapes de couplage / déprotection / lavage (Schéma 18). Classiquement les groupements *t*Boc (*Tert*-butoxycarbonyl) et Fmoc (9-fluorenylmethyloxycarbonyl) sont utilisés pour masquer la position N-term d'un peptide.



**Schéma 18** : Synthèse peptidique en phase solide (SPPS).

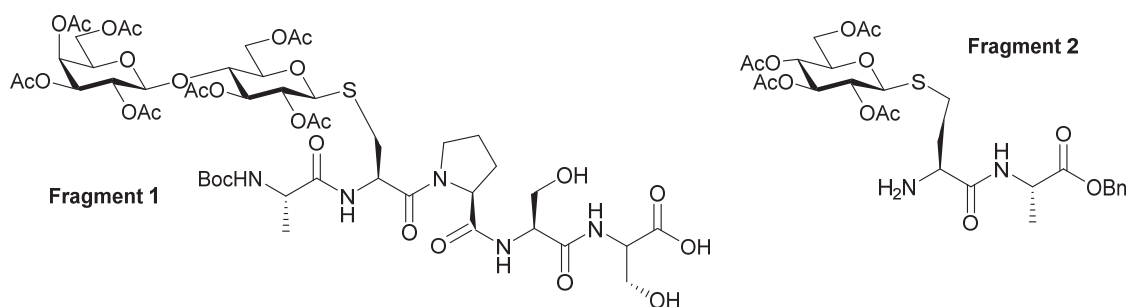
Cette méthodologie a été appliquée à la synthèse de peptides glycosylés, en incorporant l'acide aminé glycosylé sur la chaîne peptidique en construction. Plusieurs voies de synthèse permettent la thioglycosylation d'AA : 1) l'alkylation directe de la fonction thiol ou thioacétate anomérique *via* un AA halogéné (typiquement la Bromo-Alanine ou Iodo-Alanine) ; 2) la condensation de Mitsunobu permet le transfert de la fonction thiol du sucre sur la fonction alcool de l'AA dans des conditions douces de réaction ; ou 3) l'addition type Michael du thioglycoside sur la double liaison électrophile de la déhydroalanine. (Pachamuthu, **2006**) La synthèse de thioglycopeptides par SPPS est une méthode pratique car elle permet la glycosylation directe d'un acide aminé qui peut dès lors être introduit dans la chaîne peptidique et ce de manière totalement spécifique. L'inconvénient majeur de cette technique reste dans l'impossibilité de synthétiser des peptides de plus de 50 AA, ce qui limite son utilisation pour la synthèse de protéines glycosylées. Une alternative à cette limitation repose sur la technique de Ligation Chimique Native (LCN) qui consiste à combiner plusieurs fragments peptidiques (Schéma 19).





**Schéma 19 :** Mécanisme de Ligation Chimique Native.

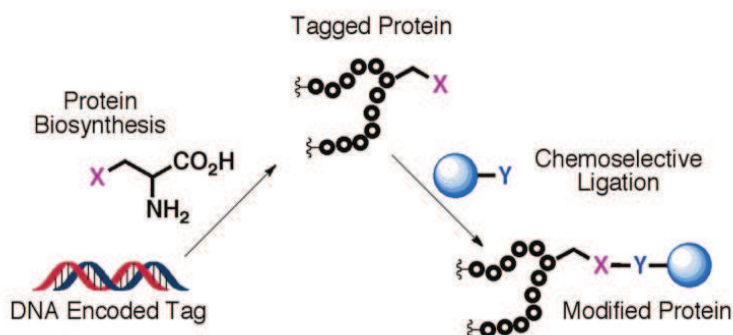
La synthèse de thioglycosides est généralement réalisée par réaction de type  $S_N2$  entre un thiol accepteur et un donneur de glycosyl halogéné. Ainsi un peptide, reconstitué par LCN, portant une fonction thiol peut être utilisé dans la réaction de glycosylation en présence d'un donneur de glycoside activé *via* un promoteur basique (tel que  $Na_2CO_3$ ). En suivant cette méthodologie, la synthèse d'un glycopeptide analogue à la glycoprotéine Tamm-Horsfall (protéine de l'urine humaine) a pu être obtenue à partir de 2 fragments peptidiques thioglycosylés (Figure 14). (Zhu, 2004)



**Figure 14 :** Fragments peptidiques glycosylés par LCN, briques élémentaires d'un analogue de la protéine Tamm-Horsfall.

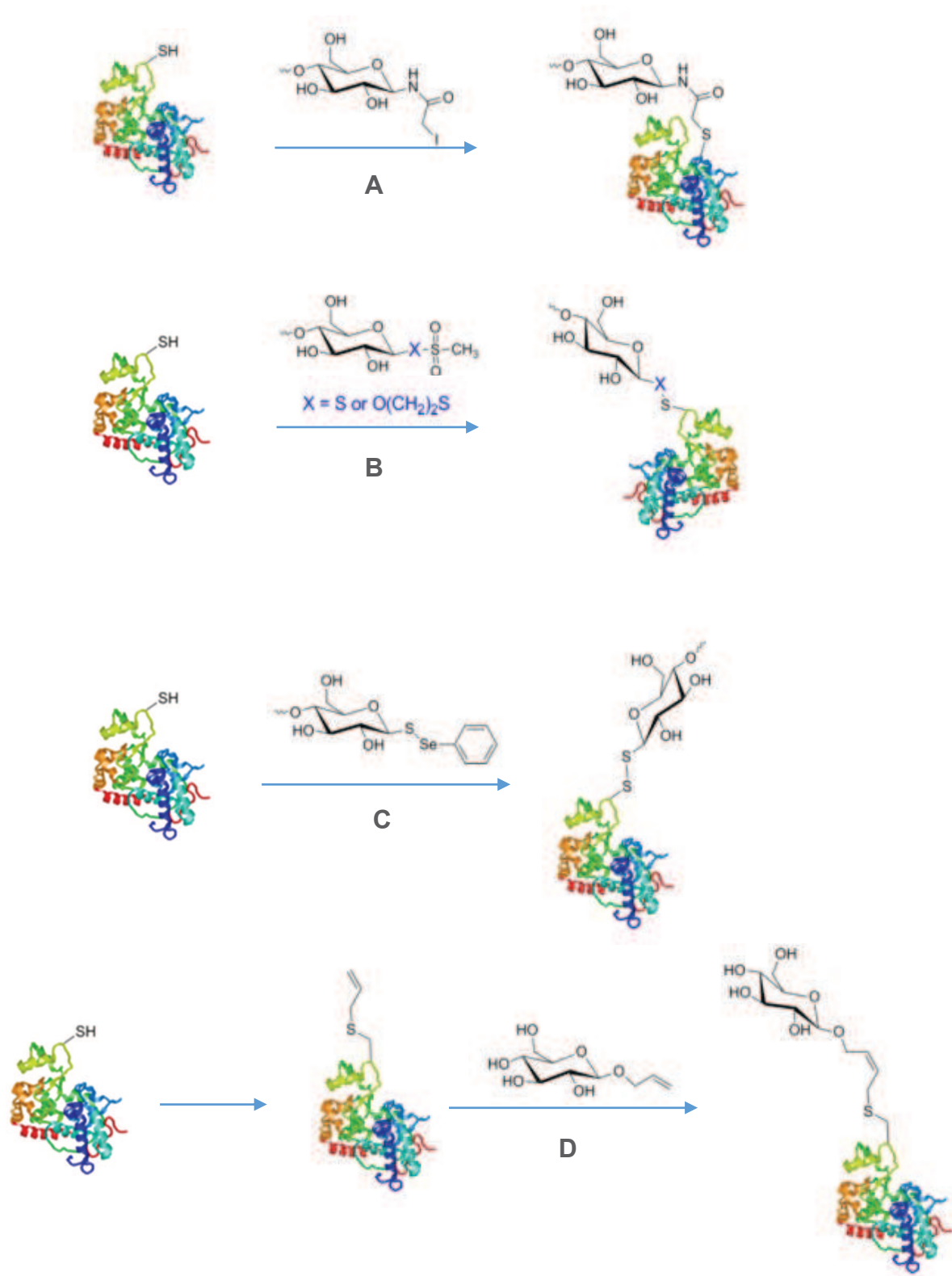


Enfin la synthèse de néothioglycoprotéines a été envisagée par la conjugaison chimiosélective de glycanes sur une protéine entière. Plusieurs méthodes ont été développées et reposent sur la réactivité de la fonction thiol d'un résidu Cystéine présent sur la protéine. En 2011 l'équipe du Prof. Davis proposa une stratégie innovante de modification de protéines appelée « Tag-and-Modify » qui consiste à introduire de manière spécifique (par mutagenèse dirigée) un acide aminé cible sur la séquence protéique. La présence de ce marqueur permet par la suite le couplage chimique sélectif avec un donneur glycosylé approprié (Figure 15).(Chalker, 2011)



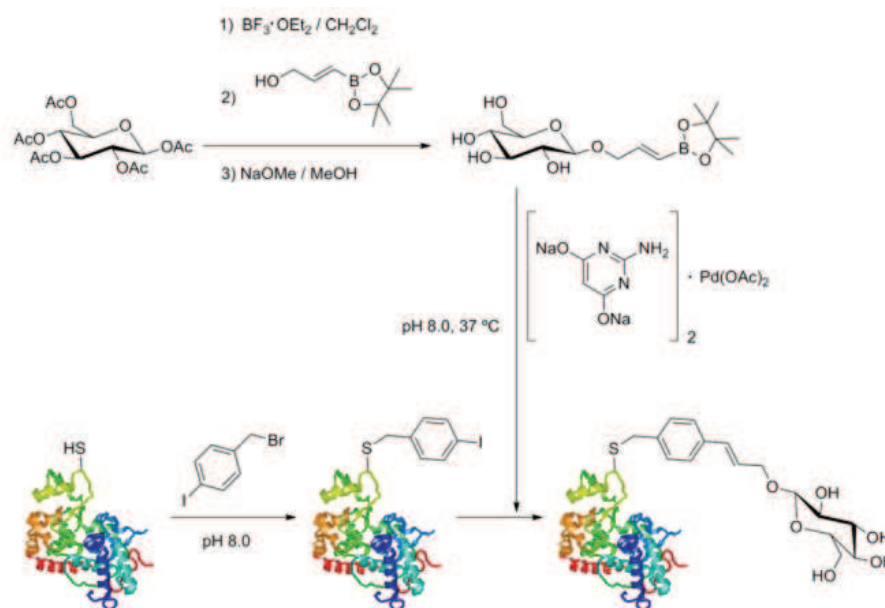
**Figure 15 :** Stratégie de « Tag and Modify » pour le couplage sélectif glycan-protéine.(Chalker, 2011)

L'avantage de cet acide aminé repose sur sa relative faible abondance dans les séquences protéiques ce qui permet d'assurer l'introduction spécifique et homogène de glycanes. Les protéines portant un marqueur de type Cystéine peuvent entre autre être glycosylées par i) réaction d'alkylation de la fonction thiol *via* une unité sucre portant un groupement iodoacétamide (**A**, Figure 16), ii) ligation chimique en créant un pont disulfure à partir d'un donneur de type glycoside de méthaneithiosulfonate (**B**) ou sélénylsulfide (**C**) ou iii) par réaction de métathèse (**D**). (Pachamuthu, 2006; Villalonga, 2014)



**Figure 16 :** Stratégies de thioglycosylation de protéines par voie chimique. Adaptée de(Villalonga, 2014)

L'équipe du Prof. Davis a reporté la glycosylation d'une protéine en appliquant la réaction de couplage de Suzuki-Miyaura. Le résidu Cys de la protéine a dans un premier temps été fonctionnalisé par un groupement iodoarylé, la réaction de couplage Pallado-catalysée a par la suite été réalisée sur le dérivé glycosylé vinylique correspondant (Schéma 20). (Chalker, **2009**) Par la suite ils ont étendu leur concept en incorporant des AA non-naturels (AAns) permettant la modification des protéines *via* une variété de transformations bio-orthogonales, telles que les réactions de couplage *via* les métaux de transition à partir de marqueurs halogénés, ou les réactions de chimie click *via* des AA portant un groupement azido sur leur chaîne latérale. (Chalker, **2011**)

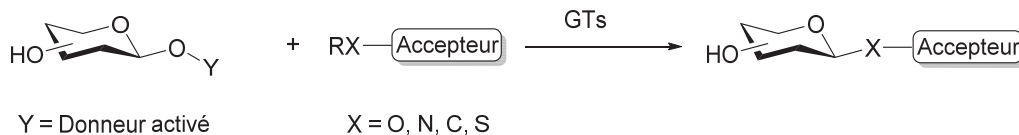


**Schéma 20** : Glycosylation enzymatique de protéines *via* la réaction de métathèse de Suzuki-Miyaura. (Villalonga, **2014**)

A l'heure actuelle de nombreuses voies de synthèse chimique de thioglycopeptides et thioglycoprotéines sont disponibles et ont prouvé leur efficacité, à l'inverse les méthodologies par voie enzymatique ont suscité peu d'intérêt. Les GT et GH sont classiquement utilisées pour la *N*- et *O*-glycosylation de séquences peptidiques. La découverte récente de *S*-GT naturelles et le développement des thioglycoligases sont le point de départ d'une nouvelle catégorie de biocatalyseurs innovants pour l'accès aux peptides et protéines *S*-glycosylés.

## 2.2. Les outils enzymatiques, nouvelles perspectives ?

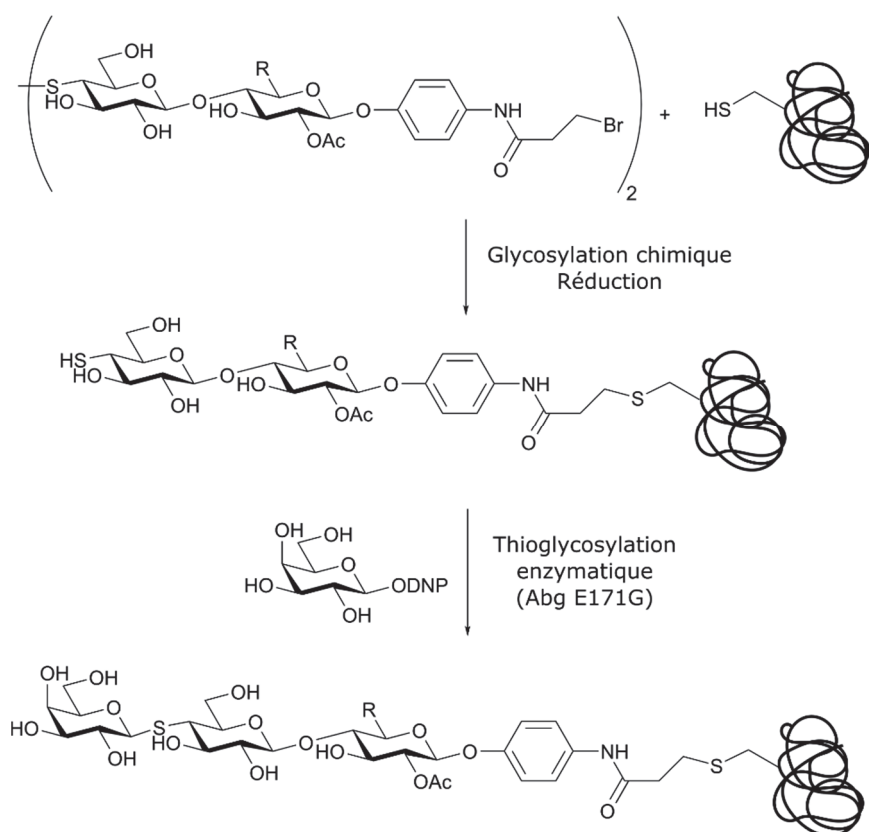
Les glycosyltransférases catalysent la formation de liaison *O*-, *N*-, *C*- et *S*-glycosidiques entre un donneur de glycoside activé (nucléotide-sucre de manière générale) et une variété d'accepteurs (sucres, lipides, peptides, ...) avec une forte régio- et stéréosélectivité (Figure 17). (Lairson, **2008**)



**Figure 17** : Réaction générale catalysée par les glycosyltransférases.

En synthèse les GT ont suscité un intérêt particulier grâce à leur flexibilité vis-à-vis de l'accepteur comme outils de glycodiversification. A l'heure actuelle de nombreuses GT ont été décrites pour la synthèse d'oligosaccharides, de glycopeptides ou protéines ou encore de petites molécules d'intérêt thérapeutique.(Gantt, **2011**; Schmaltz, **2011**) L'emploi de ce type de biocatalyseur offre donc des perspectives intéressantes pour la synthèse de nouveaux glycoconjugués naturels ou artificiels, tels que les composés thioglycosylés.(Thibodeaux, **2008**) Afin d'améliorer la versatilité de ces enzymes envers leurs substrats des méthodes de bioingénierie ont été développées et ont fait l'objet de nombreuses revues.(Chang, **2011**) A l'heure actuelle deux exemples de S-GT appliqués à la synthèse de S-glycoconjugués sont rapportés dans la littérature. En 2008, des études de criblage ont été réalisées sur un triple mutant de l'enzyme OleD (GT1), produite chez *Streptomyces antibioticus*, afin de détecter une activité de transférase sur un panel de 130 accepteurs.(Gantt, **2008**) Ainsi 4 composés présentant un atome de soufre se sont révélés positifs au screening confirmant la capacité d'OleD type sauvage et mutée à reconnaître ce type d'accepteur et donc de catalyser des S-glycosylations. Plus récemment, Wang *et coll.* ont cloné et exprimé le gène codant pour la S-GT ThuS et ont rapporté la thioglycosylation *in-vitro* du peptide thurandacine A.(Wang, **2014**) Si les techniques de bioingénierie disponibles à l'heure actuelle s'avèrent être de précieux outils pour améliorer le potentiel en biocatalyse à grande échelle des glycosyltransférases, de nombreux obstacles restent à contourner (difficultés d'expression, problèmes de solubilité, de coûts des donneurs de glycosyl activés).(Schmaltz, **2011**)

Dans le domaine des glycosidases, Mullegger *et coll.* ont proposé une application intéressante des thioglycoligases pour la génération de thioglycoprotéines en suivant un processus en 3 étapes (Schéma 21).(Mullegger, **2006b**) Dans un premier temps l'équipe s'est appliquée à la conception d'une néoglycoprotéine pouvant servir d'accepteur de glycosides et permettant ainsi le suivi de la réaction enzymatique. Leur choix s'est porté sur l'endo-xylanase de *B. circulans* (Bcx), protéine de petite taille et caractérisée. Cette protéine peut servir de bon modèle puisqu'étant une enzyme, il est facile de suivre l'effet de modifications chimiques ou enzymatiques sur l'intégrité structurale de la protéine par contrôle de son activité enzymatique. Enfin Bcx, dont la structure tridimensionnelle a été largement étudiée, ne présente aucun résidu Cystéine sur sa séquence protéique. L'introduction d'un motif thiol unique a été réalisée par mutagenèse dirigée résultant le mutant Bcx S22C. Dans un second temps une étape de conjugaison chimique est réalisée à partir d'un thioglycoconjugué halogéné (synthétisé à partir du 4'-déoxy-4'-thiocellobioside) pour former la néoglycoprotéine SG2-Bcx S22C. Enfin la protéine ainsi glycosylée a pu servir d'accepteur pour la réaction enzymatique catalysée par la thioligase Abg E171G en présence du donneur activé DNPGal, *via* la fonction thiol de l'extrémité non-réductrice du disaccharide.



**Schéma 21** : Thioglycosylation chimio-enzymatique de l'endo-xylanase Bcx de *B. circulans*.

L'équipe a pu confirmer la production de la protéine décorée d'un trisaccharide par spectrométrie de masse MALDI-TOF et ainsi démontrer pour la première fois le potentiel des thioglycoligases dans la génération de protéines S-glycosylées. Malgré l'originalité de la méthode, qui reste le seul exemple décrit dans la littérature, la technologie des thioglycoligases a été utilisée ici pour étendre la chaîne oligosaccharidique déjà greffée sur la protéine par conjugaison chimique. Il serait alors intéressant de développer une méthode permettant la glycosylation directe de glycanes sur les protéines à partir des thioglycoligases. Compte-tenu du potentiel de synthèse des thioligases, la validation d'une telle méthodologie permettrait alors d'étendre le spectre d'applications de ces biocatalyseurs innovants, de la synthèse de thioglycosides simples aux thioglycoprotéines, comme alternative efficace et éco-compatible aux procédés de chimie classique.

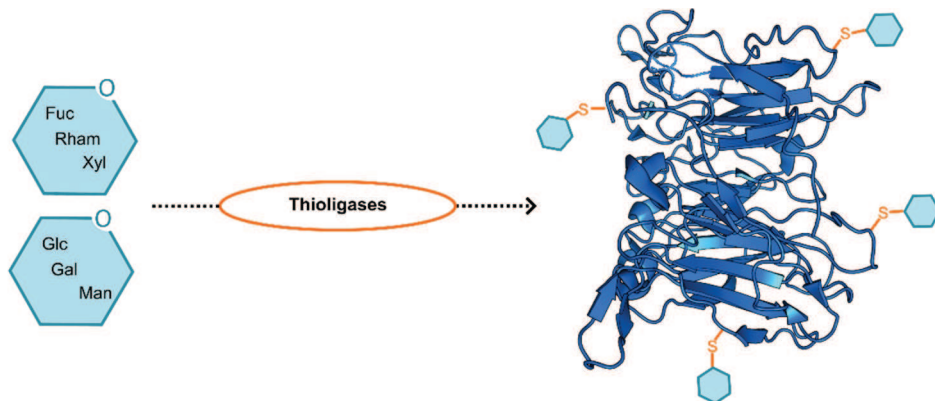
### III. Objectifs de la thèse

Le challenge à l'heure actuelle repose sur la compréhension des mécanismes qui gouvernent la structure, la biosynthèse et les fonctions biologiques des glycanes et requière donc l'accès à des formes homogènes et stables de ces glycoconjugués. (Gamblin, **2009**) Pour répondre à ce défi de nombreuses méthodologies de synthèse ont été proposées pour faciliter le couplage d'unités sucres individuelles ou de séquences d'oligosaccharides sur un acide aminé spécifique d'une séquence peptidique. (Boutureira, **2015**) Parmi ces méthodes les thioglycosides ont été utilisés pour générer des analogues S-glycosylés de protéines

naturelles présentant une meilleure résistance vis-à-vis de l'hydrolyse et compatibles avec la réalisation d'étude *in-vivo*.(Pachamuthu, **2006**) De plus, il a récemment été démontré que des formes naturelles de thioglycoprotéines étaient retrouvées dans le monde du vivant, il est donc d'un intérêt majeur de comprendre le rôle de ce nouveau mode de glycosylation.(Lafite, **2012**)

Parallèlement les enjeux environnementaux et économiques de la société actuelle ont propulsé l'utilisation des enzymes comme alternative éco-compatible aux procédés de chimie classique par exemple en limitant le rejet de déchets polluants et en développant des procédés moins énergivores.(Menon, **2012**) Dans le domaine des glycosciences, les glycosyltransférases et glycoside hydrolases ont prouvé leur potentiel en synthèse en facilitant l'accès à de nombreux glycoconjugués complexes de manière régio- et stéréospécifiques.(Hancock, **2006**) Les enzymes capables de la création de liaisons S-glycosidiques sont de deux types, les S-GT naturelles découvertes récemment, et les thioligases, glycosidases modifiées par bioingénierie qui ont prouvé leur potentiel pour la synthèse de thioglycosides.(Guillotin, **2014**) A l'heure actuelle un seul exemple dans la littérature rapporte l'utilisation d'une S-GT pour le couplage d'une unité sucre sur un acide aminé *via* une liaison S-glycosidique.

C'est dans ce contexte que notre projet s'inscrit en proposant une nouvelle voie d'accès enzymatique aux acides aminés S-glycosylés à partir des thioglycoligases. Les outils enzymatiques développés et le concept validé, ce dernier pourra être étendu à la synthèse de thioglycopeptides et thioglycoprotéines (Figure 18).



**Figure 18 :** Les thioglycoligases nouvel outil enzymatique pour la synthèse de S-glycoprotéines.

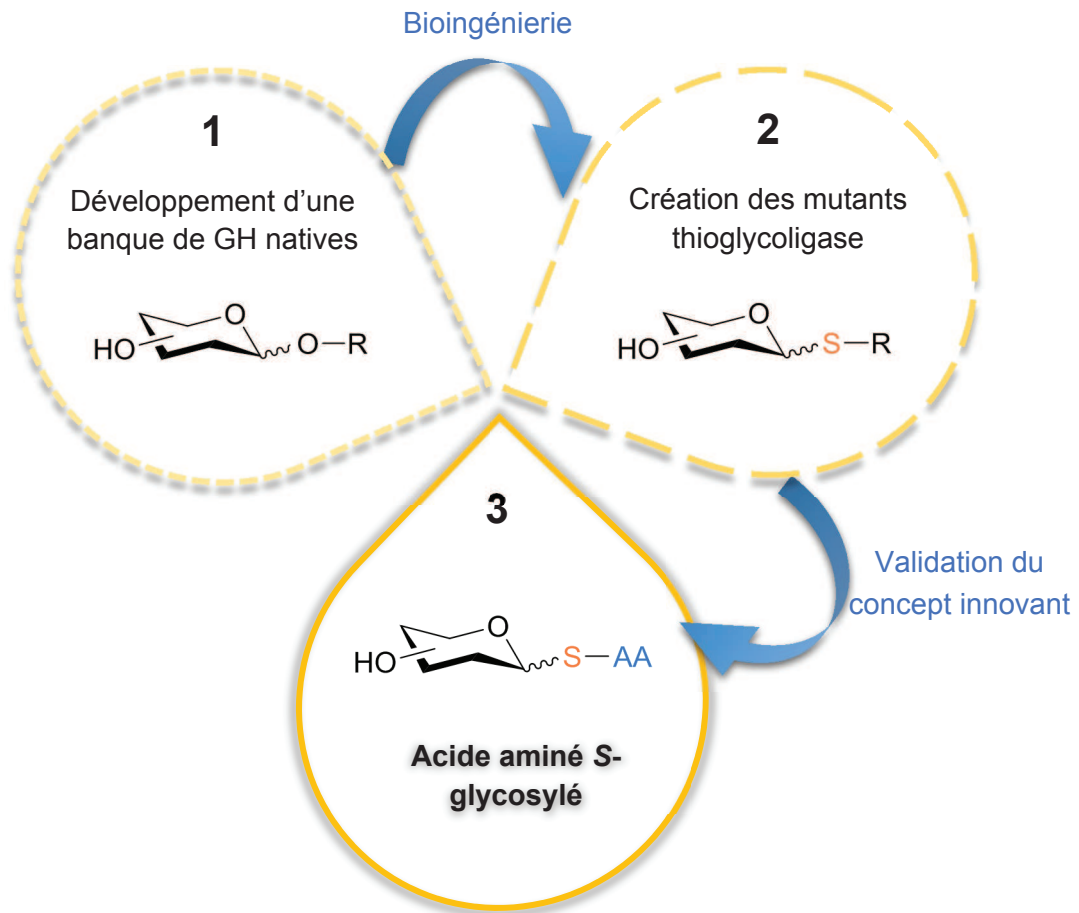
Ainsi le premier chapitre aura pour objet l'élaboration d'une variété de glycosidases natives qui servira de banque pour la génération des mutants thioglycoligases. La stratégie de production des biocatalyseurs en type sauvage sera décrite, de la sélection d'un micro-organisme « source » à partir duquel seront sélectionnés les gènes d'intérêts, au choix du système de clonage d'ADN et de surexpression des protéines. Chaque glycosidase produite et purifiée sera caractérisée pour déterminer la spécificité de substrat, les conditions optimales d'hydrolyse (pH, température, co-solvant) ainsi que les constantes cinétiques. Des essais de cristallogénèse seront également décrits pour obtenir de plus amples informations sur les



structures tridimensionnelles de ces protéines et comprendre leur mode d'action. Par la suite, le potentiel en biocatalyse des enzymes sera évalué sur un panel d'alcools aliphatiques et les meilleurs candidats seront valorisés pour développer une voie d'accès chimio-enzymatique aux glycéroglycosides, composés d'intérêt dans le domaine de la cosmétique, dont les propriétés biologiques seront testées grâce à un partenariat établi avec l'entreprise GLYcoDiag.

Le second chapitre portera sur la conception de mutants thioligases par mutagenèse dirigée à partir de la banque de glycosidases natives disponible. Seront alors décrits la stratégie de préparation des mutants, depuis la préparation des amorces nécessaires à la modification génétique à l'amplification de l'ADN par PCR, ainsi que la production des enzymes correspondantes et leur caractérisation en activité d'hydrolyse. Par la suite les mutants seront testés en thioglycosylation sur un panel de thiols aliphatiques et aromatiques et leur influence sur les propriétés cinétiques des enzymes sera déterminée. Pour inscrire nos enzymes dans une démarche de synthèse éco-compatible, ils seront utilisés en biocatalyse pour la production de dérivés S-glycosylés. Enfin des études de cristallogenèse à partir de mutants dépourvus d'activité d'hydrolyse seront réalisées pour obtenir des co-cristaux en présence de substrat qui permettront d'obtenir de plus amples informations sur la structure du site actif et sur le mode de reconnaissance du substrat.

Enfin le dernier chapitre sera quant à lui consacré l'utilisation du meilleur candidat thioglycoligase pour réaliser le couplage glycosidique sur un acide aminé et ainsi proposer la première voie de synthèse enzymatique d'un acide aminé S-glycosylé. Seront alors décrits la synthèse d'un acide aminé non-naturel ainsi que la mise au point d'une méthode de suivi de la réaction enzymatique permettant la détection et l'identification des produits résultants (Schéma 22).



**Schéma 22** : Présentation des principaux objectifs de la thèse.





## Bibliographie

- Abbott, D. W., Macauley, M. S., Vocadlo, D. J., and Boraston, A. B.; *J. Biol. Chem.* **(2009)**, *284*, 11676-11689
- Agerbirk, N., and Olsen, C. E.; *Phytochemistry* **(2012)**, *77*, 16-45
- Ajisaka, K., and Fujimoto, H.; *Carbohydr. Res.* **(1989)**, *185*, 139-146
- Ajisaka, K., Nishida, H., and Fujimoto, H.; *Biotechnol. Lett.* **(1987)**, *9*, 243-248
- Almendros, M., Danalev, D., Francois-Heude, M., Loyer, P., Legentil, L., Nugier-Chauvin, C., Daniellou, R., and Ferrieres, V.; *Org. Biomol. Chem.* **(2011)**, *9*, 8371-8378
- Almendros, M., François-Heude, M., Legentil, L., Nugier-Chauvin, C., Daniellou, R., and Ferrières, V.; *ARKIVOC* **(2013)**, *2*, 123-132
- Ansari, S. A., and Husain, Q.; *Biotechnol. Adv.* **(2012)**, *30*, 512-523
- Armstrong, Z., Reitingner, S., Kantner, T., and Withers, S. G.; *ChemBiochem* **(2010)**, *11*, 533-538
- Armstrong, Z., and Withers, S. G.; *Biopolymers* **(2013)**, *99*, 666-674
- Arnold, F. H.; *Nature* **(2001)**, *409*, 253-257
- Arnold, F. H., and Volkov, A. A.; *Curr. Opin. Chem. Biol.* **(1999)**, *3*, 54-59
- Basso, A., Ducret, A., Gardossi, L., and Lortie, R.; *Tetrahedron Lett.* **(2002)**, *43*, 2005-2008
- Ben-David, A., Shoham, G., and Shoham, Y.; *Chem. Biol.* **(2008)**, *15*, 546-551
- Bergquist, P. L., Morgan, H. W., and Saul, D.; *Curr. Biotechnol.* **(2014)**, *3*, 45-59
- Bhalla, A., Bansal, N., Kumar, S., Bischoff, K. M., and Sani, R. K.; *Biores. Technol.* **(2013)**, *128*, 751-759
- Bissaro, B., Durand, J., Biarnés, X., Planas, A., Monsan, P., O'Donohue, M. J., and Fauré, R.; *ACS Catal.* **(2015)**, 4598-4611
- Bojarova, P., and Kren, V.; *Trends Biotechnol.* **(2009)**, *27*, 199-209
- Bourquelot, E., and Bridel, M.; *Ann. Chim. Phys.* **(1913)**, *29*, 145 - 218
- Boutureira, O., and Bernardes, G. J.; *Chem. Rev.* **(2015)**, *115*, 2174-2195
- Brachet, E., Brion, J.-D., Messaoudi, S., and Alami, M.; *Advanced Synthesis & Catalysis* **(2013a)**, *355*, 477-490
- Brachet, E., Brion, J. D., Alami, M., and Messaoudi, S.; *Chem. Eur. J.* **(2013b)**, *19*, 15276-15280
- Bruneau, A., Roche, M., Hamze, A., Brion, J. D., Alami, M., and Messaoudi, S.; *Chem. Eur. J.* **(2015)**, *21*, 8375-8379
- Campo, V. L., Aragao-Leoneti, V., Teixeira, M. B. M., and Carvalho, I.; *New Dev. Med. Chem.* **(2010)**, *1*, 133-151
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B.; *Nucleic Acids Res.* **(2009)**, *37*, D233-238
- Chahid, Z., Montet, D., Pina, M., and Graille, J.; *Biotechnol. Lett.* **(1992)**, *14*, 281-284
- Chalker, J. M., Bernardes, G. J. L., and Davis, B. G.; *Acc. Chem. Res.* **(2011)**, *44*, 730-741
- Chalker, J. M., Wood, C. S., and Davis, B. G.; *J. Am. Chem. Soc.* **(2009)**, *131*, 16346-16347
- Champion, E., Andre, I., Moulis, C., Boutet, J., Descroix, K., Morel, S., Monsan, P., Mulard, L. A., and Remaud-Simeon, M.; *J. Am. Chem. Soc.* **(2009)**, *131*, 7379-7389

- Champion, E., Guerin, F., Moulis, C., Barbe, S., Tran, T. H., Morel, S., Descroix, K., Monsan, P., Mourey, L., Mulard, L. A., Tranier, S., Remaud-Simeon, M., and Andre, I.; *J. Am. Chem. Soc.* **(2012)**, *134*, 18677-18688
- Chang, A., Singh, S., Phillips Jr, G. N., and Thorson, J. S.; *Curr. Opin. Biotechnol.* **(2011)**, *22*, 800-808
- Cheng, F., Zhu, L., and Schwaneberg, U.; *Chem. Commun. (Cambridge, U.K.)* **(2015)**, *51*, 9760-9772
- Chiffoleau-Giraud, v., Spangenberg, P., and Rabiller, C.; *Tetrahedron-Asymmetry* **(1997)**, *8*, 2017-2023
- Choi, S. P., Nguyen, M. T., and Sim, S. J.; *Biores. Technol.* **(2010)**, *101*, 5330-5336
- Cobucci-Ponzano, B., and Moracci, M.; *Nat. Prod. Rep.* **(2012a)**, *29*, 697-709
- Cobucci-Ponzano, B., Perugini, G., Strazzulli, A., Rossi, M., and Moracci, M.; *Methods Enzymol.* **(2012b)**, *510*, 273-300
- Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M., and Moracci, M.; *Biochemistry* **(2003)**, *42*, 9525-9531
- Crout, D. H. G., and Vic, G.; *Curr. Opin. Chem. biol.* **(1998)**, *2*, 98-111
- Debruyne, C. K., and Yde, M.; *Carbohydr. Res.* **(1977)**, *56*, 153-164
- Defaus, S., Gupta, P., Andreu, D., and Gutierrez-Gallego, R.; *Analyst* **(2014)**, *139*, 2944-2967
- Dintinger, T., Dutheilbouedec, D., Bouchonneau, M., and Tellier, C.; *Biotechnol. Lett.* **(1994)**, *16*, 689-692
- Driguez, H.; *Top. Curr. Chem.* **(1997)**, *187*, 85-116
- Driguez, H.; *Chembiochem* **(2001)**, *2*, 311-318
- Ducros, V. M. A., Tarling, C. A., Zechel, D. L., Brzozowski, A. M., Frandsen, T. P., von Ossowski, I., Schüle, M., Withers, S. G., and Davies, G. J.; *Chem. Biol.* **(2003)**, *10*, 619-628
- Ernst, B., and Magnani, J. L.; *Nat. Rev. Drug Discovery* **(2009)**, *8*, 661-677
- Farran, A., Cai, C., Sandoval, M., Xu, Y., Liu, J., Hernaiz, M. J., and Linhardt, R. J.; *Chem. Rev.* **(2015)**, *115*, 6811-6853
- Feng, H. Y., Drone, J., Hoffmann, L., Tran, V., Tellier, C., Rabiller, C., and Dion, M.; *J. Biol. Chem.* **(2005)**, *280*, 37088-37097
- Fernandez-Tejada, A., Canada, F. J., and Jimenez-Barbero, J.; *Chemistry* **(2015)**, *21*, 10616-10628
- Fernandez Mayoralas, A.; *Top. Curr. Chem.* **(1997)**, *186*, 1-20
- Ferrier, R. J., and Furneaux, R. H.; *Carbohydr. Res.* **(1976)**, *52*, 63-68
- Ferrier, R. J., Hay, R. W., and Vethaviyaear, N.; *Carbohydr. Res.* **(1973)**, *27*, 55-61
- Gamblin, D. P., Scanlan, E. M., and Davis, B. G.; *Chem. Rev.* **(2009)**, *109*, 131-163
- Gantt, R. W., Goff, R. D., Williams, G. J., and Thorson, J. S.; *Angew. Chem., Int. Ed.* **(2008)**, *47*, 8889-8892
- Gantt, R. W., Peltier-Pain, P., and Thorson, J. S.; *Nat. Prod. Rep.* **(2011)**, *28*, 1811-1853
- Gillam, S., and Smith, M.; *Genes* **(1979)**, *8*, 81-97
- Guillotin, L., Lafite, P., and Daniellou, R.; *Carbohydr. Chem.* **(2014)**, *40*, 178-194
- Hancock, S. M., D Vaughan, M., and Withers, S. G.; *Curr. Opin Chem. biol.* **(2006)**, *10*, 509-519

- Hansson, T., Andersson, M., and Adlercreutz, P.; *Enzyme Microb. Technol.* **(2001)**, 29, 527-534
- Hehre, E. J., Brewer, C. F., and Genghof, D. S.; *J. Biol. Chem.* **(1979)**, 254, 5942-5950
- Henrissat, B.; **(1991a)**, <http://www.cazy.org/>
- Henrissat, B.; *Biochem. J.* **(1991b)**, 280, 309-316
- Herr, R. R., and Bergy, M. E.; *Antimicrob. Agents Chemother.* **(1962)**, 560,
- Hoeksema, H., Bannister, B., Birkenmeyer, R. D., Kagan, F., Magerlein, B. J., Mackellar, F. A., Schroeder, W., Slomp, G., and Herr, R. R.; *J. Am. Chem. Soc.* **(1964)**, 86, 4223-4224
- Honda, Y., and Kitaoka, M.; *J. Biol. Chem.* **(2006)**, 281, 1426-1431
- Horsch, M., Mayer, C., Sennhauser, U., and Rast, D. M.; *Pharmacol. Ther.* **(1997)**, 76, 187-218
- Jahn, M., Chen, H. M., Mullegger, J., Marles, J., Warren, R. A. J., and Withers, S. G.; *Chem. Commun. (Cambridge, U.K.)* **(2004)**, 274-275
- Jahn, M., Marles, J., Warren, R. A. J., and Withers, S. G.; *Angew. Chem., Int. Ed.* **(2003)**, 42, 352-354
- Juturu, V., and Wu, J. C.; *Renewable & Sustainable Energy Rev.* **(2014)**, 33, 188-203
- Kim, Y.-W., Chen, H.-M., Kim, J. H., Muellegger, J., Mahuran, D., and Withers, S. G.; *Chembiochem* **(2007)**, 8, 1495-1499
- Kim, Y. W., Chen, H. M., Kim, J. H., and Withers, S. G.; *FEBS Lett.* **(2006a)**, 580, 4377-4381
- Kim, Y. W., Lee, S. S., Warren, R. A., and Withers, S. G.; *J. Biol. Chem.* **(2004)**, 279, 42787-42793
- Kim, Y. W., Lovering, A. L., Chen, H. M., Kantner, T., McIntosh, L. P., Strynadka, N. C. J., and Withers, S. G.; *J. Am. Chem. Soc.* **(2006b)**, 128, 2202-2203
- Kittl, R., and Withers, S. G.; *Carbohydr. Res.* **(2010)**, 345, 1272-1279
- Kobayashi, S., Kashiwa, K., Kawasaki, T., and Shoda, S.; *J. Am. Chem. Soc.* **(1991)**, 113, 3079-3084
- Koshland, D. E.; *Biol. Rev. Cambridge Philos. Soc.* **(1953)**, 28, 416-436
- Kwan, D. H., and Withers, S. G.; *J. Carbohydr. Chem.* **(2011)**, 30, 181-205
- Lafite, P., and Daniellou, R.; *Nat. Prod. Rep.* **(2012)**, 29, 729-738
- Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G.; *Annu. Rev. Biochem.* **(2008)**, 77, 521-555
- Lehmann, M., and Wyss, M.; *Curr. Opin. Biotechnol.* **(2001)**, 12, 371-375
- Lian, G., Zhang, X., and Yu, B.; *Carbohydr. Res.* **(2015)**, 403, 13-22
- Lodge, J. A., Maier, T., Liebl, W., Hoffmann, V., and Strater, N.; *J. Biol. Chem.* **(2003)**, 278, 19151-19158
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B.; *Nucleic Acids Res.* **(2014)**, 42, D490-D495
- Lote, C. J., and Weiss, J. B.; *Biochem. J.* **(1971a)**, 123, 25P
- Lote, C. J., and Weiss, J. B.; *FEBS Lett.* **(1971b)**, 16, 81-85
- Mackenzie, L. F., Wang, Q. P., Warren, R. A. J., and Withers, S. G.; *J. Am. Chem. Soc.* **(1998)**, 120, 5583-5584
- Malet, C., and Planas, A.; *FEBS Lett.* **(1998)**, 440, 208-212
- Mandal, S., and Nilsson, U. J.; *Org. Biomol. Chem.* **(2014)**, 12, 4816-4819

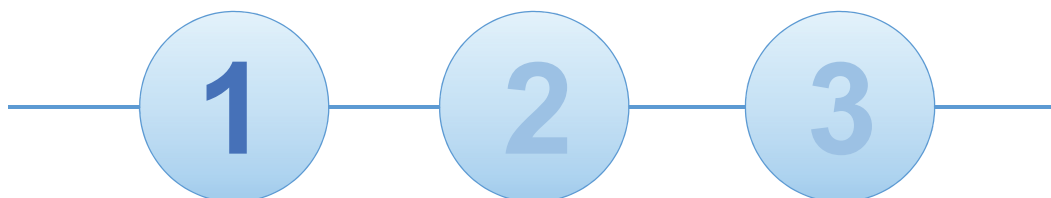
- Mayer, C., Zechel, D. L., Reid, S. P., Warren, R. A. J., and Withers, S. G.; *FEBS Lett.* **(2000)**, *466*, 40-44
- Menon, V., and Rao, M.; *Prog. Energy Combust. Sci.* **(2012)**, *38*, 522-550
- Mitsuo, N., Takeichi, H., and Satoh, T.; *Chem. Pharm. Bull.* **(1984)**, *32*, 1183-1187
- Mladenoska, I., Winkelhausen, E., and Kuzmanova, S.; *Food Technol. Biotechnol.* **(2008)**, *46*, 311-316
- Monsan, P., and Paul, F.; *FEMS Microbiol. rev.* **(1995)**, *16*, 187-192
- Mrazek, H., Weignerova, L., Bojarova, P., Novak, P., Vanek, O., and Bezouska, K.; *Biotechnol. Adv.* **(2013)**, *31*, 17-37
- Mullegger, J., Chen, H. M., Chan, W. Y., Reid, S. P., Jahn, M., Warren, R. A. J., Salleh, H. M., and Withers, S. G.; *Chembiochem* **(2006a)**, *7*, 1028-1030
- Mullegger, J., Chen, H. M., Warren, R. A. J., and Withers, S. G.; *Angew. Chem., Int. Ed.* **(2006b)**, *45*, 2585-2588
- Mullegger, J., Jahn, M., Chen, H. M., Warren, R. A. J., and Withers, S. G.; *Protein Eng., Des. Sel.* **(2005)**, *18*, 33-40
- Ngane, A. N., Lavault, M., Séraphin, D., Landreau, A., and Richomme, P.; *Carbohydr. Res.* **(2006)**, *341*, 2799-2802
- Ohnuma, T., Fukuda, T., Dozen, S., Honda, Y., Kitaoka, M., and Fukamizo, T.; *Biochem. J.* **(2012)**, *444*, 437-443
- Oman, T. J., Boettcher, J. M., Wang, H., Okalibe, X. N., and van der Donk, W. A.; *Nature Chemical Biology* **(2011)**, *7*, 78-80
- Pachamuthu, K., and Schmidt, R. R.; *Chem. Rev.* **(2006)**, *106*, 160-187
- Panintrarux, C., Adachi, S., Araki, Y., Kimura, Y., and Matsuno, R.; *Enzyme Microb. Technol.* **(1995)**, *17*, 32-40
- Perugino, G., Cobucci-Ponzano, B., Rossi, M., and Moracci, M.; *Advanced Synthesis & Catalysis* **(2005)**, *347*, 941-950
- Perugino, G., Trincone, A., Rossi, M., and Moracci, M.; *Trends Biotechnol.* **(2004)**, *22*, 31-37
- Pfaffli, P. J., Hixson, S. H., and Anderson, L.; *Carbohydr. Res.* **(1972)**, *23*, 195-206
- Rajan, S. S., Yang, X. J., Collart, F., Yip, V. L. Y., Withers, S. G., Varrot, A., Thompson, J., Davies, G. J., and Anderson, W. F.; *Structure* **(2004)**, *12*, 1619-1629
- Rao, F. V., Dorfmueller, H. C., Villa, F., Allwood, M., Eggleston, I. M., and van Aalten, D. M. F.; *Embo J.* **(2006)**, *25*, 1569-1578
- Rao, V., Cui, T., Guan, C. D., and Van Roey, P.; *Protein Sci.* **(1999)**, *8*, 2338-2346
- Rosnoblet, C., Peanne, R., Legrand, D., and Foulquier, F.; *Glycoconjugate J.* **(2013)**, *30*, 23-31
- Rye, C. S., and Withers, S. G.; *Curr. Opin. Chem. biol.* **(2000)**, *4*, 573-580
- Salamone, S., Guerreiro, C., Cambon, E., Andre, I., Remaud-Simeon, M., and Mulard, L. A.; *Chem. Commun. (Cambridge, U.K.)* **(2015)**, *51*, 2581-2584
- Sardzik, R., Both, P., and Flitsch, S. L.; *Nature Chemical Biology* **(2011)**, *7*, 69-70
- Schmaltz, R. M., Hanson, S. R., and Wong, C. H.; *Chem. Rev.* **(2011)**, *111*, 4259-4307
- Shim, J. H., Chen, H. M., Rich, J. R., Goddard-Borger, E. D., and Withers, S. G.; *Protein Eng., Des. Sel.* **(2012)**, *25*, 465-472
- Spiro, R. G.; *Glycobiology* **(2002)**, *12*, 43R-56R

- Stepper, J., Shastri, S., Loo, T. S., Preston, J. C., Novak, P., Man, P., Moore, C. H., Havlicek, V., Patchett, M. L., and Norris, G. E.; *FEBS Lett.* **(2011)**, 585, 645-650
- Strohmeier, G. A., Pichler, H., May, O., and Gruber-Khadjawi, M.; *Chem. Rev.* **(2011)**, 111, 4141-4164
- Thibodeaux, C. J., Melancon, C. E., and Liu, H. W.; *Angew. Chem., Int. Ed. Engl.* **(2008)**, 47, 9814-9859
- Thompson, J., Pikis, A., Ruvinov, S. B., Henrissat, B., Yamamoto, H., and Sekiguchi, J.; *J. Biol. Chem.* **(1998)**, 273, 27347-27356
- Toone, E. J., Simon, E. S., Bednarski, M. D., and Whitesides, G. M.; *Tetrahedron* **(1989)**, 45, 5365-5422
- van Rantwijk, F., Woudenberg-van Oosterom, M., and Sheldon, R. A.; *J. Mol. Catal. B: Enzym.* **(1999)**, 6, 511-532
- Varki, A.; *Glycobiology* **(1993)**, 3, 97-130
- Varrot, A., Yip, V. L. Y., Li, Y. S., Rajan, S. S., Yang, X. J., Anderson, W. F., Thompson, J., Withers, S. G., and Davies, G. J.; *J. Mol. Biol.* **(2005)**, 346, 423-435
- Venugopal, H., Edwards, P. J. B., Schwalbe, M., Claridge, J. K., Libich, D. S., Stepper, J., Loo, T., Patchett, M. L., Norris, G. E., and Pascal, S. M.; *Biochemistry* **(2011)**, 50, 2748-2755
- Vic, G., and Crout, D. H. G.; *Tetrahedron-Asymmetry* **(1994)**, 5, 2513-2516
- Vic, G., and Thomas, D.; *Tetrahedron Lett.* **(1992)**, 33, 4567-4570
- Viebel, S.; *Enzymologia* **(1936)**, 1, 124-132
- Villalonga, M. L., Diez, P., Sanchez, A., Gamella, M., Pingarron, J. M., and Villalonga, R.; *Chem. Rev.* **(2014)**, 114, 4868-4917
- Vocado, D. J., Mayer, C., He, S. M., and Withers, S. G.; *Biochemistry* **(2000)**, 39, 117-126
- Vocado, D. J., Withers, S. G., Ernst, B., Hart, G. W., and Sinaý, P.; *Carbohydr. Chem. Biol.* **(2008)**, 29, 724-844
- Wada, J., Honda, Y., Nagae, M., Kato, R., Wakatsuki, S., Katayama, T., Taniguchi, H., Kumagai, H., Kitaoka, M., and Yamamoto, H.; *FEBS Lett.* **(2008)**, 582, 3739-3743
- Wang, H., Oman, T. J., Zhang, R., Garcia De Gonzalo, C. V., Zhang, Q., and van der Donk, W. A.; *J. Am. Chem. Soc.* **(2014)**, 136, 84-87
- Wang, H., and van der Donk, W. A.; *J. Am. Chem. Soc.* **(2011)**, 133, 16394-16397
- Wardrop, D. J., and Waidyarachchi, S. L.; *Nat. Prod. Rep.* **(2010)**, 27, 1431-1468
- Watts, A. G., Damager, I., Amaya, M. L., Buschiazzi, A., Alzari, P., Frasc, A. C., and Withers, S. G.; *J. Am. Chem. Soc.* **(2003)**, 125, 7532-7533
- Weiss, J. B., Lote, C. J., and Bobinski, H.; *Nat. New. Biol.* **(1971)**, 234, 25-26
- Williams, S. J., and Withers, S. G.; *Carbohydr. Res.* **(2000)**, 327, 27-46
- Witczak, Z. J., and Culhane, J. M.; *Appl. Microbiol. Biotechnol.* **(2005)**, 69, 237-244
- Witczak, Z. J., Sun, J. M., and Mielguj, R.; *Bioorg. Med. Chem. Lett.* **(1995)**, 5, 2169-2174
- Woudenberg-van Oosterom, M., van Belle, H. J. A., Van Rantwijk, F., and Sheldon, R. A.; *J. Mol. Catal. B: Enzym.* **(1998)**, 134, 267-274
- Yang, F., Lian, G., and Yu, B.; *Carbohydr. Res.* **(2010)**, 345, 309-314
- Zhu, X., Haag, T., and Schmidt, R. R.; *Org. Biomol. Chem.* **(2004)**, 2,



# **Chapitre 1**

## **Développement d'une banque de glycosides hydrolases natives : caractérisation et valorisation**







## **Chapitre 1 – Développement d'une banque de glycoside hydrolases natives : caractérisation et valorisation**

### **I. Stratégie de production : Du choix d'un microorganisme vers la conception d'une banque de glycosidases natives**

Les glycoside hydrolases constituent la 2<sup>ème</sup> classe d'enzymes impliquée dans la biosynthèse et le métabolisme des glycanes en catalysant l'hydrolyse de liaisons glycosidiques avec rétention ou inversion de la configuration anomérique. Pendant de nombreuses années l'intérêt porté aux glycosidases s'est focalisé sur leur utilité dans la dégradation de produit issu de la biomasse ou dans l'industrie agro-alimentaire.(Cairns, **2010**) Elles ont cependant prouvé leur potentiel pour la synthèse de dérivés glycosylés par hydrolyse inverse ou transglycosylation, apparaissant alors comme une alternative de choix aux procédés de synthèse classiques qui nécessitent de nombreuses étapes de protection/déprotection et de purification des sous-produits ce qui conduit dans la plupart des cas à de faibles rendements.(Bojarova, **2009**; Wang, **2009**) Une large gamme de GH est à l'heure actuelle disponible et sont de manière générale plus facile d'accès puisque ce sont des protéines majoritairement cytosoliques,(Meynial-Salles, **1996**) à l'inverse des GT, protéines généralement membranaires, dont la production et la purification plus complexes entraînent souvent une perte de l'activité catalytique.(Meynial-Salles, **1996**; Toone, **1989**) Enfin les glycosidases ne requièrent pas l'utilisation de nucléotide-sucres qui représentent une limite en terme de coût par rapport aux GT.(Gantt, **2011**) En tenant compte de l'ensemble de ces considérations, nous nous sommes naturellement orientés vers la technologie des glycosidases et sur leur potentiel pour réaliser notre projet qui est porté sur la thioglycosylation enzymatique d'acides aminés.

L'objet de ce premier chapitre consiste à décrire la stratégie mise en place pour obtenir une banque de glycosidases sauvages qui servira de base pour la conception des mutants thioglycoligases correspondants. Les outils de biologie moléculaire développés depuis ces dernières années offrent une variété de modes de production d'enzymes recombinantes. Le choix doit donc être judicieux pour favoriser une production efficace des protéines et faciliter leur purification. Ces différents aspects seront abordés au cours de ce chapitre. Après clonage et expression chez *E. coli* Rosetta-(DE3) les enzymes ainsi produites ont été caractérisées afin de définir leur spécificité ainsi que leurs constantes cinétiques ( $K_M$  et  $k_{cat}$ ). Enfin, dans un souci de valorisation des différents biocatalyseurs générés au laboratoire, le potentiel de transglycosylation de ces enzymes a été évalué et les meilleurs candidats ont été exploités pour la synthèse de glycosides glycérol, composés très largement utilisés notamment dans le domaine cosmétique.(Mladenoska, **2008**)

## 1. *Dictyoglomus thermophilum*

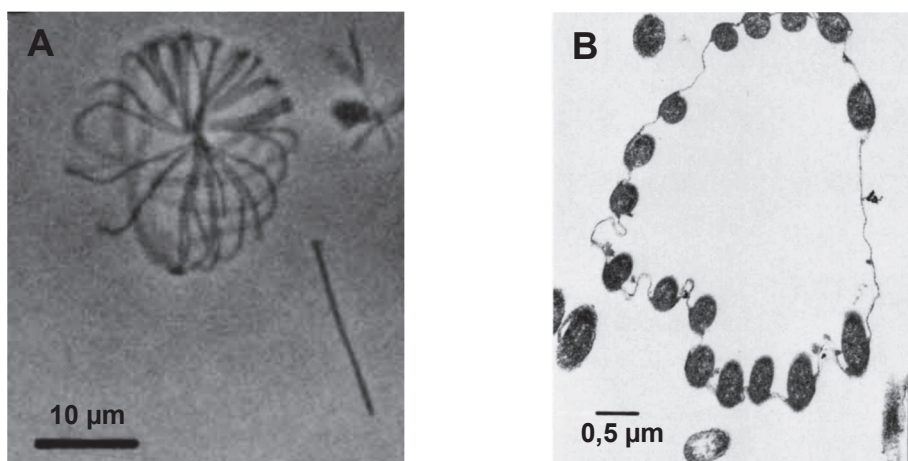
De manière générale les qualités recherchées lors du design d'un biocatalyseur, notamment pour envisager son utilisation à l'échelle industrielle, sont : i) la facilité de production, ii) la robustesse (résistance à des conditions de température, pression et aux solvants) et iii) la réutilisation sur plusieurs cycles catalytiques. D'un point de vue historique, les premières glycosidases exploitées en synthèse étaient extraites de plantes (Jack bean, amande) ou de bactéries, telles qu'*E. coli* ou *Aspergillus oryzae*, néanmoins leur utilisation était limitée par de faible rendement de production du à leur manque de stabilité vis-à-vis des solvants organiques.(Dion, **1999**)

Dans les années 1970, l'apparition des méthodes de séquençage d'ADN proposées par Gilbert (Maxam, **1977**) et Sanger (Sanger, **1978**) couplée au développement des techniques de biologie moléculaire a favorisé l'accès au génome d'une multitude de microorganismes (bactéries, champignons,...) ouvrant donc le champ d'investigation à une large gamme de nouvelles glycosidases plus performantes. Ainsi dans l'industrie du papier, l'utilisation de xylanases ou cellulases thermophiles a permis de mettre en évidence leur résistance face aux effets dénaturants des conditions de hautes pressions et de hautes températures mais aussi et surtout l'augmentation de leur activité d'hydrolyse.(Li, **2015**; Zeikus, **1979**)

De la même manière, pour la synthèse de glycosides d'alkyle les glycosidases thermophiles se sont révélées être des biocatalyseurs d'avenir. Comme décrit précédemment, la réaction de transglycosylation est une réaction sous contrôle cinétique qui s'effectue à de fortes concentrations en alcool et une faible proportion en eau, l'efficacité de la réaction va donc dépendre en partie des propriétés du catalyseur. Les études réalisées sur diverses glycosidases provenant de micro-organismes tels que *Thermus thermophilus*.(Dion, **1999**) *Thermotoga neapolitana*(Turner, **2007**) ou *Dictyoglomus thermophilum*(Zou, **2012**) ont toutes présentées des activités de transglycosylation supérieures à leurs analogues mésophiles grâce notamment à leur résistance aux solvants organiques.

En considérant les qualités présentées par les enzymes thermorésistantes, nous avons orienté notre sélection vers un micro-organisme extremophile. De nombreuses possibilités étaient alors à notre portée (*Thermotoga maritima*, *Pyrococcus furiosus*, *Thermus thermophilum*,...) et pour restreindre notre choix nous avons décidé d'opter pour un micro-organisme encore peu étudié et dont le génome séquencé présente une large gamme de GH. En suivant ce raisonnement la bactérie thermophile *Dictyoglomus thermophilum* s'est révélée être une candidate intéressante, son génome entièrement séquencé en 2014 est disponible commercialement. (Coil, **2014**; Saiki, **1985**)

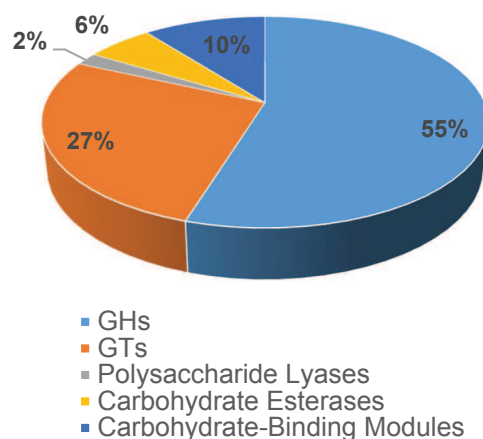
Cette bactérie anaérobie, en forme de bâtonnet (Figure 19), découverte en 1985 au Japon dans des sources chaudes volcaniques possède sur son génome une grande variété de CAZymes (Figure 20) répertoriées dans la classification CAZy dont de nombreuses GT et GH. A l'heure actuelle parmi les 57 glycosidases potentielles (provenant de 31 familles différentes) exprimées chez *D. thermophilum* seulement quelques représentantes sont à l'étude.



**Figure 19 :** (A) cellules de *D. thermophilum* regroupées sous forme de sphère ;  
(B) coupe des cellules en forme de bâtonnet. (Saiki, **1985**)

Parmi ces dernières, la xylanase GH11 a suscité beaucoup d'intérêt (Mathrani, **1992**) ; récemment les études menées sur cette dernière, alors identifiée XYNB, et son mutant XYNB-DS (présentant un pont disulfure) ont démontré leur stabilité sous de fortes conditions de pression et de température (500 MPa, 80°C) ainsi qu'une augmentation de leurs activités sous ces mêmes conditions. (Li, **2015**) En 2013 une endo-glucanase de la famille GH5 (Cel5H) a été caractérisée et présente une forte activité sur une gamme de température de 50 à 85°C en condition acide de pH (4,6 – 5,6) et tolérante à de fortes concentrations en sel. (Shi, **2013**) Si ces enzymes sont étudiées pour leur activité d'hydrolyse (génération de bio-ethanol,...) une seule glycosidase de *D. thermophilum* a été décrite en tant que transglycosylase.

Dans un souci de proposer de meilleurs biocatalyseurs pour la préparation de dérivés glycosylés Zou et coll. (Zou, **2012**) ont produit une  $\beta$ -D-glucosidase, DTGH, et ont comparé son activité de transglucosylation avec diverses glucosidases thermophiles et la  $\beta$ -glucosidase commerciale extraite d'amande. L'équipe a ainsi démontré que DTGH était capable de transglycosyler des alcools de longueurs de chaîne variables (Méthyl - Octyl) mais que contrairement aux autres glucosidases décrites, son activité de synthèse était conservée même à très haute température (100 °C). Les essais réalisés en milieu biphasique (eau/octanol) et à une haute concentration en glucose ont souligné les propriétés de stabilité remarquables et unique de cette enzyme. Cette glycosidase est le premier exemple d'enzyme thermophile présentant de telles caractéristiques pour la synthèse de dérivés glycosylés. Au vu des propriétés décrites pour les GH issues de *D. thermophilum*, notre choix de bâtir une banque de biocatalyseurs à partir du génome de cette bactérie n'en fut que conforté.



**Figure 20** : Répartition des Carbohydrate-Active EnzYmes potentielles chez *D. thermophilum* (<http://www.cazy.org/> , 07/2015).

## 2. De la sélection des gènes à l'utilisation des outils de biologie moléculaire

En étudiant les différentes glycosidases potentielles exprimées chez *D. thermophilum* (souche H-6-12 ATCC 35947, DSM), la classification CAZy s'est avérée être un précieux outil car l'ensemble des GH du microorganisme ont été répertoriées dans leur famille d'appartenance. Ainsi en cherchant une  $\beta$ -D-glucosidase de la famille GH1, reconnue pour leur grande versatilité vis-à-vis du substrat, deux gènes candidats (situés sur les loci *dicth\_0359* et *dicth\_1689* du génome) étaient disponibles et notre choix s'est porté sur la glucosidase codée par *dicth\_0359* (48 % d'homologie de séquence avec le gène de la  $\beta$ -D-Glc de *Caldilinea aerophila*) car la seconde (DTGH) avait déjà été étudiée. (Zou, **2012**) En suivant ce processus quatre autres gènes de *D. thermophilum* ont été sélectionnés formant ainsi notre banque virtuelle de cinq glycosidases. Une comparaison par homologie des séquences protéiques a été réalisée par un outil d'alignement BLAST du serveur Uniprot (Basic Local Alignment Search Tool, <http://www.uniprot.org/blast/>) afin déterminer les protéines les plus proches (Tableau 2).

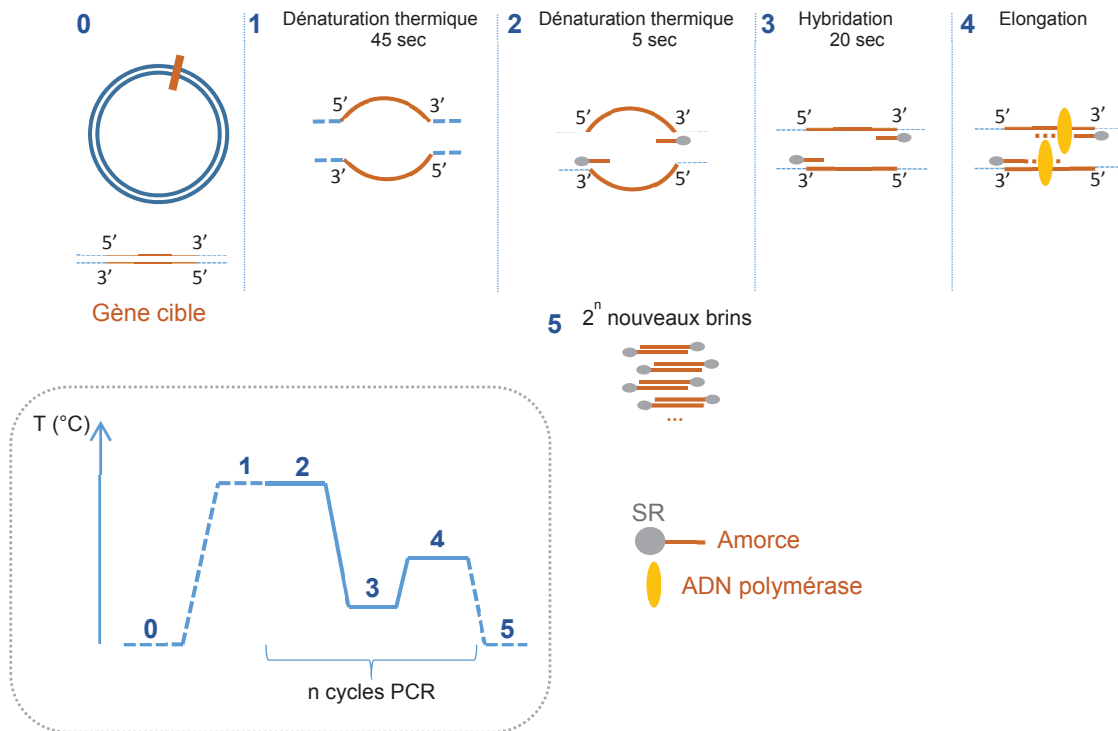
Locus	Annotation	Code Uniprot	Famille GH	Séquence proche	Identité (%)	Structure Proche (PDB)	Identité (%)
<i>dicth_0043</i>	$\alpha$ -L-fucosidase	B5YAV7	GH29	S0EVD8 $\alpha$ -L-fucosidase	90		
<i>dicth_0289</i>	$\alpha$ -L-rhamnosidase	B5YC64	GH78	F3ZY56 $\alpha$ -L-rhamnosidase	52	3CIH	21
<i>dicth_0359</i>	$\beta$ -D-glucosidase	B5YCI2	GH1	O58I04 $\beta$ -D-glucosidase	45	1VFF	44
<i>dicth_1638</i>	$\alpha$ -D-glucosidase	B5YAI2	GH4	Q9WZL1 $\alpha$ -D-glucosidase	69	1VJT	69
<i>dicth_1692</i>	$\beta$ -D-mannosidase	B5YAN4	GH2	Q8AAK6 $\beta$ -D-mannosidase	37	2WBK	37

**Tableau 2 :** Gènes cibles de *D. thermophilum* codant pour de potentielles glycosidases et leur analogies avec des protéines déjà caractérisées.

La production de protéines recombinantes à partir du génome d'un micro-organisme suit un processus en quatre étapes : i) amplification du gène codant pour la protéine d'intérêt par PCR, ii) insertion et clonage du gène dans un vecteur plasmidique, iii) transformation de cet ADN recombinant chez un hôte d'expression (bactéries, levures, cellules de mammifères,...) et iv) la surexpression de la protéine cible. Afin de produire notre banque de GH de *D. thermophilum*, le plasmide pET28-a(+) a été retenu comme vecteur de clonage, tandis que la machinerie cellulaire d'*E. coli* Rosetta-(DE3) sera exploitée pour la surexpression des protéines.

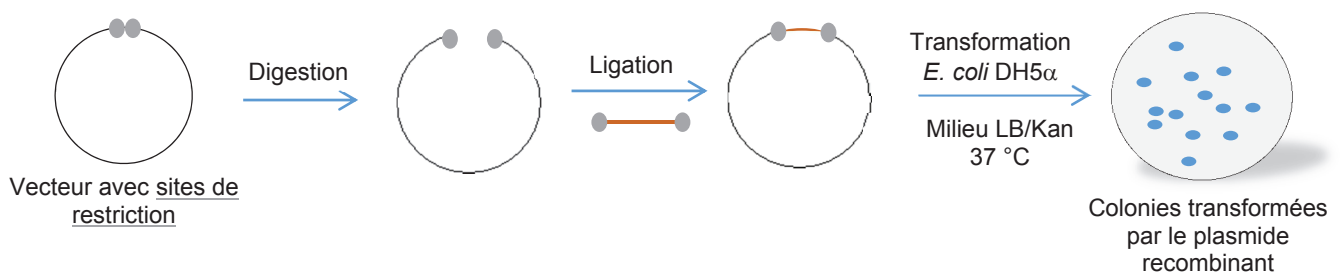
### 2.1. Plasmide de clonage pET28-a(+)

L'amplification des gènes sélectionnés à partir du génome de *D. thermophilum* par PCR nécessite le design des couples d'amorces appropriés permettant d'encadrer la séquence en nucléotides sur les 2 brins d'ADN (sens et anti-sens) et une ADN polymérase permettant de reconstruire de manière fidèle le gène cible. Les amorces sont construites de manière à incorporer un site de restriction (SR) à chaque extrémité du gène amplifié (Figure 21).



**Figure 21 :** Amplification d'un gène cible par PCR à partir de l'ADN génomique d'un micro-organisme.

La présence de ces sites de restriction sur le gène permet l'incorporation de ce dernier sur la séquence ADN du vecteur plasmidique. pET28-a(+) possède des nombreux site de restriction, ces courtes séquences de nucléotides sont reconnues comme sites de coupure d'ADN par des endonucléases spécifiques, aussi appelées enzymes de restriction. Après avoir amplifié le gène d'intérêt, encadré par la séquence en nucléotides de deux SR (reconnus par exemple par les enzymes *NotI* et *NheI*), le plasmide qui possède également ces courtes séquences peut être ouvert spécifiquement par digestion *via* les enzymes *NotI* et *NheI* (Figure 22). Le gène cible est par la suite incorporé sur la séquence de pET28-a(+) par ligation enzymatique. Le plasmide recombinant est alors transformé chez un hôte de clonage, ici *E. coli* DH5 $\alpha$  où il peut se répliquer en grande quantité. pET28-a(+) possède un gène de résistance à la Kanamycine (Kan), codant pour une kinase qui inactive l'antibiotique après phosphorylation. Ce gène sert ici d'outil de sélection car seules les bactéries *E. coli* DH5 $\alpha$  transformées par le plasmide recombinant pourront survivre à un milieu contenant l'antibiotique Kan.

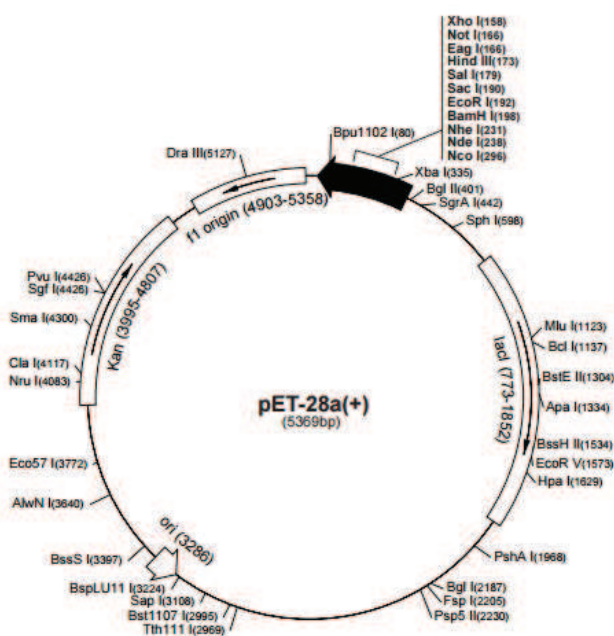


**Figure 22 :** Insertion du gène cible sur pET28-a(+) et transformation chez *E. coli* DH5 $\alpha$ .



Tout vecteur de clonage se doit de faciliter le clonage du gène cible mais aussi la surexpression et la purification de la protéine correspondante. Le vecteur pasmidique pET28-a(+) (Figure 23) présente également l'avantage de posséder sur sa séquence deux gènes clés :

- Un gène codant pour le promoteur T7 nécessaire à la fixation de l'ARN polymérase T7 (de l'hôte d'expression) et donc à la transcription du gène d'intérêt.
- Un gène codant pour une protéine de fusion (Fusion Tag) qui permet de faciliter la détection et la purification des protéines. Dans le cas de pET28-a(+) cette protéine de fusion est constituée de 6 motifs Histidine (His•Tag) qui se fixe en fonction du clonage sur l'extrémité N-terminale ou sur les deux extrémités N-term et C-term de la protéine cible. La purification de cette dernière peut alors se faire sur colonne d'affinité  $\text{Ni}^{2+}$ -NTA (Nickel nitriloacetic acid) via une chélation aux ions  $\text{Ni}^{2+}$ .



**Figure 23 :** Cartographie du plasmide de clonage pET28-a(+).

## 2.2. Vecteur d'expression : *E.coli* Rosetta-(DE3)

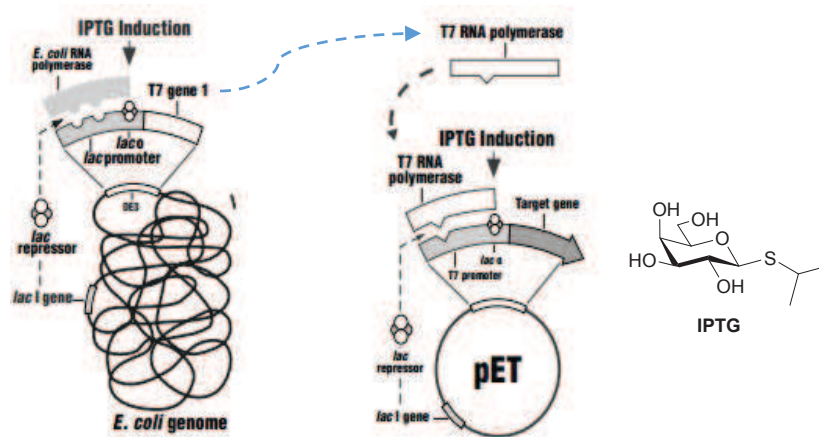
Tout comme les vecteurs de clonage il existe différents systèmes d'expression de protéines parmi lesquels se retrouvent les bactéries (Gram +/-), les levures, les champignons ou encore les cellules eucaryotes provenant d'insectes ou de mammifères. Chaque système présente des avantages et des inconvénients selon le type de protéines à produire. Ainsi les systèmes utilisant des bactéries s'avèrent beaucoup moins coûteux et plus faciles à mettre en place que les systèmes impliquant des cellules de mammifères mais ne permettent pas d'apporter de modifications post-traductionnelles sur les protéines.

A l'heure actuelle encore peu de bactéries pouvant générer des modifications post-traductionnelles ont été répertoriées. Aucune donnée à ce sujet n'est accessible concernant



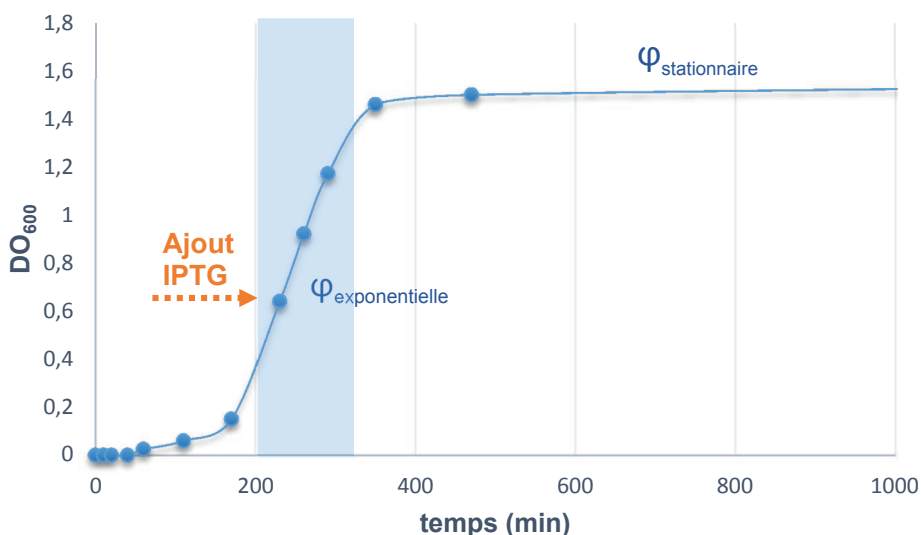
*D. thermophilum*, l'emploi de systèmes eucaryotes s'est donc révélé inutile et dans un souci de facilité de production le système bactérien d'*E.coli* a été retenu. Plus de 300 souches de cette bactérie sont disponibles, parmi lesquelles BL21 (DE3) qui est la plus utilisée pour la production de protéines recombinantes. Nous avons cependant porté notre choix sur une souche dérivée de BL21 : Rosetta-(DE3) puisqu'elle présente sur son génome à la fois le gène de l'ARN polymérase T7 (Schéma 23), compatible avec l'utilisation du plasmide pET28-a(+), ainsi que le plasmide pRARE, ce plasmide permettant de produire les 10 ARNt peu présents chez *E. coli*. La souche Rosetta ainsi « armée » est capable de traduire l'ensemble des codons d'acides aminés (AAs) provenant du gène cible hétérologue (d'origine procaryote ou eucaryote). La souche Rosetta-(DE3) contient également un gène de résistance à l'antibiotique Chloramphénicol (Cam), codant pour une acétyltransférase, qui sert d'outil de sélection. Ainsi après transformation du plasmide recombinant chez Rosetta-(DE3) les bactéries sont déposées sur un milieu nutritif solide en présence des deux antibiotiques Kan et Cam à 37 °C. Dans ces conditions seules les bactéries Rosetta-(DE3) contenant le plasmide recombinant pourront se développer.

La surexpression de la protéine d'intérêt est initiée lors de la croissance bactérienne par l'ajout d'un inducteur non métabolisable par la machinerie cellulaire d'*E. coli*, l'isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Ainsi lors de l'induction, la transcription du gène T7 est activée chez *E. coli*. L'ARN polymérase T7 résultante peut alors se fixer au promoteur T7 présent sur la séquence de pET28-a(+) et transcrire le gène de la protéine cible (Schéma 23).



**Schéma 23** : Induction en cascade de l'IPTG. (Extrait du manuel Novagen pET System 11<sup>ème</sup> édition).

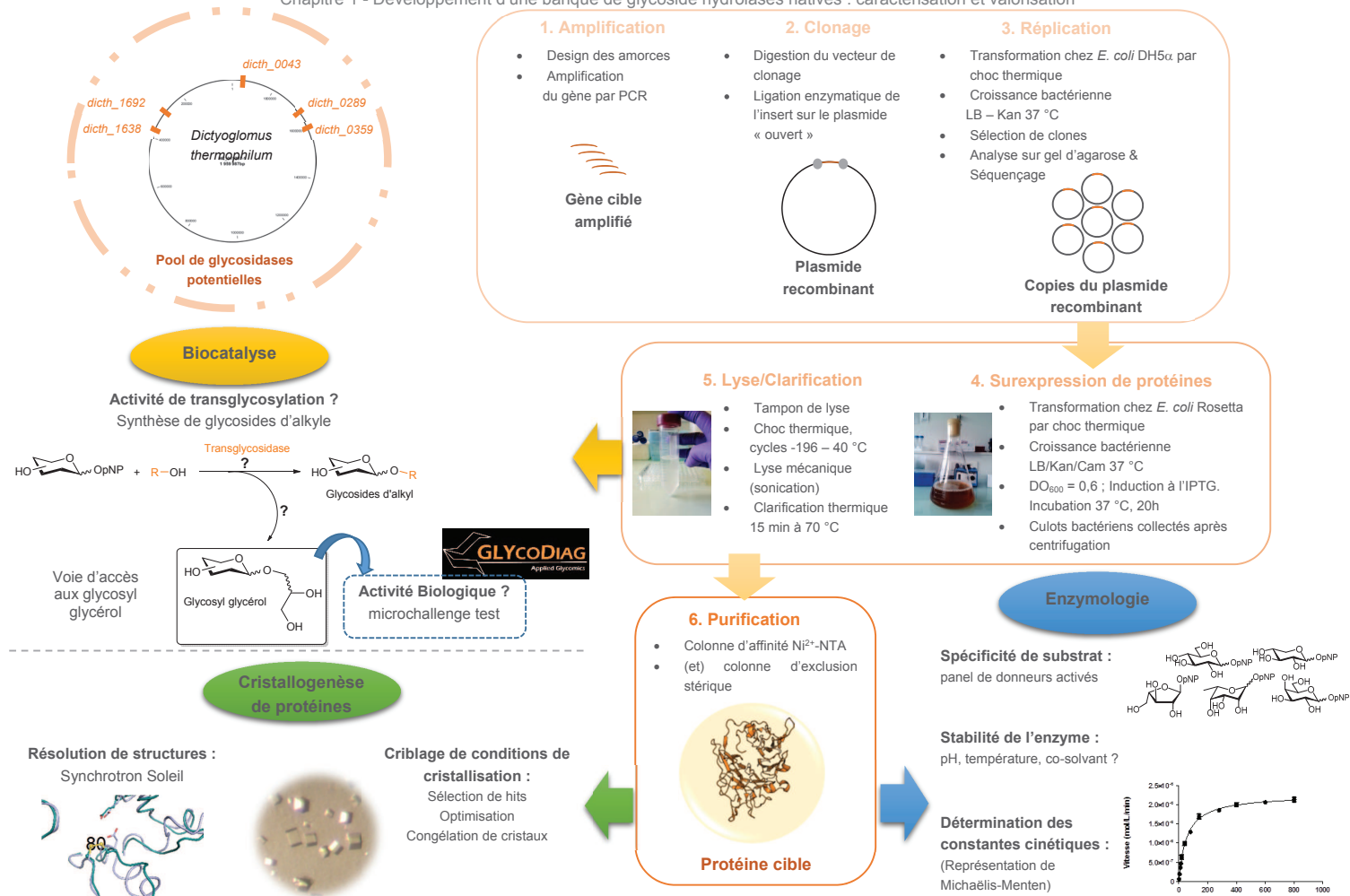
Pour assurer une production maximale de la protéine recombinante, la croissance bactérienne, réalisée en milieu LB – Kan/Cam, est suivie par mesure de la densité optique (DO) et l'ajout de l'inducteur est effectué lorsque la phase exponentielle de croissance bactérienne est atteinte ( $DO_{600} = 0,5 - 0,7$ ) (Figure 24).



**Figure 24 :** Courbe de croissance bactérienne.

Ainsi, tous les éléments ont été réunis pour la production de la banque de glycosidases recombinantes sélectionnées, depuis l'origine du microorganisme en passant par le système de clonage et d'expression des protéines. Toutes les enzymes ont été clonées et produites en suivant le même protocole décrit dans la partie expérimentale. Chaque enzyme a par la suite été étudiée afin de déterminer i) sa spécificité vis-à-vis du substrat, ii) ses conditions optimales d'activité (température et pH), iii) son aptitude à tolérer des solvants organiques. Des essais de cristallogenèse ont également été réalisés dans le but d'obtenir les structures tridimensionnelles de ces biocatalyseurs pour connaître la géométrie du site actif en distinguant les résidus catalytiques, des résidus responsables de la fixation du substrat. Enfin le potentiel de transglycosylation des enzymes sera évalué en utilisant une variété d'accepteur pour la synthèse de glycosides d'alkyle. Cette stratégie de production et caractérisation des glycosidases recombinantes de *D. thermophilum* peut être résumée selon le schéma décrit dans la page suivante de ce manuscrit (Schéma 24, Stratégie de production et de caractérisation d'une banque de GH thermostables).

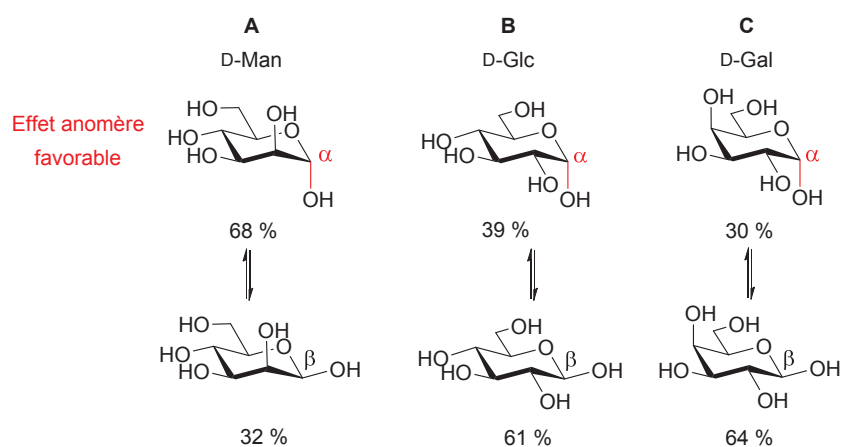
## Chapitre 1 - Développement d'une banque de glycoside hydrolases natives : caractérisation et valorisation



## II. Caractérisation de nouvelles glycosidases recombinantes

### 1. DtMan, $\beta$ -D-mannosidase de *D. thermophilum*

Dans la Nature, les liaisons de type  $\beta$ -D-mannopyranosidiques sont fréquemment rencontrées. Chez les plantes ce sont les constituants de ( $\beta$ -1,4) - oligosaccharides, (Gomes, **2007**) mais ces motifs sont également présents dans les trisaccharides de cœur de *N*-glycoprotéines, sur les polysaccharides antigéniques de levures (Vinogradov, **2000**) et *Salmonelles* (Lindberg, **1988**) ou encore sous forme de glycolipides. (Sugita, **1988**) Les  $\beta$ -mannanases et  $\beta$ -mannosidases (EC 3.2.1.25) sont responsables de leur métabolisme et sont largement utilisées dans l'industrie papetière ou pour la production de carburant par dégradation de la biomasse. (Dhawan, **2007**) De par leur présence sur des séquences antigéniques de pathogènes (i.e., *Salmonelles*) la synthèse de dérivés  $\beta$ -mannosylés biologiquement actifs suscite un fort intérêt, mais reste un exercice périlleux. Par voie chimique de nombreuses méthodes ont été décrites mais reflètent toutes de la difficulté d'obtenir la bonne configuration de liaison. Dans le cas du mannose le groupe hydroxyle 2 est en position axiale, l'effet anomère qui tient compte des facteurs électroniques et stériques est favorable à la configuration  $\alpha$  (à l'inverse du glucose et galactose par exemple, Schéma 25). D'un point de vue stérique, la configuration 1,2-*cis* est moins stabilisante que la configuration 1,2-*trans*. (Drew, **2001**)



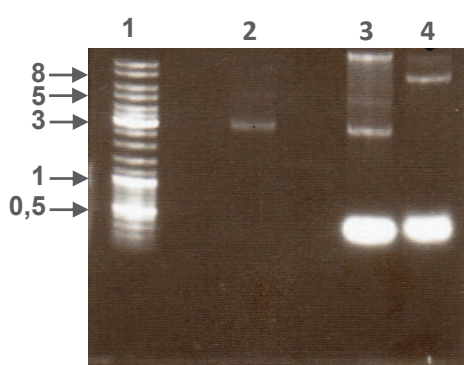
**Schéma 25** : Effet anomère favorable pour le mannose (A), glucose (B) et galactose (C).

A la fin du XXe siècle l'accès à ces composés par voie enzymatique s'est développé avec la caractérisation de  $\beta$ -mannosidases provenant de divers micro-organismes (*Bacteroides thetaiotaomicron*, *Cellulomonas fimi*, *Streptomyces sp*, *Thermotoga neapolitana*,...) présentant des activités de transglycosylation. (Zhang, **2009**) La plupart de ces biocatalyseurs a été répertoriée dans la famille GH2 des glycoside hydrolases, à l'exception de deux mannosidases de *Pyrococcus furiosus* classifiées dans la famille des GH1. (Park, **2011**)

*D.thermophilum* présente sur son génome quatre gènes codant pour de potentielles  $\beta$ -mannosidases ou  $\beta$ -mannanases et à l'heure actuelle aucun d'entre eux n'a été cloné et exprimé chez *E. coli* (<http://www.cazy.org/>, 07/2015). Dans un souci de pouvoir comparer nos

enzymes avec des modèles déjà caractérisés nous avons sélectionné le gène *dicth\_1692* annoté pour une potentielle exo- $\beta$ -mannosidase à rétention de configuration (code Uniprot B5YAN4) appartenant à la famille GH2. Des analyses basées sur l'homologie de séquence en acides aminés à l'aide d'un logiciel de BLAST (Basic Local Alignment Search Tool) ont révélé que la protéine B5YAN4 présentait 37 % d'identité de séquence avec l'unique exo- $\beta$ -mannosidase caractérisée et de structure résolue de GH2, la BtMan2A de *B. tethaiotaomicron*. (Tailford, **2007**) L'alignement de ces 2 séquences est présenté en annexe de ce manuscrit (Annexe 01, p.247).

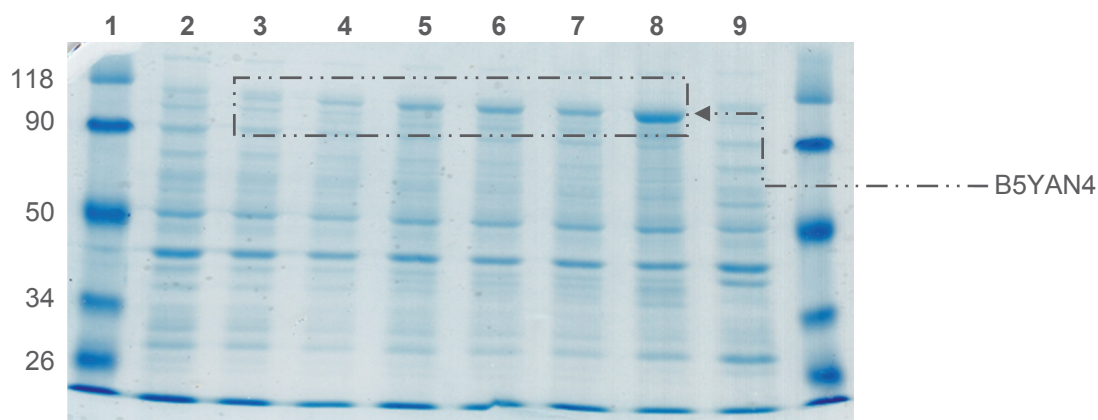
Le gène *dicth\_1692* constitué de 2439 paires de bases (pb), codant pour 813 acides aminés, a été amplifié par PCR (Polymerase Chain Reaction) à partir de l'ADN génomique de *D. thermophilum* H-6-12 ATCC 35947, DSM 3960. Les amorces nécessaires à cette étape ont été construites de manière à présenter les sites de restriction *NheI* et *NotI* sur les extrémités du gène (cf. Partie expérimentale p.195, et produites par l'entreprise Eurofins Genomics (<http://www.eurofinsgenomics.eu/>). La ligation du gène, ou insert, a été réalisée après digestion du vecteur pET28-a(+) par les enzymes de restriction spécifiques *NheI* et *NotI*. Après transformation du plasmide chez *E. coli* DH5 $\alpha$ , deux colonies bactériennes ont pu être isolées, parmi la dizaine de clones, et mises en culture dans un milieu nutritif LB/Kan. L'ADN plasmidique a été extrait et purifié à partir d'un kit commercial (E.N.Z.A Minikit, OmegaBiotech) puis analysé par électrophorèse sur gel d'agarose 0,5 % contenant du SYBRsafe et visualisé aux UVs ( $\lambda = 254$  nm) à l'aide d'un transilluminateur équipé d'une caméra (Figure 25). Parmi les deux clones sélectionnés un seul a présenté la bande de taille attendue ( $\approx 7800$  pb, piste 4). Le second clone (piste 3), quant à lui, possède une bande de taille identique au plasmide vide linéarisé et une deuxième bande de taille supérieure correspondant à l'ADN plasmidique circulaire intact (ou ADN superenroulé) qui n'a pas été dégradé par l'enzyme de digestion avant le dépôt sur gel.



**Figure 25** : Analyse par électrophorèse sur gel d'agarose (1%) du clonage de *dicth\_1692*. (1) marqueur de taille (en kpb) ; (2) plasmide pET28-a(+) vide ( $\approx 5300$  pb) ; (3) clone négatif présentant une bande de la taille du plasmide vide et une bande  $> 7800$  pb; (4) clone positif, une seule bande correspondant à la construction pET28-a(+)/*dicth\_1692*.

L'ADN plasmidique du clone retenu a été séquencé (par l'entreprise Eurofins Genomics), et le résultat de l'alignement gène parent – fils a permis de confirmer le clonage de *dict\_1692* (Annexe 02, p.249). Après transformation chez *E. coli* Rosetta-(DE3), les bactéries ont été mises en culture (37 °C) et l'expression de la protéine cible a été initiée par induction à l'IPTG (0,2 mM concentration finale).

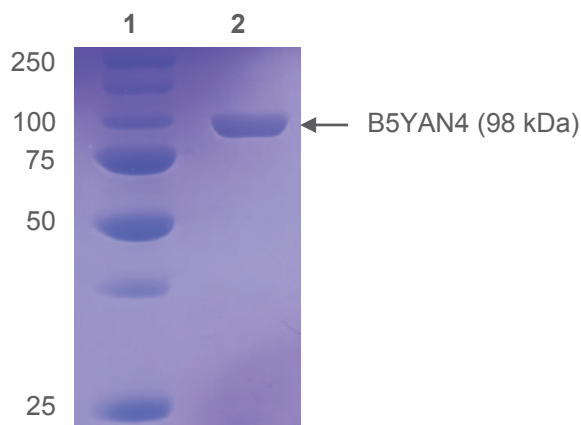
Afin de déterminer les meilleures conditions d'expression de la protéine chez Rosetta-(DE3) un suivi a été réalisé après induction par prélèvement d'échantillons à différents temps t. Le contenu des échantillons a été analysé et comparé à un échantillon blanc non induit sur gel SDS-PAGE. Cette étude révèle d'une part que l'induction à l'IPTG est nécessaire à l'expression de la protéine chez *E. coli* Rosetta-(DE3) et d'autre part que le maximum d'expression est obtenu 20h après l'ajout de l'inducteur à une température constante de 37 °C (Figure 26).



**Figure 26 :** Suivi de l'expression de B5YAN4 chez *E. coli* sur gel SDS-PAGE. (1) Marqueur de taille (en kDa) ; (2), (3), (4), (5), (6), (7) et (8) Echantillon induit à T0, 30 min, 1h, 2h, 3h 4h et 20h après ajout de l'IPTG ; (9) Echantillon blanc non induit.

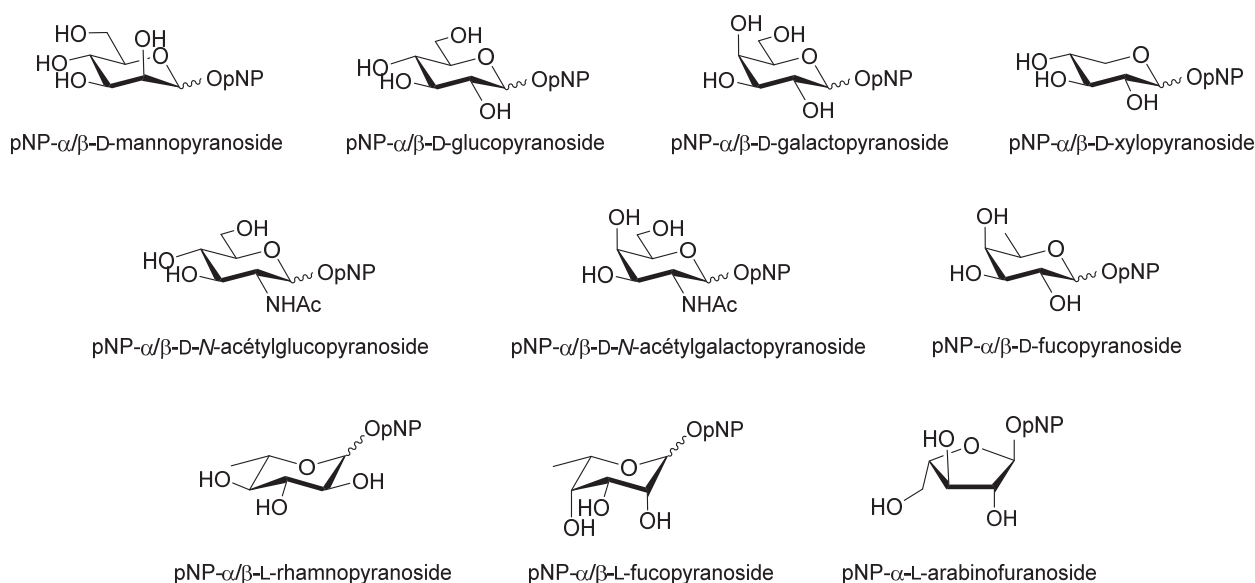
La protéine est extraite des bactéries en suivant la méthode classique de lyse par choc thermique, en réalisant 3 cycles de congélation/décongélation (-196 à 40 °C), et sonication. Après centrifugation le surnageant subit une étape de clarification thermique. Ce procédé consiste à incuber la solution protéique dans un bain à 70 °C pendant 15 min, la protéine de *D. thermophilum* étant thermorésistante elle ne sera pas dégradée contrairement à la grande majorité des protéines d'*E. coli*. L'enzyme est récupérée sous forme pure après purification par chromatographie d'affinité (NiNTA) et analysée sur gel SDS-PAGE. La séquence de B5YAN4 est constituée de 813 acides aminés ce qui représente une masse théorique de 96 kDa, l'analyse sur gel révèle la présence d'une bande unique de taille inférieure à 100 kDa ce qui correspond à la taille de la protéine cible (Figure 27). Après dosage, le rendement de production de DtMan s'élève à 8 mg de protéine par litre de culture bactérienne.





**Figure 27 :** Analyse sur gel d'électrophorèse SDS-PAGE (10 % acrylamide) de la protéine B5YAN4. (1) Marqueur de tailles (en kDa) ; (2) solution récupérée après purification sur colonne de  $\text{Ni}^{2+}$  contenant la protéine B5YAN4 (98 kDa).

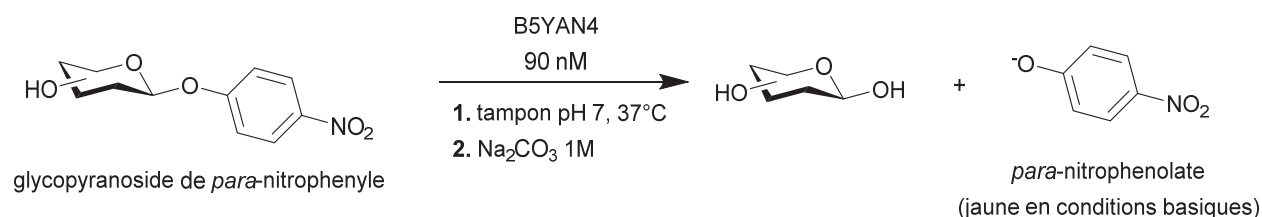
La spécificité de la protéine B5YAN4 vis-à-vis de différents substrats a été déterminée qualitativement en mesurant l'activité d'hydrolyse de l'enzyme sur un panel de glycopyranosides de *para*-nitrophényle (pNP-glycopyranosides) (Figure 28).



**Figure 28 :** Panel de pNP-glycopyranosides testés comme substrats des GH de *D. thermophilum*.

Le choix s'est porté sur ce type de substrat car lors de l'hydrolyse enzymatique il y a libération de *para*-nitrophénolate (pNP, à  $\text{pH} > \text{pK}_{\text{a pNP}} = 7,65$ ) composé qui absorbe à une longueur d'onde  $\lambda = 405 \text{ nm}$  et qui confère à la solution une couleur jaune (Schéma 26). Ainsi l'activité de l'enzyme peut facilement être observée visuellement et d'autre part la quantité de pNP relarguée quantifiée par mesure d'absorbance est directement liée à l'activité enzymatique. Une vingtaine de substrats ont été testés en présence de B5YAN4 qui ne s'est révélée active

que sur le pNP- $\beta$ -D-mannopyranoside (pNPMan) confirmant l'activité  $\beta$ -mannosidase de l'enzyme nommée par la suite **DtMan**.

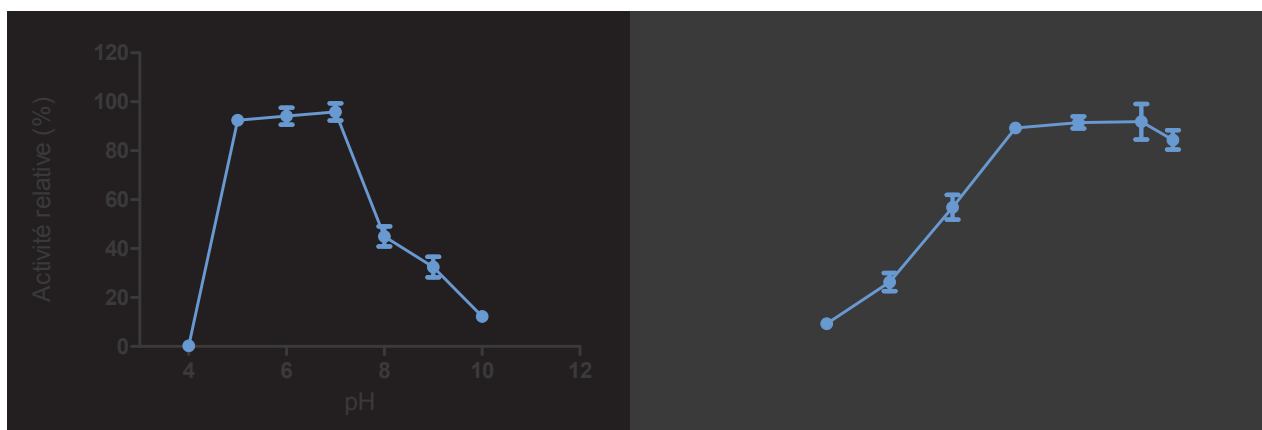


**Schéma 26** : Détermination de la spécificité de substrat de B5YAN4 sur le panel de pNP-sucres.

### 1.1. Conditions optimales de réaction et paramètres cinétiques

Afin de déterminer les constantes cinétiques de l'enzyme, les conditions optimales de réaction ont dans un premier temps été étudiées. Ainsi l'activité d'hydrolyse a été mesurée pour différentes valeurs de pH et de températures. La libération de pNP lors de l'hydrolyse a pu être quantifiée par simple mesure d'absorbance à  $\lambda = 405$  nm, ainsi l'activité catalytique de l'enzyme a pu être définie en rapportant la quantité d'enzyme nécessaire pour la libération d'1  $\mu$ mol de pNP par minute.

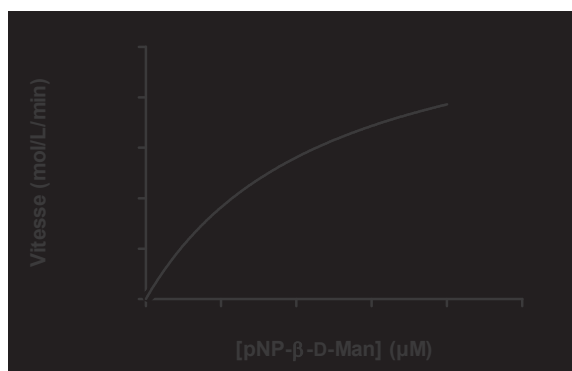
Le maximum d'activité de l'enzyme est atteint à pH 7. A partir de cette valeur maximale, la courbe d'activité relative de l'enzyme en fonction du pH a été établie (**A**, Figure 29). En milieu tamponné pH 5 à 7, 80 % de l'activité enzymatique est conservée, en revanche l'enzyme semble plus sensible pour des milieux alcalins (pH > 9) ou moins de 30 % de l'activité d'hydrolyse est conservée. La même représentation a été utilisée pour évaluer l'influence de la température sur l'activité du biocatalyseur (**B**, Figure 29). De par son caractère thermorésistant, l'enzyme présente une activité maximale pour des températures comprises entre 70 – 95 °C avec un pic d'hydrolyse atteint à 90 °C. Ces valeurs sont en accords avec les  $\beta$ Man thermostables de la famille GH2 décrites dans la littérature,  $pH_{Opt} \approx 7,0$  et  $T_{Opt} = 80 - 100$  °C.(Bauer, **1996**; Duffaud, **1997**; Zhang, **2009**)



**Figure 29** : Détermination des conditions de pH (**A**) et de températures (**B**) optimales.



A partir de ces résultats préliminaires, les paramètres cinétiques de **DtMan** vis-à-vis du pNPMan ont dans un premier temps été calculés dans les conditions optimales de réactivité de l'enzyme, c'est-à-dire à pH 7 (tampon imidazole.HCl 20 mM) et 70 °C. L'activité enzymatique mesurée sur une gamme de concentration en substrat (0,05 – 5 mM) permet de tracer la représentation de Michaëlis-Menten qui donne accès aux paramètres cinétiques de l'enzyme (Figure 30). Afin d'éviter la consommation trop rapide du substrat et assurer le suivi d'hydrolyse sur un temps d'incubation de 30 min, l'enzyme a été introduite à une concentration finale de 0,7 nM.



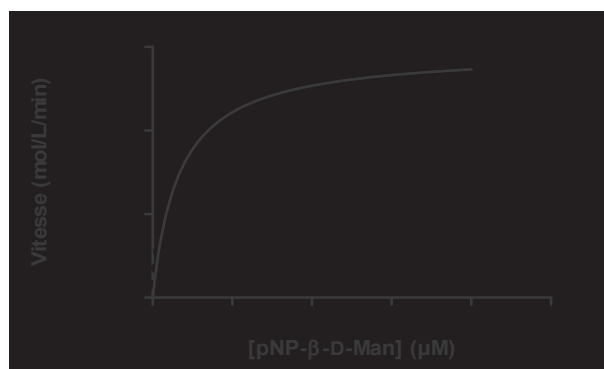
**Figure 30 :** Courbe de Michaëlis-Menten d'activité d'hydrolyse de DtMan en présence d'une gamme de concentration en pNPMan à pH 7 et 70 °C. Les essais ont été réalisés en triplicat.

Les valeurs de  $K_M$  et  $V_{max}$  obtenues pour ce substrat sont de l'ordre de 0,48 mM et 0,01 mM.min<sup>-1</sup>. Ces deux valeurs ont permis de calculer le turnover par minute,  $k_{cat} = 297 \text{ s}^{-1}$ , ainsi que la spécificité de l'enzyme pour le pNPMan,  $k_{cat}/K_M = 619 \text{ mM}^{-1}.\text{s}^{-1}$ . En comparant ces résultats avec les  $\beta$ -mannosidases décrites dans la littérature nous avons pu mettre en évidence que l'affinité de DtMan ( $K_M$ ) pour le pNPMan était du même ordre de grandeur que les valeurs reportées (0,2 – 0,8 mM) excepté pour la  $\beta$ Man de *T. neapolitana*. D'autre part DtMan possède une efficacité catalytique vis-à-vis du pNPMan 15 à 30 fois supérieure aux autres  $\beta$ Man thermostables décrites, mais reste comparable à la BtMan2A avec laquelle elle partage la plus grande homologie de séquence (Tableau 3).

Souche	$K_M$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$ (mM <sup>-1</sup> .min <sup>-1</sup> )
<i>D. thermophilum</i>	0,48 ± 0,19	297 ± 33	619
<i>B. tethaiotaomicron</i> (BtMan2A)(Tailford, 2007)	0,19	128	674
<i>T. maritima</i> (Zhang, 2009)	0,49	8	16
<i>P. furiosus</i> (Bauer, 1996)	0,79	30	38
<i>T. neapolitana</i> (Duffaud, 1997)	3,1	68	22

**Tableau 3 :** Caractéristiques de  $\beta$ -mannosidases provenant de différents microorganismes.

De manière générale, les glycosidases thermostables caractérisées possèdent une forte activité catalytique à haute température. Or l'utilisation de telles conditions de réaction représente une dépense énergétique non négligeable à plus grande échelle qui ne correspond pas aux enjeux environnementaux actuels (limitations des déchets, du coût énergétique,...). Afin d'inscrire notre projet dans une démarche éco-compatible, nous avons tenté de trouver un compromis entre l'efficacité de l'enzyme et des conditions de réaction moins énergivores. Dans cette optique, les paramètres cinétiques d'hydrolyse de DtMan ont été calculés en présence d'une gamme de concentration de pNPMAN (0,05 - 0,8 mM) en solution dans un tampon imidazole.HCl 20 mM pH 7 à 37 °C (Figure 31). Malgré la baisse de la température l'enzyme a dû être introduite en faible quantité (0,9 nM) pour permettre de ralentir son activité et assurer le suivi d'hydrolyse du pNPMAN dans le temps.



**Figure 31** : Courbe de Michaelis-Menten d'activité d'hydrolyse de DtMan (0,9 nM) en présence d'une gamme de concentration en pNPMAN à pH 7 et 37 °C. Les essais ont été réalisés en triplicat.

Dans ces conditions de réaction l'enzyme ne fonctionne qu'à 10 % de son maximum (**B**, Figure 29, p. 74) donc comme nous pouvions le prévoir une diminution nette des constantes catalytiques est observée par rapport à l'étude réalisée à 70 °C (Tableau 4). En revanche il est à noter que malgré la baisse d'activité de DtMan la valeur du ratio  $k_{cat}/K_M$  reste jusqu'à 9 fois supérieure à celles des  $\beta$ Man thermostables caractérisées dans leurs conditions optimales de réaction, mettant ainsi en exergue l'extrême efficacité catalytique de DtMan vis-à-vis du pNPMAN. A ce jour aucune  $\beta$ Man présentant une telle activité à basse température n'a été décrite dans la littérature.

Température	$K_M$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $mM^{-1}.min^{-1}$ )
70 °C	$0,48 \pm 0,19$	$297 \pm 33$	619
37 °C	$0.066 \pm 0.03$	$11 \pm 2$	167

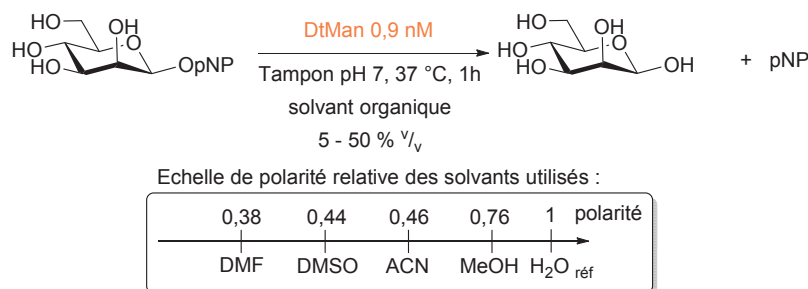
**Tableau 4** : Etude comparative des constantes cinétiques de DtMan à 37 et 70 °C.

Ainsi, la mannosidase thermostable de *D. thermophilum*, DtMan, est caractérisée par une spécificité vis-à-vis des liaisons de type exo- $\beta$ -D-mannopyranosidiques mais également par

une efficacité catalytique nettement supérieure aux  $\beta$ Man thermostables caractérisées et ce quelle que soit les conditions de réactions.

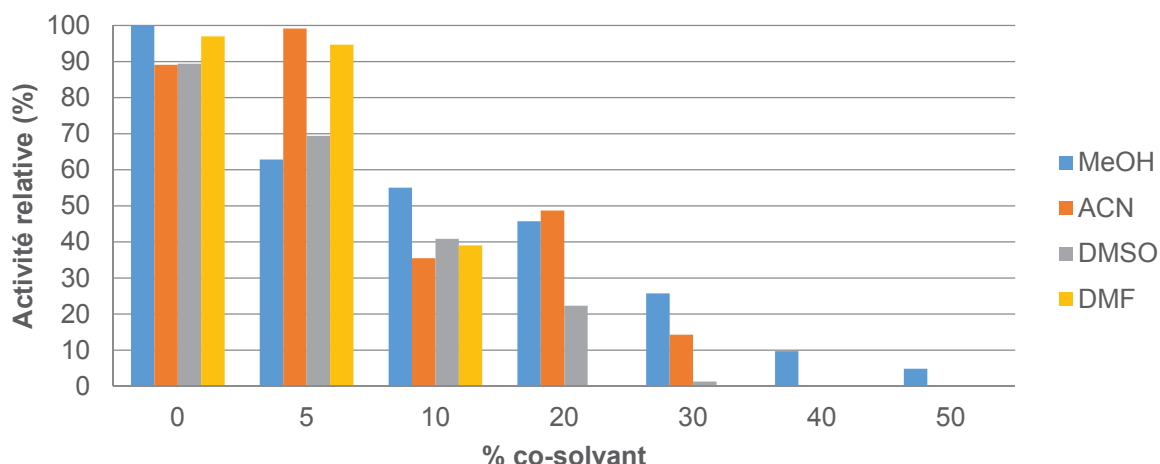
## 1.2. Influence du co-solvant

L'activité de l'enzyme a été mesurée en présence de différents co-solvants. En effet dans les réactions de glycosylation enzymatiques, l'utilisation de solvants organiques peut s'avérer nécessaire pour faciliter la solubilisation d'accepteurs hydrophobes et diminuer l'activité de l'eau. Dans cette optique, différents solvants (plus ou moins polaires) ont été testés dans des proportions variant entre 5 et 50 %  $v/v$  (Schéma 27).



**Schéma 27** : Etude de l'activité d'hydrolyse de DtMan en présence de co-solvant. Echelle de polarité relative extraite de (Reichardt, **2003**).

Pour chaque solvant un suivi cinétique en continu de l'hydrolyse du pNPMan a été réalisé à pH 7 (tampon imidazole.HCl 20 mM) pendant 1h à 37 °C et les résultats obtenus ont été comparés aux échantillons témoins sans solvant organique. En présence de co-solvant l'activité enzymatique de DtMan diminue sans surprise (Figure 32). Néanmoins, 30 % de cette activité est conservée lorsque le milieu réactionnel est composé à 10 % de co-solvant et ce quel que soit le système organique utilisé. Le méthanol permet d'obtenir les meilleurs résultats puisque 5 % de l'activité enzymatique de DtMan sont conservés dans un milieu tampon / MeOH 50:50. Ce co-solvant pourra donc être utilisé préférentiellement dans les essais de glycosylation futurs pour favoriser la solubilité d'accepteurs.



**Figure 32 :** Activité d'hydrolyse de DtMan en présence de différents co-solvants organiques à pH 7. Expérience réalisée sur un essai unique.

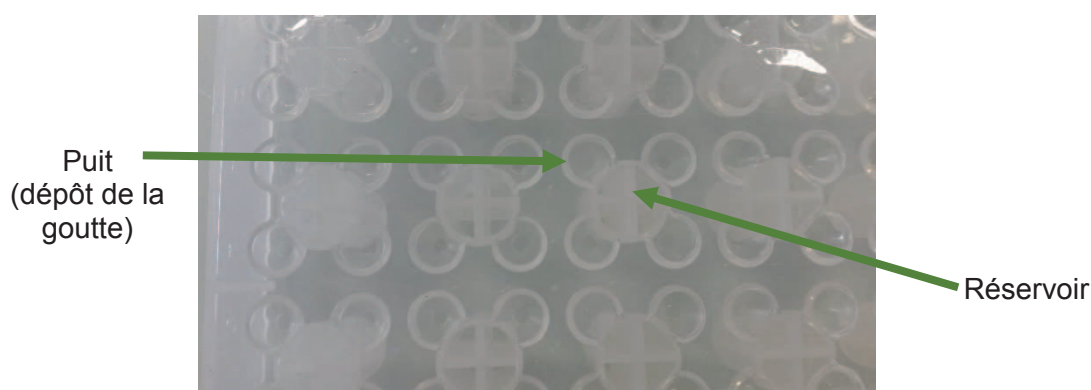
### 1.3. Cristallogénèse de DtMan

La caractérisation d'une enzyme ne s'arrête pas à la simple détermination de ces constantes cinétiques. La création de la classification CAZy à la fin du XX<sup>ème</sup> siècle le prouve puisqu'elle a permis de regrouper de nombreuses enzymes en différentes familles selon leurs similarités structurales et d'activité. Par exemple, les enzymes appartenant au clan-A des GH présentent un mode de repliement commun de type  $(\beta/\alpha)_8$  et agissent toutes selon un mécanisme avec rétention de configuration. Tandis que les GH du clan-I se replient selon un mode  $(\alpha + \beta)$  et catalysent l'hydrolyse de la liaison glycosidique avec inversion de la configuration anomérique. Ainsi la cristallogénèse de protéine est une étape essentielle à la caractérisation d'un nouveau biocatalyseur. A partir des cristaux obtenus, il est possible de résoudre la structure tridimensionnelle de l'enzyme qui est nécessaire à la compréhension des mécanismes moléculaires engagés lors de l'acte catalytique. Dans notre cas l'obtention de cristaux de DtMan nous permettra de confirmer son appartenance à la famille GH2 mais aussi d'obtenir des informations supplémentaires sur la structure de son site actif. La position des acides aminés catalytiques pourra être déterminée ainsi que l'identification des résidus ne jouant pas un rôle direct sur l'acte catalytique mais sur la stabilisation du substrat et/ou de l'intermédiaire glycosyl-enzyme. En termes d'activité, meilleure sera l'affinité de l'enzyme pour son substrat, faible valeur de  $K_M$ , plus grande sera la valeur du ratio  $k_{cat}/K_M$  et par conséquent meilleure sera l'efficacité de l'enzyme.

Cette étude requiert une protéine très pure et concentrée de manière à cribler le plus grand nombre de conditions de cristallisation. Pour ce faire DtMan a été fraîchement produite à partir d'une colonie d' *E. coli* Rosetta-(DE3) inoculée dans un milieu nutritif LB – Kan/Cam et laissée incubée une nuit à 37 °C. Par la suite 6 L de milieu nutritif ont étéensemencés avec cet inoculum de manière à produire, après induction à l'IPTG, la protéine en très grande quantité. Après traitement des extraits cellulaires par lyse, clarification thermique et purification sur colonne d'affinité, la protéine a subi une étape supplémentaire de purification sur colonne d'exclusion stérique. Cette dernière étape permet d'éliminer les résidus ayant pu être

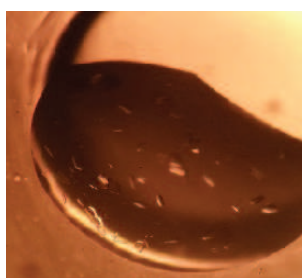
accumulés au cours des traitements précédents, tels que l'imidazole utilisés dans les éluants lors de l'étape de purification par chromatographie d'affinité. Au terme de ce processus une fraction pure de DtMan a été récupérée de concentration 0,45 mg/ml (2 ml). La protéine est concentrée pour obtenir une concentration finale de 9 mg/ml (100  $\mu$ L).

Les conditions de cristallisation de l'enzyme ont été testées à partir du kit commercial Structure Screen 1 (Molecular Dimensions) contenant 100 compositions de solutions. La plaque de cristallisation a été préparée *via* la technique de la goutte assise, en déposant 90  $\mu$ l de solution de cristallisation dans un réservoir et sur le puit associé est déposé une goutte de composition protéine / tampon 1 :1 (volume final de goutte 2  $\mu$ l) (Figure 33).



**Figure 33** : Description d'une boîte de cristallisation, technique de la goutte assise.

Les boîtes sont ensuite recouvertes d'un film adhésif pour favoriser le phénomène de diffusion de vapeur goutte – réservoir et limiter le phénomène d'évaporation, puis laissées au repos à température ambiante. Après 15 jours d'incubation, l'analyse par microscopie des plaques a permis de mettre en évidence 1 hit de cristallisation pour la solution contenant du tampon NaHepes 100 mM pH 7,5 ;  $\text{MgCl}_2$  200 mM et 30 %  $\text{v/v}$  de PEG 400 (Figure 34).



**Figure 34** : Cristaux de DtMan *wt* obtenus en milieu tamponné NaHepes 100 mM pH 7,5 ;  $\text{MgCl}_2$  200 mM et 30 %  $\text{v/v}$  de PEG 400.

A partir de ce premier hit, les conditions ont été modulées de manière à optimiser la formation des cristaux de DtMan. En suivant le processus décrit précédemment une plaque de cristallisation a été préparée à partir de trois concentrations en protéine (5, 10 et 15 mg/ml) et quatre solutions de cristallisation (Tableau 5) selon la méthode d'ensemencement. Dans un

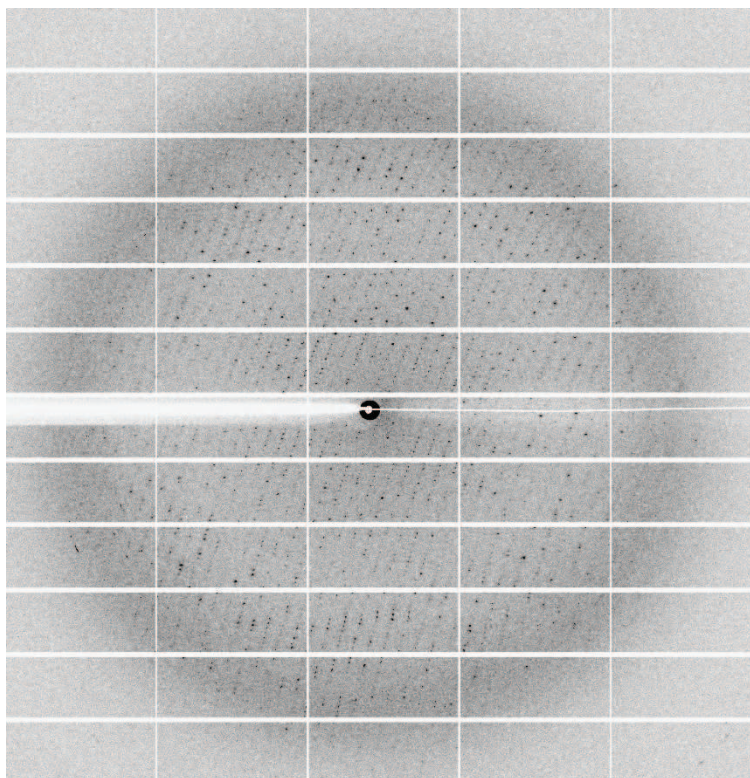
premier temps, la plaque est préparée comme décrit précédemment. Par la suite un cristal obtenu lors du premier essai de cristallisation est récupéré et broyé dans une goutte contenant son milieu de cristallisation. Une moustache de chat est alors plongée dans ce « broyat » de cristal et utilisée pour ensemencher chaque goutte de la plaque de cristallographie. L'avantage de cette méthode est de pouvoir initier plus rapidement la cristallisation des protéines et peut conduire à la formation de plus gros cristaux. La boîte a été laissée au repos à température ambiante.

Solutions de cristallisation		
Précipitant	Tampon	Additif
24 % PEG 400	NaHepes 100 mM pH 7,5	200 mM MgCl <sub>2</sub>
26 % PEG 400	NaHepes 100 mM pH 7,5	100 mM MgCl <sub>2</sub>
32 % PEG 400	NaHepes 100 mM pH 7,5	200 mM MgCl <sub>2</sub>
36 % PEG 400	NaHepes 100 mM pH 7,5	200 mM MgCl <sub>2</sub>

**Tableau 5 :** Conditions de cristallisation testées à partir du premier hit de DtMan *wt*.

Un cristal a été obtenu dans les conditions 26 % PEG 400, NaHepes 100 mM pH 7,5 et MgCl<sub>2</sub> 200 mM. Avant congélation du cristal dans l'azote liquide pour son analyse RX celui-ci doit être plongé dans une solution de cryoprotectant pour éviter sa dégradation. Le PEG 400 peut être utilisé comme système de cryoconservation, ainsi le cristal a été directement congelé dans l'azote liquide et analysé par la suite sur la ligne Proxima 1 du Synchrotron Soleil (<http://www.synchrotron-soleil.fr/>). L'analyse RX s'est avérée concluante car une image de diffraction a été obtenue (résolution  $\approx 3 \text{ \AA}$ ) ce qui est une preuve que le cristal de la protéine diffracte (Figure 35). Néanmoins, les données n'ont pas été collectées car les taches diffuses observées sur l'image de diffraction peuvent être le résultat d'un mauvais empilement des protéines formant le cristal ou d'une mauvaise congélation.





**Figure 35 :** Image de diffraction de DtMan *wt* cristallisée dans la solution 26 % PEG 400, NaHepes 100 mM pH 7,5 et MgCl<sub>2</sub> 200 mM.

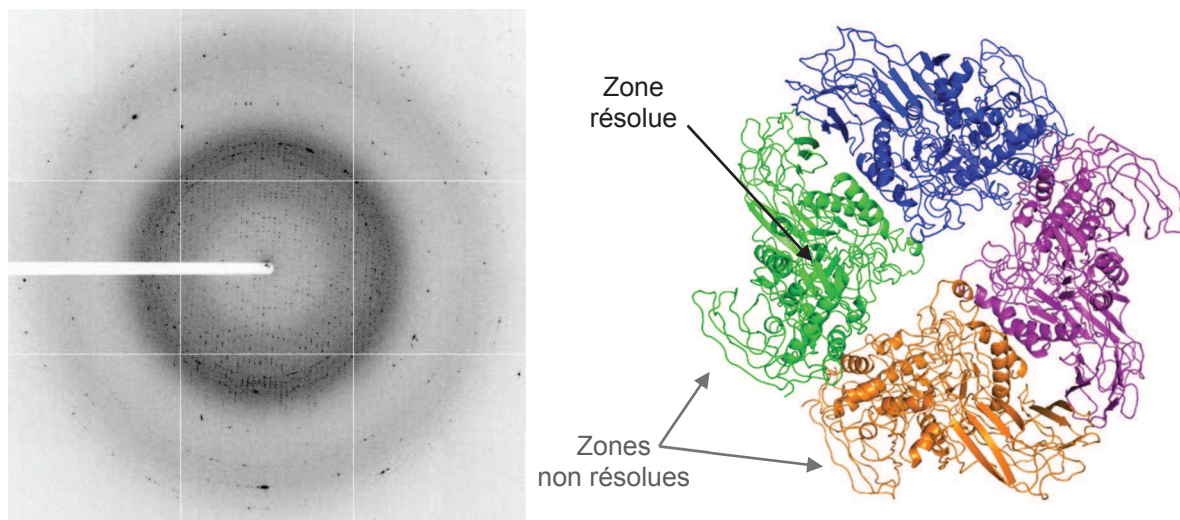
Pour résoudre ce problème deux nouvelles boîtes de cristallisation ont par la suite été réalisées à partir de trois concentrations de DtMan (10, 15 et 20 mg/ml) en suivant les six conditions décrites dans le tableau 6 :

Solution de cristallisation		
Précipitant	Tampon	Additif
24 % PEG 400	NaHepes 100 mM pH 7,5	100 / 200 mM MgCl <sub>2</sub>
25 % PEG 400	NaHepes 100 mM pH 7,5	100 / 200 mM MgCl <sub>2</sub>
26 % PEG 400	NaHepes 100 mM pH 7,5	100 / 200 mM MgCl <sub>2</sub>

**Tableau 6 :** Variation des conditions de cristallisation autour de la solution 26 % PEG 400, NaHepes 100 mM pH 7,5 et MgCl<sub>2</sub> 200 mM.

Les plaques ont été préparées en suivant la méthode d'ensemencement décrite précédemment et laissées au repos à température ambiante. Après analyse par microscopie un cristal en forme de « brique », obtenu dans les conditions 24 % PEG 400, NaHepes 100 mM pH 7,5 et MgCl<sub>2</sub> 200 mM a été plongé dans le système de cryoconservation PEG 400 et dans l'huile. Après congélation dans l'azote liquide le cristal a été analysé sur la ligne Proxima 2 du Synchrotron Soleil. Une image de diffraction à 2,7 Å a été obtenue (**A**, Figure 36) et les

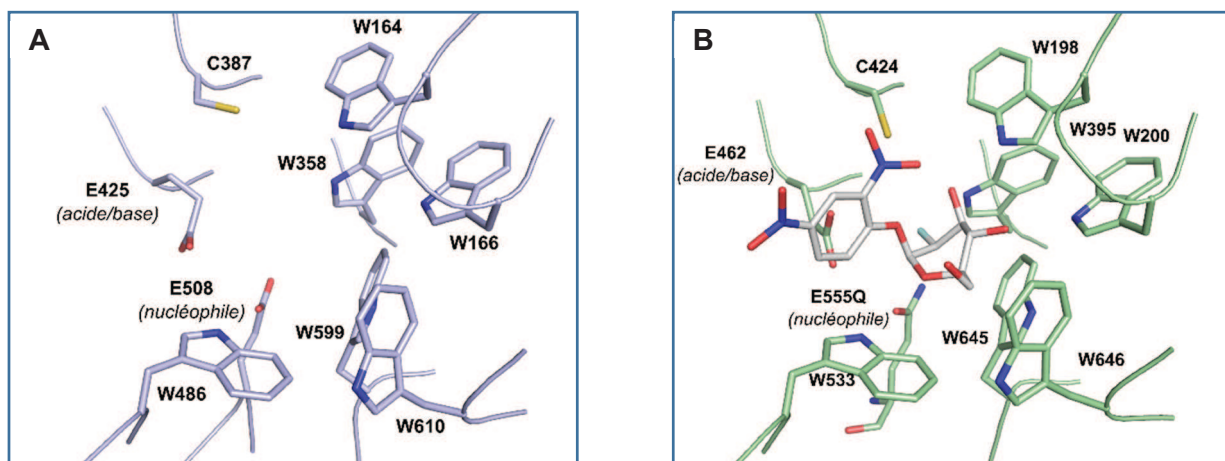
données collectées. Le traitement des données n'étant pas complet au moment de l'écriture de ce manuscrit la description détaillée du modèle de DtMan *wt* n'a pas pu être présentée. Néanmoins les premières constatations montrent que la protéine s'organise sous la forme d'un tétramère (**B**, Figure 36).



**Figure 36 :** (A) Image de diffraction de DtMan *wt* avec une résolution de 2,7 Å et (B) Tetramère de DtMan *wt*.

Si certaines zones restent non résolues, la structure du site actif de l'enzyme a pu être bien définie (**A**, Figure 37). Sa comparaison avec la structure tridimensionnelle de l'enzyme la plus proche, BtMan2A co-cristallisée avec le 2,4-DNP-2fluoro-mannose (**B**, Figure 37), a permis d'identifier les acides aminés responsables de l'activité catalytique de l'enzyme, le Glutamate E425 jouant le rôle d'acide/base et le Glutamate E508 celui du nucléophile. Le rôle des six résidus Tryptophane voisins (W164, W166, W358, W486, W589 et W610) a pu être attribué à la stabilisation du pNPMAN chez DtMan *wt* comme en atteste la fixation du substrat par ces mêmes résidus dans le cas de BtMan2A.





**Figure 37 :** (A) Structure du site actif de DtMan wt et (B) Structure du site actif de BtMan2A co-cristallisée avec le 2,4-DNP-2-fluoromannose (pdb 2WBK (Offen, 2009)).

Le gène de *D. thermophilum*, *dichth\_1692*, codant pour une exo- $\beta$ -mannosidase a été cloné et exprimé chez *E. coli* Rosetta-(DE3). Le caractère thermorésistant de la protéine a été confirmé puisque son activité optimale est atteinte à pH 7 et 90 °C. L'enzyme présente une spécificité stricte vis-à-vis du pNP- $\beta$ -D-Man et une forte efficacité catalytique supérieure aux autres  $\beta$ Man déjà décrites.

## 2. DtRha, $\alpha$ -L-rhamnosidase de *D.thermophilum*

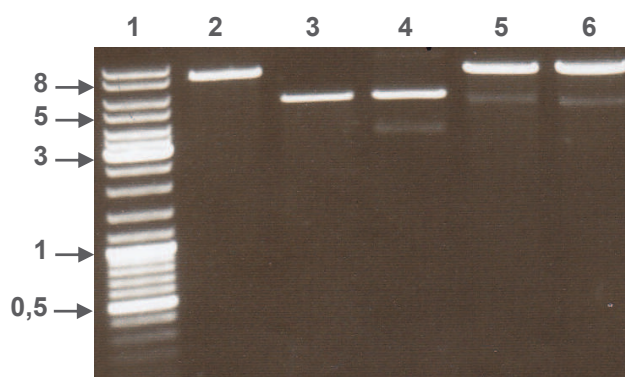
Les  $\alpha$ -L-rhamnosidases (EC 3.2.1.40) sont des enzymes ubiquitaires dans le monde du vivant et sont responsables de l'hydrolyse du motif  $\alpha$ -L-rhamnose de nombreux dérivés glycosylés. En effet le rhamnose est largement distribué chez les plantes et les bactéries sous forme de polysaccharides (pectines, antigène-O de bactéries pathogènes), dérivés glycosylés (flavonoïdes, terpènes) ou de biosurfactants (rhamnolipides).(Cui, 2007) Depuis de nombreuses années les rhamnosidases sont utilisées comme biocatalyseurs dans de nombreux procédés industriels. Dans l'industrie alimentaire, elles sont utilisées pour la production de jus de fruits ou pour l'amélioration de l'arôme du vin par déglycosylation de terpènes. Le  $\alpha$ -L-rhamnose ainsi produit peut être directement valorisé dans l'industrie chimique comme précurseur chiral en synthèse ou pour la glycodiversification de produits naturels.(Thibodeaux, 2008) Les  $\alpha$ -rhamnosidases sont également largement exploitées dans l'industrie pharmaceutique pour la dérhamnosylation de flavonoïdes (rutine, hesperidine, naringine). En effet ces composés sont reconnus pour leurs effets bénéfiques sur la santé humaine, or il a été démontré que les composés mono-glycosylés présentaient une meilleure biodisponibilité que leur analogues disaccharidiques.(Rabausch, 2014; Weignerova, 2012) Le motif  $\alpha$ -L-rhamnose est également largement rencontré chez les glycoprotéines de surface de bactéries et participe donc à la virulence de l'agent pathogène. A l'inverse, il est rarement décrit pour former une liaison directe avec un acide aminé (contrairement au glucose, mannose, galactose,...), à l'heure actuelle seuls quelques exemples font référence à ce type de glycosylation rare. Ainsi, chez *Streptococcus sanguis* une unité rhamnose est couplée au

résidu Asn, tandis que chez *Pseudomonae flagellins* et *P. aeruginosa* la rhamnosylation a lieu sur les résidus Thr et Ser respectivement. Dans cette optique, il nous a semblé intéressant de produire un biocatalyseur capable de reproduire *in vitro* la rhamnosylation d'acides aminés dans le but de développer nos connaissances sur ce type de glycosylation rare.

D'après la classification CAZy, les  $\alpha$ -L-rhamnosidases ( $\alpha$ Rha) agissent selon un mécanisme d'inversion de configuration et sont répertoriées dans 3 familles de GHs (GH78, GH28 et GH106, <http://www.cazy.org/>, 07/2015). A l'heure actuelle une vingtaine de  $\alpha$ Rha de GH78 ont été décrites provenant de divers organismes et parmi ces dernières seules 4 représentantes sont issues de microorganismes thermophiles, *Thermomicrobium* sp (Rhm A et Rhm B),(Birgisson, **2004**) *Clostridium stercorarium* (RamA) (Zverlov, **2000**) et *Aspergillus terreus*.(Weignerova, **2012**) *D.thermophilum* possède sur son génome un seul gène, *dicth\_0289*, codant pour une potentielle exo- $\alpha$ -L-rhamnosidase à inversion de configuration.

Cette enzyme identifiée par le B5YC64 (identification Uniprot) possède une identité de séquence de 51,5 % avec l'  $\alpha$ Rha de *Mahella australiensis* (F3ZY56, GH78) et de 21 % avec l'  $\alpha$ Rha de GH78 cristallisée (3CIH, identification PDB) de *B. tethaiotaomicron* (projet Protein Structure Initiative).

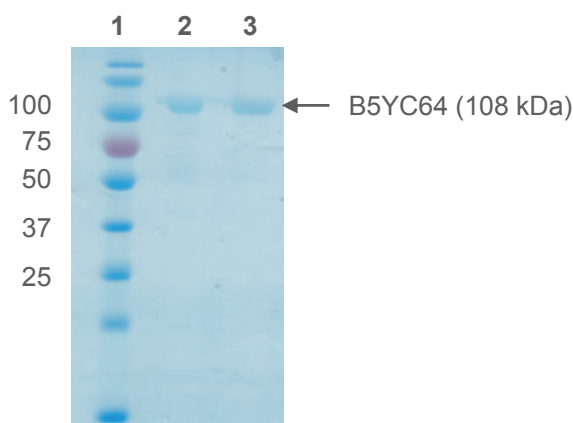
Le gène *dicth\_0289* contenant 2763 paires de bases a été amplifié à partir des amorces construites pour permettre son insertion entre les sites de restriction *Nhe* I/*Sal* I du vecteur pET28-a(+) (cf. Partie expérimentale p.195). Après transformation chez *E. coli* DH5 $\alpha$ , l'ADN plasmidique recombinant de quatre colonies isolées a été extrait, digéré et analysé sur gel d'agarose. Parmi les quatre clones sélectionnés, trois ont présenté la bande de taille attendue ( $\approx$  8100 pb) (Figure 38). Ces clones « positifs » ont été séquencés et l'alignement des séquences avec le gène *dicth\_0289* a confirmé la stricte conformité de l'ADN fils – parent (Annexe 03, p.254). Le plasmide recombinant d'un clone positif a été transformé chez *E. coli* Rosetta-(DE3) et la protéine a été exprimée après induction à l'IPTG de la culture bactérienne dans le milieu nutritif LB/Kan/Cam à 37 °C pendant 20h.



**Figure 38 :** Analyse sur gel d'agarose (1 %) du clonage de *dicth\_0289*.

(1) marqueur de taille (en kpb) ; (2), (5), (6) clones positifs, bandes correspondant à la construction pET28-a(+)/*dicth\_0289* ( $\approx$  8100 pb); (3) plasmide pET28-a(+) vide ( $\approx$  5300 pb) ; (4) clone négatif présentant une bande de la taille du plasmide vide.

Après, lyse, clarification thermique et purification sur colonne de Nickel, l'enzyme ainsi purifiée est analysée sur gel SDS-PAGE. La protéine est constituée d'un enchainement de 921 acides aminés ce qui représente un poids moléculaire théorique de 106 kDa, avec la présence des 6 motifs His (provenant du His●Tag) le poids moléculaire réel de B5YC64 est de 108 kDa. L'analyse sur gel confirme la présence de la protéine puisqu'une bande unique est observée au niveau que de la bande de la protéine de référence de 100 kDa (piste 3, Figure 39). Le rendement de production de la protéine est de 4,5 mg par litre de culture bactérienne après purification.

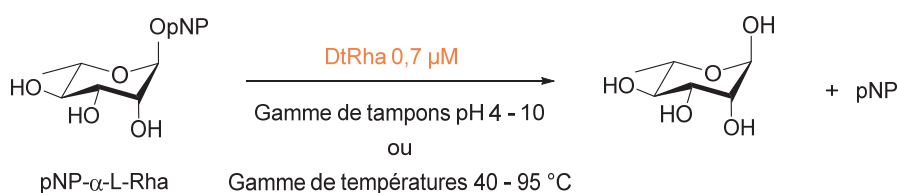


**Figure 39** : Analyse sur gel d'électrophorèse SDS-PAGE (10 % acrylamide) de la protéine B5YC64. (1) Marqueur de tailles de protéines (en kDa) ; (2) & (3) B5YC64 pure après purification sur colonne de  $\text{Ni}^{2+}$ .

L'activité de  $\alpha$ -L-rhamnosidase de B5YC64 de *D. thermophilum* a été confirmée en testant l'enzyme sur notre panel de pNPsucre (Figure 28, p.73). Après 30 minutes d'incubation dans un milieu tamponné Imidazole.HCl 20 mM pH 7 à 37 °C, seul le milieu contenant 1mM de pNP- $\alpha$ -L-rhamnopyranoside (pNPRha) a présenté une couleur jaune caractéristique de la libération de pNP et donc de l'hydrolyse par l'enzyme, ainsi nommée **DtRha**.

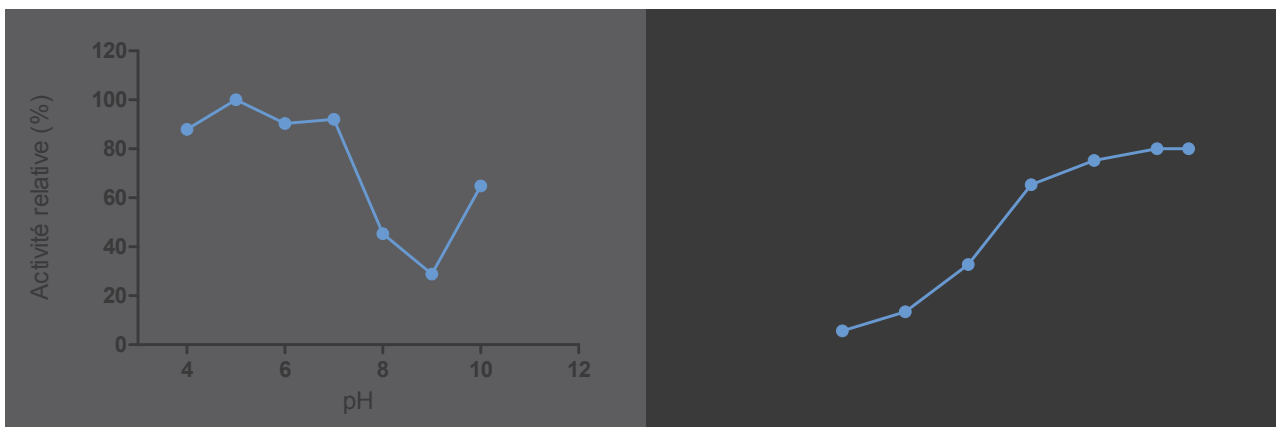
### 2.1. Conditions optimales de réaction et paramètres cinétiques

L'activité de DtRha a dans un premier temps été mesurée suivant différentes conditions de pH (pH 4 à 10) puis à partir de ce premier résultat l'hydrolyse du pNPRha par l'enzyme a été suivie sur une gamme de températures (40 à 95 °C) au pH optimal (pH<sub>Opt</sub>) de l'enzyme. (Schéma 28). Ces deux études ont permis de tracer les courbes d'activité correspondantes à partir des maxima d'hydrolyse relevés à pH 5 et 95°C (Figure 40).



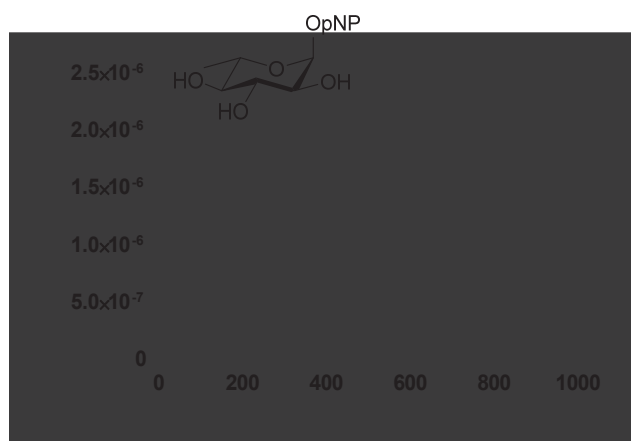
**Schéma 28** : Détermination des conditions optimales d'activité d'hydrolyse de DtRha sur une gamme de pH et de températures.

Ainsi le comportement de l'enzyme est favorisé sur une large gamme de pH à tendance acide, avec un pH<sub>Opt</sub> de 5 (**A**), ce résultat est comparable aux  $\alpha$ Rha caractérisées (pH<sub>Opt</sub> 5 à 8). (Birgisson, **2004**; Weignerova, **2012**) L'enzyme est également très stable vis-à-vis de la température puisqu'elle présente une forte activité pour des températures supérieures à 70 °C avec un pic d'activité à 95 °C (**B**), à l'heure actuelle aucune représentante thermostable de  $\alpha$ Rha présentant une telle activité à haute température n'a été décrite.



**Figure 40 :** Effets du pH (**A**) et de la température (**B**) sur l'activité de DtRha.

Les paramètres cinétiques de l'enzyme ont été déterminés en mesurant l'activité d'hydrolyse de l'enzyme (à une concentration finale de 0,2  $\mu$ M) sur une gamme de concentration de pNPRha (5 – 800  $\mu$ M) dans un tampon acide citrique/ $\text{Na}_2\text{HPO}_4$  20 mM pH 5 à 37 °C (Figure 41), le choix de la température a également été fixé pour assurer un compromis activité / stabilité de l'enzyme.



**Figure 41 :** Courbe de Michaelis-Menten d'activité d'hydrolyse de DtRha en présence d'une gamme de concentration en pNPRha à pH 5 et 37 °C. Mesures d'activité réalisées en triplicat.

Dans ces conditions, la constante de Michaelis-Menten de DtRha calculée est de  $54 \pm 3 \mu\text{M}$ , cette valeur est comparable aux  $K_M$  décrits pour les  $\alpha$ Rha mésophiles et reflète de la forte

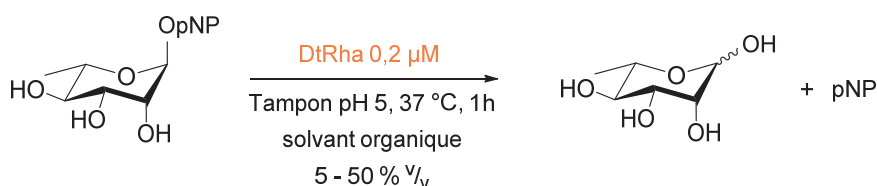
affinité de DtRha pour le pNPRha (Tableau 7). En revanche à 37 °C l'enzyme présente une efficacité catalytique ( $k_{cat}/K_M = 3,14 \text{ mM}^{-1}.\text{s}^{-1}$ ) relativement faible par comparaison avec ses analogues thermophiles (facteur 1000 avec RhmA de *T. sp*). Contrairement aux résultats observés avec DtMan ( $k_{cat} = 10 \text{ s}^{-1}$  à 37 °C), il semble que la rhamnosidase de *D. thermophilum* nécessite des conditions de hautes températures pour assurer l'hydrolyse du pNPRha.

Enzyme	$K_M$ (mM)	$k_{cat}$ ( $\text{s}^{-1}$ )	$k_{cat}/K_M$ ( $\text{mM}^{-1}.\text{s}^{-1}$ )
<i>RhaB</i> <i>B. sp</i> (Cui, <b>2007</b> )	0,28	140	500
<i>KoRha</i> <i>K. oxytoca</i> (O'Neill, <b>2015</b> )	0,21	0,98	4,67
<i>SaRha78A</i> <i>S. avermitilis</i> (Fujimoto, <b>2013</b> )	0,026	26,4	1 014
<b>DtRha</b> <i>D. thermophilum</i>	<b>0,054 ± 3</b>	<b>0,17 ± 0,01</b>	<b>3,14</b>
<i>RhmA</i> <i>T. sp</i> (Birgisson, <b>2004</b> )	0,46	460	1 000

**Tableau 7 :** Comparaison des constantes cinétiques de DtRha avec des  $\alpha$ Rha caractérisées.

## 2.2. Influence du co-solvant

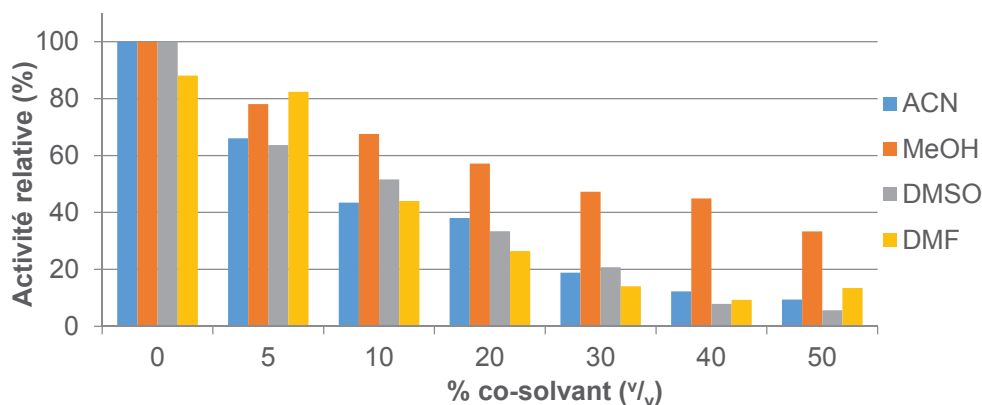
La stabilité de DtRha a été évaluée en présence des solvants décrits précédemment en suivant l'hydrolyse du pNPRha (1mM) dans un milieu tamponné acide citrique/ $\text{Na}_2\text{HPO}_4$  20 mM pH 5 à 37 °C pendant 1h (Schéma 29). L'activité mesurée dans chaque cas a été comparée à celle d'un échantillon blanc sans co-solvant (100 % d'activité) (Figure 42).



**Schéma 29 :** Suivi de l'activité d'hydrolyse de DtRha en présence de co-solvant.

Le graphique représentant l'activité relative de DtRha en fonction de différentes concentrations ( $v/v$ ) en solvants permet de mettre en évidence la stabilité de l'enzyme en conditions organiques. En effet l'activité d'hydrolyse de DtRha est conservée à 5 % de son maximum quand le milieu réactionnel est constitué d'un ratio 50:50 tampon / solvant et ce quel que soit le type de solvant. Enfin l'enzyme est fortement tolérante vis-à-vis de l'ajout de méthanol dans le milieu réactionnel puisque 30 % d'activité d'hydrolyse sont conservés pour un mélange 50:50 tampon / MeOH. Ce système pourra donc être privilégié pour les essais de transglycosylation futurs, néanmoins grâce à la stabilité générale de l'enzyme l'acétonitrile, le

DMSO ou encore le DMF pourront être utilisés dans le cas d'accepteurs peu solubles dans le méthanol.



**Figure 42** : Stabilité de DtRha vis-à-vis de différents co-solvants.

### 2.3. Cristallogénèse de DtRha

Des études de cristallographie ont été réalisées sur DtRha pour tenter de résoudre sa structure tridimensionnelle et obtenir plus d'information sur la structure de son site actif. En suivant le protocole décrit pour la cristallogénèse de DtMan (p. 78), 2ml d'une fraction pure de DtRha a été récupérée à 0,89 mg/ml. Après concentration (200  $\mu$ l à 10 mg/ml) les conditions de cristallisation de l'enzyme ont été testées à partir du kit commercial JBScreen Basic HTS (JenaBioscience) contenant 96 compositions de solutions pour deux concentrations 5 et 10 mg/ml. Les plaques de cristallisation ont été préparées en associant chaque solution de cristallisation (90  $\mu$ l dans le réservoir) à une goutte (2  $\mu$ l par puit) composée du ratio 1:1 protéine / tampon de cristallisation. Après une période de repos de 15 jours à 20 °C, l'analyse par microscopie des plaques a permis de mettre en évidence 7 hits de cristallisation décrits dans le tableau 8 suivant :

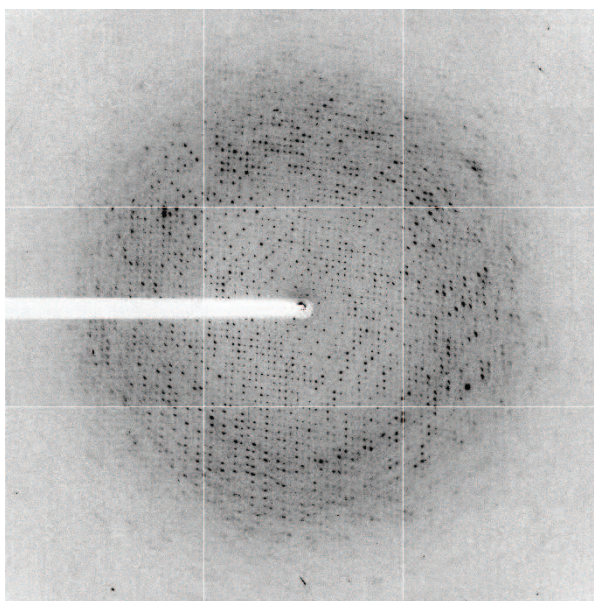
Solution de cristallisation			Forme des cristaux
Précipitant(s)	Tampon	Additif	
2 % PEG 400 2 M sulfate d'ammonium	NaHepes 100 mM pH 7,5	-	Petites briques
10 % PEG 1000 10 % PEG 8000	-	-	Aiguilles
30 % PEG 1500	-	-	Gros cristal sous forme de feuillet
8 % PEG 4000	Acétate de sodium 100 mM pH 4,6	-	Aiguilles
10 % PEG 6000	NaHepes 100 mM pH 7,5	5 % $v/v$ MPD*	Aiguilles
10 % PEG 8000	NaHepes 100 mM pH 7,5	8 % Ethylène glycol	Aiguilles
15 % PEG 8000	-	Sulfate de Lithium 500 mM	Aiguilles

\*MPD = 2-méthyl-2,4-pentanediol

**Tableau 8 :** Composition des solutions favorisant la cristallisation de DtRha.

Le cristal obtenu dans les conditions 30 % PEG 1500 a été collecté et congelé dans une solution de cryoprotectant PEG 1500, MPD (25 %  $v/v$ ) et son analyse lors d'un premier run sur la ligne Proxima 1 du Synchrotron Soleil (<http://www.synchrotron-soleil.fr/>) a permis d'obtenir une image de diffraction de résolution 3,6 Å (Figure 43). Les données obtenues n'ont cependant pas été collectées car la diffraction d'un cristal doit au moins être  $\leq 2,5$  Å pour permettre une résolution précise de la structure tridimensionnelle de la protéine. Néanmoins cette première expérience a permis de confirmer que le cristal de DtRha diffracte. A partir de ce résultat prometteur de nouvelles conditions ont été testées en modulant les conditions de cristallisation de manière à optimiser la formation des cristaux de DtRha.





**Figure 43 :** Image de diffraction de DtRha cristallisée dans les conditions PEG 1500 (30 %) et congelée dans un système PEG 1500 et MPD (25 % v/v).

En suivant le processus décrit précédemment une plaque de cristallisation a été préparée à partir de trois concentrations en protéine (5, 10 et 15 mg/ml) et 12 solutions de cristallisation (Tableau 9).

Solution de cristallisation		
Précipitant(s)	Tampon	Additif
24 à 34 % PEG 1500	-	-
30 % PEG 1500	NaHepes 100 mM pH 7,5	-
30 % PEG 1500	NaHepes 100 mM pH 7,5	100 mM MgCl <sub>2</sub>
30 % PEG 1500	NaHepes 100 mM pH 7,5	200 mM MgCl <sub>2</sub>
30 % PEG 1500	Acétate de sodium 100 mM pH 4,6	-
30 % PEG 1500	Acétate de sodium 100 mM pH 4,6	200 mM MgCl <sub>2</sub>
30 % PEG 1500	Acétate de sodium 100 mM pH 4,6	100 mM MgCl <sub>2</sub>

**Tableau 9 :** Solutions de cristallisation optimisées de DtRha.

Une seconde boîte a par la suite été préparée à partir d'une fraction de DtRha à 20 mg/ml et de quatre conditions de cristallisation (34 % PEG seul ; 36% PEG seul ; 30 % PEG 1500 et



NaHepes 100 mM pH 7,5 ; 30 % PEG 400, NaHepes 100 mM pH 7,5 et  $\text{MgCl}_2$  100 mM) selon la méthode d'ensemencement. Un cristal, en forme de demi-lune, obtenu dans les conditions 30 % PEG 1500 et NaHepes 100 mM pH 7,5 a été plongé dans le système de cryoconservation MDP 25 % et dans l'huile, congelée dans l'azote liquide et analysé sur la ligne Proxima 2 du Synchrotron Soleil. Une image de diffraction à 3 Å a été obtenue mais les taches diffuses et dédoublées laissent supposer de la présence de plusieurs cristaux superposés. Les données n'ont pas pu être collectées. Néanmoins l'augmentation de la résolution de diffraction prouve que l'optimisation des conditions de cristallisation de DtRha est en bonne voie. A l'heure actuelle de nouvelles conditions sont en cours de test.

### 3. DtGly, $\beta$ -D-glucosidase de *D. thermophilum*

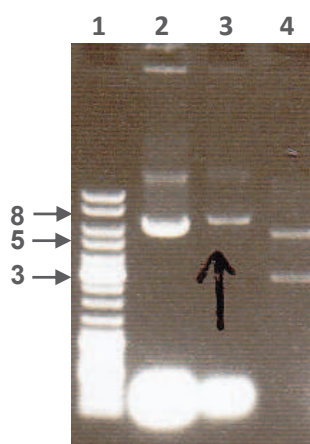
Les  $\beta$ -D-glucosidases (EC 3.2.1.21) catalysent l'hydrolyse de liaisons glycosidiques à l'extrémité non réductrice d'oligosaccharides ou de dérivés glycosylés. Dans la nature, ces enzymes sont présentes de manière ubiquitaire dans les différents domaines du monde du vivant où elles jouent un rôle important dans de nombreux processus biologiques. Chez les mammifères, les  $\beta$ -glucosidases interviennent principalement dans le métabolisme des glycolipides et des dérivés glucosylés (issus de l'alimentation). Chez l'Homme, La  $\beta$ -glucosidase acide, ou glucocérébrosidase (GBA1) est responsable de la maladie de Gaucher. Cette maladie de surcharge lysosomale est la conséquence d'un dysfonctionnement dans le transport ou l'activité de GBA1 dans les lysosomes provoquant une accumulation glycocéramides dans les tissus du système nerveux ou de la moelle osseuse.(Butters, **2007**) Ce sont chez les plantes que les  $\beta$ -glucosidases présentent la plus grande variété de fonctions biologiques. En effet elles sont, entre autre, impliquées dans les processus de défense, de métabolisme des parois cellulaires, de signalisation ou encore dans les processus de métabolisme secondaire (dégradation d'alkaloïde glycosylés par exemple). Les  $\beta$ -glucosidases de souches microbiennes suscitent quant à elle beaucoup plus d'intérêt pour leurs applications potentielles plutôt que dans l'exploration de leurs rôles biologiques. Parmi ces dernières, les enzymes responsables de la dégradation des polysaccharides complexes constituant les parois cellulaires de plantes sont les plus étudiées pour proposer des alternatives de choix dans la production de bioéthanol, notamment, à partir de la dégradation directe de produits issus de la biomasse.(Cairns, **2010**)

Dans la littérature les  $\beta$ -D-glucosidases ( $\beta$ -glu) décrites ont été affiliées à 6 familles de glycoside hydrolases : GH1, GH3, GH5, GH9, GH30 et GH116.(Cantarel, **2009**; Cobucci-Ponzano, **2010**; Henrissat, **1991**) Les représentantes des familles GH1, GH3, GH5 et GH30 sont regroupées dans le clan-A des GH qui englobe les enzymes suivant un mécanisme avec rétention de configuration et possédant une structure tridimensionnelle de type  $(\beta/\alpha)_8$ . La famille GH1 représente l'une des familles les plus étudiées de par notamment la diversité de spécificité de substrat de ses glycosidases, à l'heure actuelle plus de 300 enzymes issues de cette famille ont été caractérisées dont 47 sont de structures connues. Hormis leur activité d'hydrolyse, ces enzymes sont largement étudiées pour la synthèse de glucosides d'alkyle et de dérivés glycosylés pour des applications industrielles. Si les  $\beta$ -glucosidases mésophiles ont prouvé leur efficacité en transglycosylation, telle que la  $\beta$ -glucosidase d'amande intensivement

étudiée et disponible commercialement, les recherches se sont progressivement portées vers les biocatalyseurs issus de micro-organismes thermophiles. De nombreux exemples de  $\beta$ -glu thermostables de GH1 utilisées en transglycosylation sont disponibles dans la littérature (BglA de *T. maritima*, (Goyal, **2001**) Ttb-gly de *T. thermophilus*, (Dion, **1999**) CelB de *P. furiosus* (Hansson, **2001a**)...) et rendent compte des avantages multiples de ces enzymes, plus grande stabilité et résistance vis-à-vis des solvants organiques en autre.

Sur le génome de *D. thermophilum* quatre gènes, *dicth\_0359* / *dicth\_1689* et *dicth\_0262* / *dicth\_1949*, codent pour des  $\beta$ -glu potentielles à rétention de configuration appartenant aux familles GH1 et GH3 respectivement. Nous avons porté notre sélection sur une représentante de GH1, en utilisant l'ensemble des données actuellement disponibles (structures et enzymes caractérisées) sur cette famille comme outil de comparaison lors de sa caractérisation. Le clonage de *dicth\_1689* et la caractérisation de la  $\beta$ -glu correspondante DtGH (Zou, **2012**) venant d'être réalisés au démarrage de notre projet nous avons porté notre étude sur le clonage du gène *dicth\_0359*. La potentielle exo- $\beta$ -glucosidase correspondante, B5YCI2, possède 48% d'identité de séquence avec une  $\beta$ -glu non caractérisée de *Caldilinea aerophila* et 44% d'identité de séquence avec la  $\beta$ -glu cristallisée de *P. horikoshii* (1VFF, identification PDB). (Matsui, **2000**)

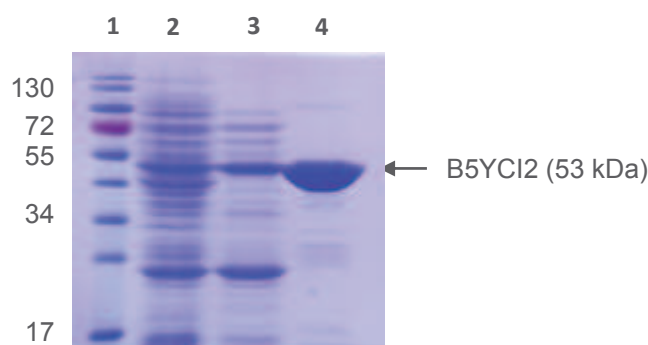
Le gène constitué d'un enchaînement de 1257 paires de bases a été amplifié par PCR à partir des amorces construites pour présenter les sites de restrictions des enzymes *NheI* et *NotI* aux extrémités de l'insert (cf. Partie expérimentale p.195). Après digestion du vecteur pET28-a(+) par les enzymes de restrictions *NheI/NotI* la ligation de l'insert est effectuée et le plasmide recombinant est transformé chez *E. coli* DH5 $\alpha$ . A l'issue de cette étape deux colonies bactériennes ont été isolées parmi la vingtaine de clones obtenus, chaque ADN plasmidique ainsi produit a pu être extrait, et analysé sur gel d'agarose (Figure 44). Un seul clone a présenté la bande de taille désirée ( $\approx$  6500 pb) qui a pu être séquencé et l'alignement des séquences en nucléotides ADN parent – ADN recombinant a permis de confirmer le clonage du gène *dicth\_0359* (Annexe 04, p.258).



**Figure 44 :** Analyse par électrophorèse du clonage de *dicth\_0359*.

(1) Marqueur de taille (en kpb) ; (2) clone négatif ; (3) clone positif bande unique 6500 pb ;  
(4) pET28-a(+) vide (5300 pb).

Enfin, le plasmide a été transformé chez *E. coli* Rosetta-(DE3) et l'expression de la protéine a été initiée par l'ajout d'IPTG dans le milieu de culture bactérien composé de LB/Kan/Cam à 37 °C. Au terme des 20h d'incubation à 37 °C, la protéine est récupérée après lyse et clarification thermique puis purifiée par chromatographie d'affinité (Ni-NTA). La protéine B5YCI2 est constituée d'une séquence de 418 acides aminés ce qui représente une masse théorique de 49 kDa et une masse réelle de 53 kDa avec l'incorporation du His•Tag. L'analyse sur gel SDS-PAGE révèle la présence d'une bande inférieure à 55 kDa à chaque étape qui a suivi la culture bactérienne ce qui correspond à la protéine cible (Figure 45). Le rendement de production de la protéine est de 2 mg par litre de culture bactérienne.



**Figure 45 :** Analyse sur gel SDS-PAGE de la production de B5YCI2.

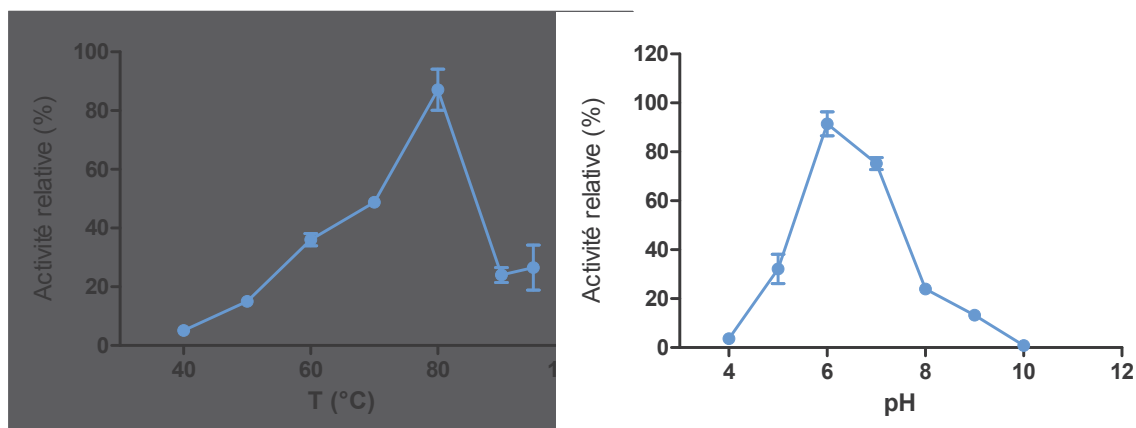
(1) Marqueur de taille protéique (en kDa) ; (2) surnageant de lyse ; (3) surnageant de clarification ; (4) fraction protéique collectée après purification sur colonne de Nickel.

L'activité d'hydrolyse de B5YCI2 a été testée sur le panel de pNPsucre disponibles au laboratoire dans un tampon imidazole.HCl 20 mM pH 7 à 37 °C (Schéma 30). Après 30 minutes d'incubation, la réaction enzymatique est stoppée par ajout de Na<sub>2</sub>CO<sub>3</sub> 1 M et la quantité de pNP libre en solution est quantifiée par mesure d'absorbance à  $\lambda = 405$  nm. Ce test préliminaire a permis de mettre en évidence le caractère versatile de l'enzyme ainsi que sa grande spécificité vis-à-vis des liaisons  $\beta$ -glycosidiques. En effet B5YCI2 catalyse l'hydrolyse du pNP- $\beta$ -D-Glc, pNP- $\beta$ -D-Gal et pNP- $\beta$ -D-Fuc mais également, à des taux plus faibles, celle du pNP- $\beta$ -D-Man et du pNP- $\beta$ -D-Xyl. De manière surprenante, l'enzyme catalyse plus efficacement l'hydrolyse du pNPFuc, l'activité relative de l'enzyme par rapport à ce maximum a pu être déterminée pour l'ensemble des pNPsucre testés (Figure 46). Cette caractéristique est également décrite chez l'analogue DTGH et est redondante chez de nombreuses représentantes de GH1. (Cairns, **2010**; Zou, **2012**) L'activité de  $\beta$ -glucosidase de B5YCI2 a été confirmée, cependant, compte-tenu de la diversité de substrats reconnus par l'enzyme nous avons choisi de l'identifier comme la **DtGly** de *D. thermophilum*.

Substrat	Activité relative (%)
PNP-α-D-Gal	0
PNP-β-D-Gal	98
PNP-α-D-Glc	0
PNP-β-D-Glc	85
PNP-α-D-Man	10
PNP-β-D-Man	0
PNP-α-L-Rha	0
PNP-β-D-GlcNac	0
PNP-α-D-GlcNac	0
PNP-α-D-Xyl	0
PNP-β-D-Xyl	18
PNP-α-L-Fuc	0
PNP-β-L-Fuc	0
PNP-β-D-Fuc	100
PNP-α-L-Galf	0

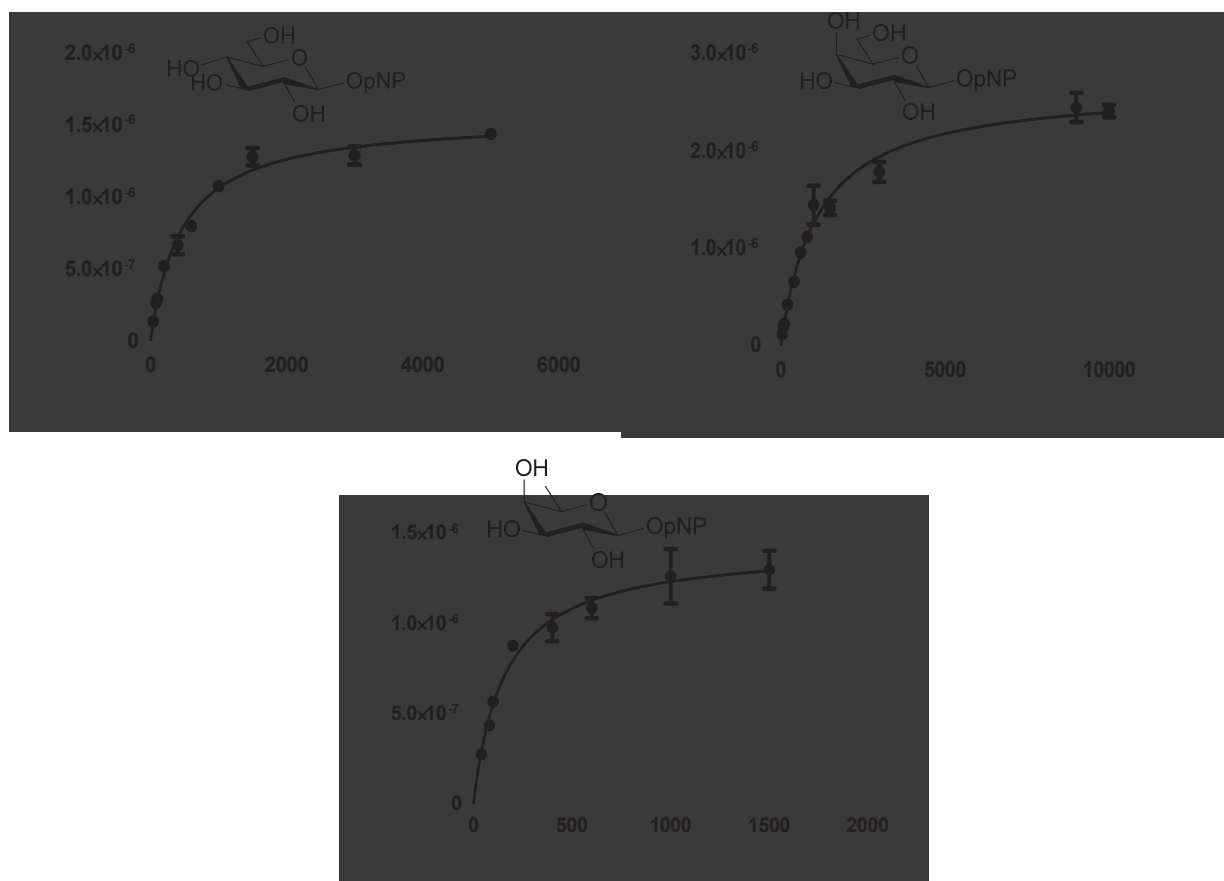
### 3.1. Conditions optimales de réaction et paramètres cinétiques

105



**Figure 47** : Influence de la température **(A)** et du pH **(B)** sur B5YCI2.

A partir de ces paramètres les constantes cinétiques de DtGly en activité d'hydrolyse ont été calculées pour chacun des trois meilleurs substrats (pNPGlc, pNPGal et pNPFuc) sur une gamme de concentration de 0,05 à 10 mM dans un tampon acide citrique/ $\text{Na}_2\text{HPO}_4$  20 mM pH 6 à 37 °C (Figure 48). L'enzyme a été introduite dans une concentration finale de 4,1 nM et les mesures ont été réalisées en triplicat.



**Figure 48** : Courbes de Michaëlis-Menten de DtGly en activité d'hydrolyse sur le pNPGlc, pNPGal et pNPFuc à pH 6 et 37 °C.

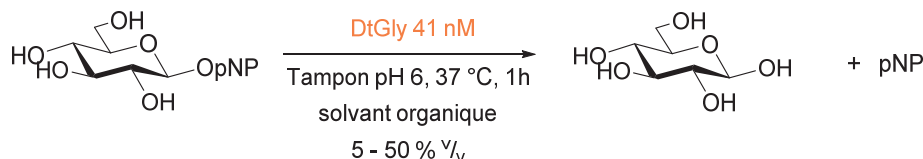
Les résultats obtenus, présentés dans le tableau 10, mettent en évidence la dépendance de l'activité enzymatique par rapport au motif sucre du substrat. Ainsi compte-tenu des valeurs de  $K_M$ , l'affinité de DtGly pour ces trois substrats peut être exprimée selon l'ordre suivant : pNPFuc > pNPGlc > pNPGal. Par ailleurs, le ratio  $k_{cat}/K_M$  est le paramètre le plus significatif pour rendre compte de la spécificité de l'enzyme pour son substrat. Les valeurs obtenues pour les trois pNPsucres sont en accord avec les conclusions basées sur l'analyse du  $K_M$  puisque la spécificité de l'enzyme vis-à-vis du pNPFuc est jusqu'à 2 fois supérieure que pour le pNPGlc et 4 fois supérieure à celle du pNPGal.

Substrat	$K_M$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $mM^{-1}.s^{-1}$ )
pNP- $\beta$ -D-Glc	$0,46 \pm 0,04$	$30,77 \pm 0,7$	66
pNP- $\beta$ -D-Gal	$1,55 \pm 0,13$	$65,04 \pm 1,46$	42
pNP- $\beta$ -D-Fuc	$0,16 \pm 0,025$	$28,29 \pm 1,22$	175

**Tableau 10** : Constantes cinétiques calculées pour les différents substrats de DtGly.

### 3.2. Influence du co-solvant

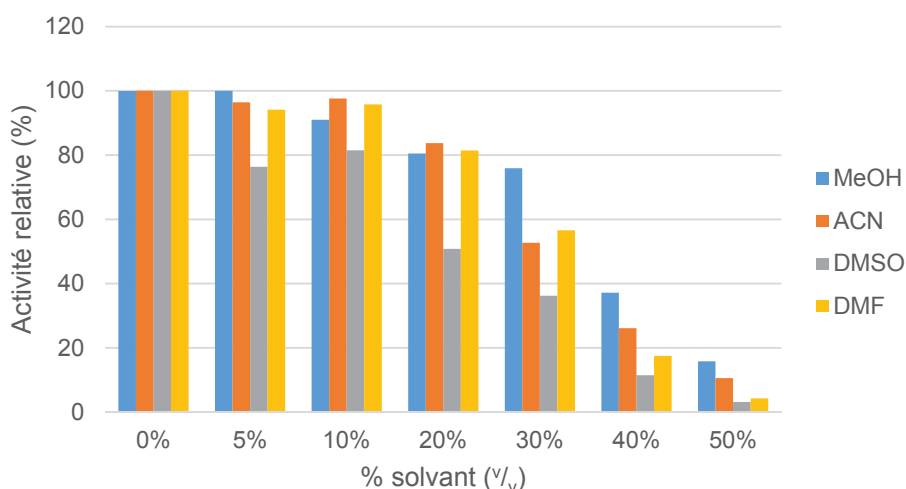
La stabilité de DtGly vis-à-vis de différents co-solvant a été testées après 1h d'incubation dans un milieu tamponné Acide citrique/ $Na_2HPO_4$  20 mM pH 6 contenant 1 mM de pNPGlc en présence de concentrations croissantes en solvant à 37 °C (Schéma 31).



**Schéma 31** : Etude de l'activité d'hydrolyse de DtGly en présence de différents co-solvants à pH 6 et 37 °C.

Le graphique obtenu met en exergue la haute tolérance de DtGly par rapport aux solvants organiques, bien supérieure aux deux glycosidases étudiées précédemment, DtMan et DtRha (Figure 49). Quel que soit le type de solvant, l'activité de l'enzyme est conservée au-delà de 20 % quand le milieu réactionnel est composé jusqu'à 40 % en phase organique. Un autre fait remarquable est à noter sur ce graphique dans le cas du méthanol ou DMF. Quand 5% de méthanol sont ajoutés dans le milieu, l'activité de l'enzyme est réduite de 10%, en revanche quand le méthanol est ajouté en plus grande quantité (jusqu'à 20%) le maximum d'activité de DtGly est de nouveau atteint. La même observation peut être faite avec le DMF. Ainsi pour

limiter l'effet inhibiteur de l'ajout d'un co-solvant sur l'activité enzymatique de DtGly, le méthanol et le DMF seront à privilégier et ce à une concentration n'excédant pas les 20% v/v.



**Figure 49 :** Etude de l'influence de divers co-solvant sur l'activité de DtGly.

### 3.3. Cristallogénèse de DtGly

Compte-tenu des propriétés intéressantes de DtGly et notamment sa versatilité vis-à-vis du substrat des études de cristallogénèse ont été réalisées sur cette protéine afin d'obtenir de plus amples informations sur la structure de son site actif le mode de reconnaissance des différents substrats. En suivant le protocole décrit pour la cristallogénèse de DtMan (p. 78), une fraction pure de DtGly a été récupérée de concentration 0,45 mg/ml (2ml). Après concentration, 250 µl de DtGly à 3,2 mg/ml sont utilisés pour tester 100 solutions de cristallisation à partir du kit commercial Structure Screen 1 (Molecular Dimensions). La plaque de cristallisation a été préparée en associant chaque solution de cristallisation (90 µl dans le réservoir) à une goutte (2 µl par puit) composée du ratio 1:1 protéine / tampon de cristallisation. Après une période de repos de 15 jours à 20 °C, l'analyse de la plaque par microscopie n'a révélé la croissance d'aucun cristal. Néanmoins la cristallogénèse des protéines étant une science empirique, il n'existe pas de règles théoriques ou de base de données permettant de prédire précisément les conditions de cristallisation d'une protéine. Plusieurs paramètres sont à prendre en compte tels que la concentration de la protéine, la température, le pH ou encore les additifs utilisés (sels, solvants organiques, polymères organiques, précipitants). Actuellement les conditions du kit commercial JBScreen Basic HTS sont en cours de test ainsi que l'influence de la température sur la cristallisation de DtGly, les boîtes préparées ont été laissées au repos à 4 °C.

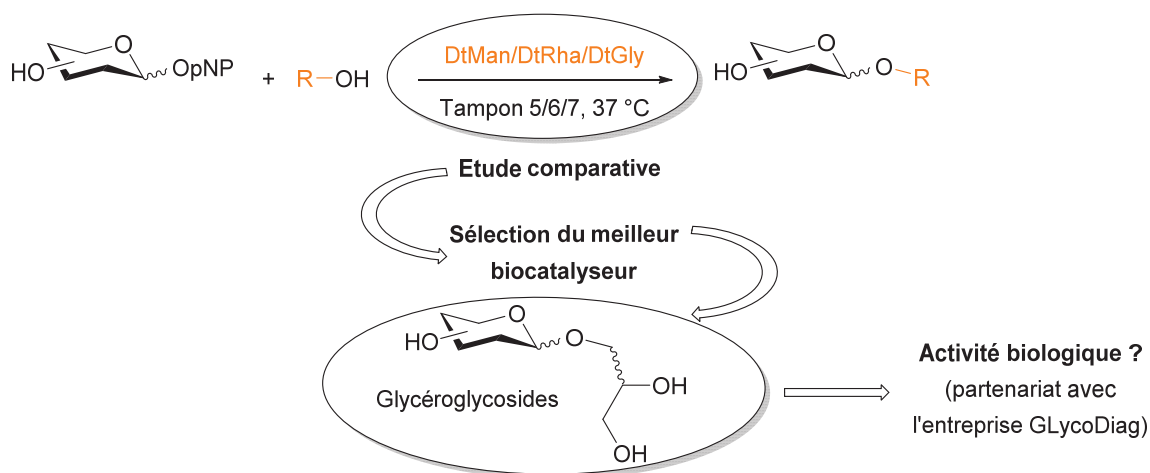
## III. Valorisation de biocatalyseurs pour la synthèse de glycosides d'alkyle, applications en cosmétique

Depuis de nombreuses années les glycosidases sont considérées comme une alternative de choix aux procédés classiques de chimie dans la mesure où elles permettent la synthèse « one-pot » de dérivés glycosylés. Parmi ces derniers, l'accès aux glycosides d'alkyle suscite



un fort intérêt. Ces tensioactifs biodégradables possèdent des propriétés remarquables qui en font de bons agents émulsifiant et d'hydratation recherchés dans les domaines agroalimentaire, pharmaceutique ou cosmétique. Ces composés sont également connus pour leurs propriétés dermatologiques qui les rendent compatibles dans la formulation de produits de soin en santé humaine. (Mladenoska, **2008**) La nécessité de proposer des biocatalyseurs de synthèse toujours plus performants dans la chimie des sucres et l'intérêt biologique de ces O-glycoconjugués nous a encouragé à évaluer le potentiel de nos glycosidases nouvellement caractérisées.

Nous nous sommes donc intéressés dans un premier temps à l'activité de transglycosylase de nos différentes enzymes en utilisant des alcools simples comme accepteurs. Dans un second temps nous avons sélectionné la meilleure candidate afin de synthétiser par voie chimio-enzymatique des glycosides de glycérol dont les propriétés antimicrobiennes ont pu être déterminées grâce à une collaboration réalisée avec l'entreprise GLYcoDiag (Schéma 32).



**Schéma 32 :** Etude du potentiel de synthèse des glycosidases de *D. thermophilum* et application à la synthèse de glycéroglycosides.

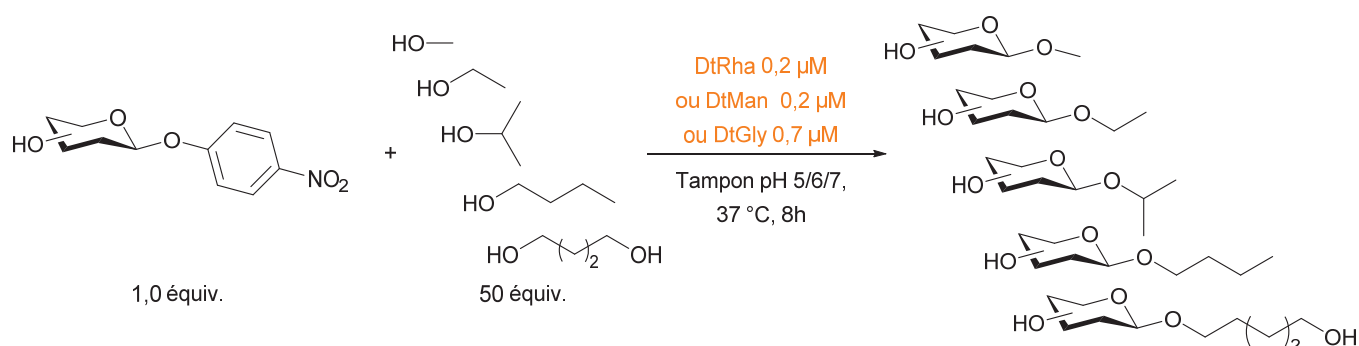
## 1. Criblage d'activité transglycosylase

La transglycosylation est une réaction régie par un contrôle cinétique. Pour favoriser la réaction dans le sens de la formation du produit de transglycosylation les essais ont été réalisés en présence de concentrations élevées en donneur et accepteur de glycosides. Ainsi, les pNPsucre, respectivement substrat des glycosidases produites, ont été utilisés comme donneurs activés de glycoside à une concentration finale dans le milieu réactionnel de 15 mM. Les accepteurs de type alcool primaire ou secondaire ont été ajoutés en large excès par rapport au donneur de glycoside (50 équivalents) afin de favoriser la réaction dans le sens de la transglycosylation. Un degré de diversité a été introduit en faisant varier la longueur de la chaîne alkyle de l'accepteur afin d'évaluer un éventuel effet sur le rendement de transglycosylation. Chaque enzyme a été fraîchement préparée à partir d'une culture bactérienne d'*E. coli* Rosetta-(DE3) transformées par le plasmide recombinant correspondant.



Après lyse thermique et mécanique les surnageant cellulaires ont été clarifiés thermiquement permettant ainsi d'obtenir les trois solutions protéiques de DtRha (0,24 mg/ml, soit une concentration molaire de 2  $\mu$ M), DtMan (0,2 mg/ml, 2  $\mu$ M) et DtGly (0,35 mg/ml, 0,7  $\mu$ M).

L'ensemble des réactions a été conduit à 37 °C pour limiter l'hydrolyse spontanée du donneur de glycoside, et inscrire nos biocatalyseurs dans une démarche éco-compatible, et en milieu tamponné correspondant aux conditions optimales de pH de chaque enzyme (Schéma 33). La formation des produits de transglycosylation a été suivie par chromatographie sur couche mince (CCM), de manière générale l'absence de produit de départ est observée après 8h de réaction. Dans chaque cas, le rendement a été calculé après isolement des produits (Tableau 11).



**Schéma 33** : Synthèse de glycosides d'alkyle.

Les essais de transglycosylation à partir de DtMan (entrée A) ou DtRha (entrée B) n'ont conduit à la détection d'aucun produit de transglycosylation même après 8h de réaction à 37 °C. En revanche dans le cas de DtGly une activité transglycosidase a pu être mise en évidence (entrée C). En présence de méthanol l'enzyme est capable de catalyser la réaction de transglycosylation à partir du pNPGlc et du pNPGal conduisant aux  $\beta$ -1-O-glycoside de méthyle **1** et **6** correspondants avec un rendement proche de 60 %. Les mêmes résultats sont observés dans le cas de l'éthanol où les produits de transglycosylation du Glc et Gal **2** et **7** ont été obtenus avec des rendements équivalents. La réaction de transglycosylation est moins efficace lorsque l'isopropanol est utilisé comme entité acceptrice de sucre. Un alcool secondaire étant moins réactif et moins accessible qu'un alcool primaire les rendements des transglucosylation et transgalactosylation conduisant aux produits **3** et **8** ne dépassent pas les 33 %.

Par la suite l'influence de la longueur de chaîne sur l'activité enzymatique a été évaluée en utilisant le butanol comme accepteur. Ce composé peu polaire et peu miscible à l'eau (i.e. polarité par rapport à l'eau : 4 ; solubilité dans l'eau : 77 g/l soit 0,43 %) a conduit à la formation d'un mélange biphasique pendant la réaction de transglycosylation. Malgré tout, l'enzyme a été capable du transfert des unités glucose et galactose conduisant aux produits réactionnels **4** et **9** avec des rendements comparables à ceux observés pour le MeOH et EtOH. Ce résultat intéressant montre que l'enzyme a la capacité de réaliser l'acte catalytique à l'interface des

phases aqueuse et organique et prouve une fois de plus sa grande stabilité en milieu organique comme décrit précédemment dans ce chapitre (p. 96).

Enfin la capacité de DtGly à transférer deux unités sucres sur un même accepteur a été testée à partir du 1,6-hexanediol. La présence des deux fonctions hydroxyle confère à la molécule une forte polarité, malgré une chaîne alkyl à 6 carbones, et permet ainsi une bonne interaction avec l'enzyme. Ainsi après réaction en présence de pNPGlc et pNPGal deux produits ont pu être isolés de manière quasi-quantitative et leur analyse en RMN a permis de caractériser les produits monoglycosylés **5** et **10**.

	A	B	C	
	pNPMan	pNPRha	pNPGlc	pNPGal
MeOH	n.d.	n.d.	<b>1</b> (54 %)	<b>6</b> (58 %)
EtOH	n.d.	n.d.	<b>2</b> (57 %)	<b>7</b> (57 %)
iPrOH	n.d.	n.d.	<b>3</b> (33 %)	<b>8</b> (31 %)
BuOH	n.d.	n.d.	<b>4</b> (64 %)	<b>9</b> (79 %)
1,2-Hexanediol	n.d.	n.d.	<b>5</b> (86 %)	<b>10</b> (99 %)

**Tableau 11** : Rendement de transglycosylation de DtMan (**A**), DtRha (**B**) et DtGly (**C**), n.d. : produit non détecté.

Afin de s'assurer que la non-réactivité de DtMan et DtRha n'était pas dû à une température réactionnelle trop faible, les mêmes essais ont été réalisés à 60 °C mais n'ont abouti à la formation d'aucun produit de transglycosylation. Lors de la mise au point de nos conditions de réaction nous nous sommes focalisés sur les 2 paramètres « contrôlables » influençant la réaction, à savoir le donneur et l'accepteur de glycoside. Un 3<sup>ème</sup> facteur joue également un rôle important dans cette réaction, il s'agit bien entendu de l'enzyme elle-même. Lors de l'étape de déglycosylation (clivage de l'intermédiaire glycosyl-enzyme) une molécule d'accepteur et une molécule d'eau vont entrer en compétition. Cette compétition est dépendante des propriétés intrinsèques de l'enzyme vis-à-vis de ces 2 entités, caractérisées par le facteur de sélectivité S (Figure 50), précédemment décrit dans l'introduction de ce manuscrit (p. 20). (Hansson, **2001b**)

$$\frac{r_S}{r_H} = S \frac{a_{ROH}}{a_w} \quad \frac{r_S}{r_H} = S_c \frac{[alcool]}{[eau]}$$

**Figure 50** : Mécanisme de transglycosylation et facteurs de sélectivité enzymatique.

Ainsi si l'enzyme présente une meilleure sélectivité vis-à-vis de l'alcool la réaction sera favorisée dans le sens de la transglycosylation, dans le cas inverse celle-ci sera orientée vers l'hydrolyse. Dans le cas de DtMan et DtRha ce facteur de sélectivité ne semble pas en faveur

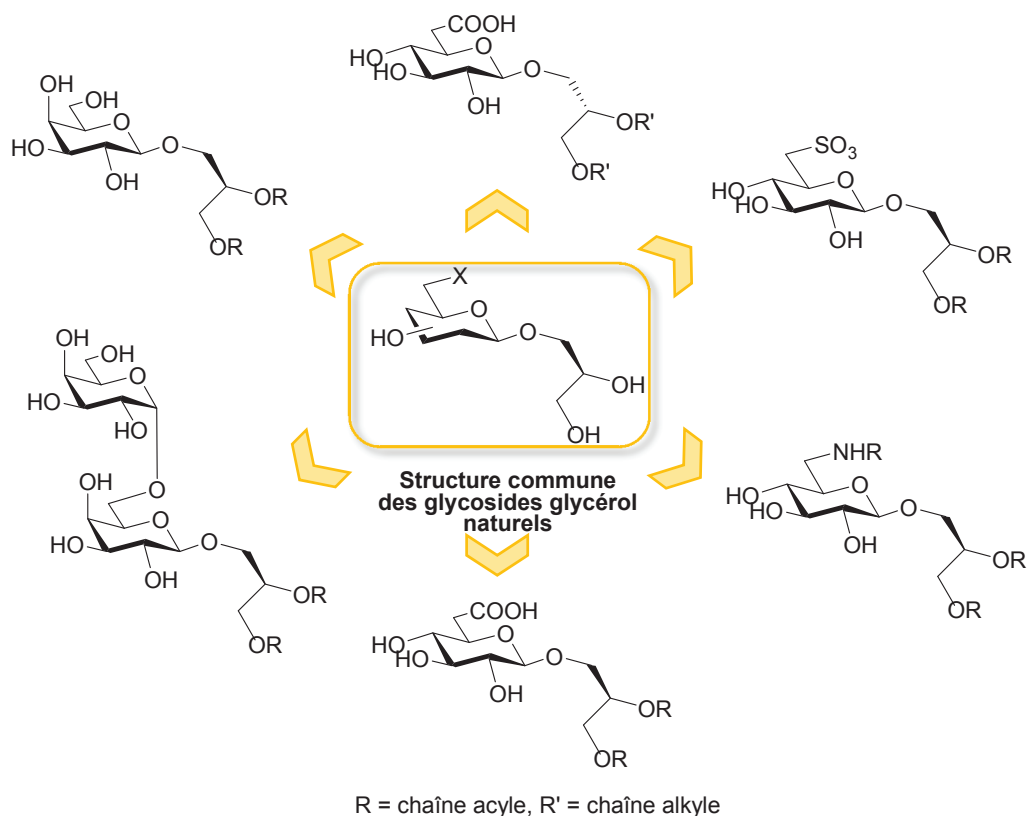
de la réaction de transglycosylation. Cette hypothèse peut être appuyée par les études en présence de co-solvant réalisée sur les trois glycosidases. DtMan et DtRha sont les deux enzymes présentant la moins bonne tolérance à la présence de solvant organique dans le milieu réactionnel (perte de 40 % et 30 % d'activité en présence de 10 % de MeOH respectivement). En revanche pour DtGly qui présente une très bonne activité en transglycosylation nous avons pu démontrer que son activité restait quasiment inchangée même à de forte concentration en solvant organique notamment dans le cas du méthanol. Par ailleurs si DtRha et DtMan possèdent une activité même faible de transglycosylation il ne faut pas oublier que le produit formé peut lui-même être reconnu comme substrat par l'enzyme et donc être hydrolysé à son tour.

Enfin, il est bon de rappeler que parmi l'ensemble des glycosidases disponibles la plupart d'entre elles ne présentent principalement qu'une activité d'hydrolyse et que seulement une minorité de GH a été décrite comme transglycosylases. A l'heure actuelle aucune règle pré-établie ne permet de prédire le potentiel de synthèse d'une glycosidase, ni les résidus clés intervenant dans ce processus. Afin de mieux comprendre les rouages moléculaires influençant le ratio hydrolyse/transglycosylation de nombreuses études, basées sur le principe d'évolution dirigée des enzymes, ont été réalisées ces dernières années. (Bissaro, **2015**; Feng, **2005**; Kone, **2009**; Nashiru, **2001**; Teze, **2014**) Dans la plupart des cas, ces recherches ont abouti à la génération de mutants présentant de meilleures activités de transglycosylation que leurs analogues en type sauvage en remplaçant des résidus situés sur la seconde sphère du site actif. Néanmoins ces recherches sont l'accomplissement d'un énorme travail depuis l'analyse de la structure tridimensionnelle des protéines, la génération de mutants par mutagenèse aléatoire ou dirigée suivi de leur criblage en transglycosylation. La génération de transglycosylases efficaces n'est pas l'objectif recherché ici mais pourra faire l'objet d'une thèse future.

Compte-tenu de ces résultats, parmi les trois enzymes produites DtMan et DtRha se caractérisent par une activité hydrolase exclusive tandis que seule la DtGly présente une activité de transglycosidase. Ce biocatalyseur sera donc utilisé pour la synthèse de O-glycoconjugués.

## 2. Application à la synthèse de glycosides de glycérols

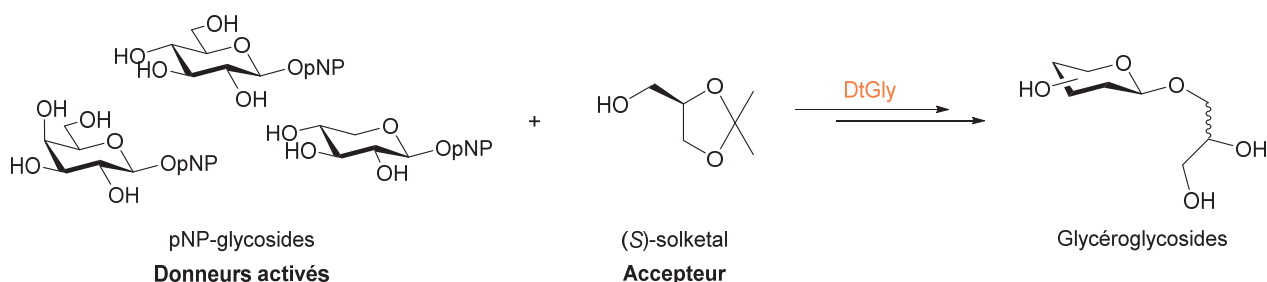
Soucieux de valoriser les biocatalyseurs développés au laboratoire nous nous sommes intéressés à la synthèse chimio-enzymatique de glycosides de glycérols (GGs). Les glycolipides sont les constituants naturels des membranes cellulaires de plantes, algues ou bactéries et assurent entre autre un rôle dans la régulation de la pression osmotique des cellules. (Goedl, **2008**) De manière générale leurs structures s'articule autour d'une unité Glc ou Gal (mono- ou oligosaccharidique) de configuration  $\alpha$  ou  $\beta$  fixée sur le motif 1,2-diacyl-*sn*-glycérol (Figure 51). Cependant des variations peuvent être rencontrées notamment sur le motif sucre (mannose ou rhamnose) ou encore sur le type de fonctions rencontrées (amine, éther, ...). (Holzl, **2007**; Zhang, **2014**)



**Figure 51 :** Structures usuelles de glycosylglycérols naturels. (Zhang, **2014**)

Ces composés suscitent un intérêt particulier puisqu'ils possèdent des propriétés biologiques remarquables telles que des propriétés anti-tumorales, (Colombo, **1996**; Sooriyaarachchi, **2009**) anti-virales (Chirasuwan, **2009**) ou encore anti-inflammatoires. (Zhang, **2014**) Cependant leur faible abondance et les difficultés rencontrées pour leur isolement limite leur étude et a donc favorisé le développement de nombreuses voies de synthèse totale. (Zhang, **2014**) L'intermédiaire clé d'accès à ces composés reste le motif glycosylglycérol qui a aussi démontré des propriétés anti-tumorales mais est également fortement étudié pour ses propriétés tensio-actives, recherchées dans de nombreux domaines, notamment dans les produits de cosmétique comme agent hydratant. (Wei, **2013**)

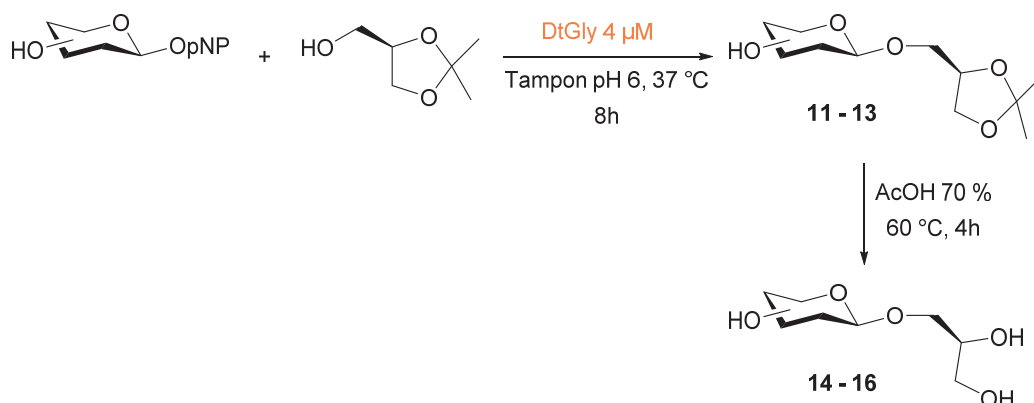
Dans cette optique nous avons souhaité nous intéresser aux propriétés antimicrobiennes de ces GGs en proposant une voie d'accès chimio-enzymatique en deux étapes à partir des pNPGlc, et pNPGal, le pNPXyl et du solketal commercial, forme protégée du glycérol. L'énantiomère (S) du solketal a été sélectionné puisque c'est la forme la plus fréquemment rencontrée dans les structures de GGs et dans un souci de simplification des étapes de purification (Schéma 34).



**Schéma 34** : Synthèse de glycéroglycosides biocatalysée par DtGly.

### 2.1. Voie de synthèse chimio-enzymatique

Notre voie de synthèse de glycoside-S-glycérol s'articule autour de deux étapes. Dans un premier temps le couplage du (S)-solketal commercial sur les pNPsucre sélectionnés est réalisé par transglycosylation en présence de DtGly (4  $\mu$ M) à 37 °C dans un tampon acide citrique/ $\text{Na}_2\text{HPO}_4$  20 mM pH 6 conduisant aux produits de transglycosylation **11**, **12** et **13**. Le groupement protecteur isopropylidène de chaque composé est ensuite clivé par traitement acide dans un milieu acide acétique (70%) à 60 °C conduisant aux dérivés glycosylés **14**, **15** et **16** entièrement déprotégés (**Schéma 35**).



**Schéma 35** : Voie de synthèse des glycosyl-(S)-glycérol.

En présence de pNPGlc et de pNPGal l'étape de transglycosylation a conduit à la formation des dérivés glycosylés **11** et **12** correspondant avec des rendements compris entre 60 et 70 % (Tableau 12). Cette différence de réactivité peut être corrélée avec les valeurs de  $K_M$  de l'enzyme déterminés précédemment pour les deux donneurs. Pour le pNPGlc le  $K_M$  de DtGly est de 0,46 mM contre 1,15 mM pour le pNPGal. Ce qui veut dire que l'enzyme catalyse plus rapidement l'étape de déglycosylation en présence de pNPGal, libérant ainsi la place dans le site actif pour la formation d'un nouvel intermédiaire covalent glycosyl-enzyme. Lorsque le pNPXyl est utilisé comme donneur de glycosyl l'enzyme catalyse la formation du produit de transglycosylation avec un rendement de 50 %. Ce résultat est intéressant car lors des essais d'hydrolyse DtGly avait présenté une faible activité enzymatique vis-à-vis de ce substrat. Au vu du résultat obtenu ici il semble que l'enzyme soit capable du transfert d'une unité Xylose vers un accepteur, autre que l'eau, de manière aussi efficace qu'en présence de pNPGlc. Cette hypothèse pourrait être confirmée en mesurant et en comparant les constantes

cinétiques de l'enzyme vis-à-vis de ces deux substrats en présence de solketal. La déprotection des dérivés glycosylés par traitement acide a permis d'obtenir le 1-O-( $\beta$ -D-glucopyranosyl)-(S)-glycérol **14**, le 1-O-( $\beta$ -D-galactopyranosyl)-(S)-glycérol **15** et le 1-O-( $\beta$ -D-xylopyranosyl)-(S)-glycérol **16** avec des rendements de 50 à 62 %. Malgré une réaction totale, des difficultés ont été rencontrées lors de la purification de ces trois composés justifiant les rendements en-dessous de ceux classiquement obtenus dans ce type de réaction (> 90 %). (Catani, 2007)

Donneur	Transglycosylation	Déprotection	Rendement global
pNP- $\beta$ -D-Glc	<b>11</b> (56 %)	<b>14</b> (62 %)	35 %
pNP- $\beta$ -D-Gal	<b>12</b> (73 %)	<b>15</b> (50 %)	37 %
pNP- $\beta$ -D-Xyl	<b>13</b> (51 %)	<b>16</b> (62 %)	32 %

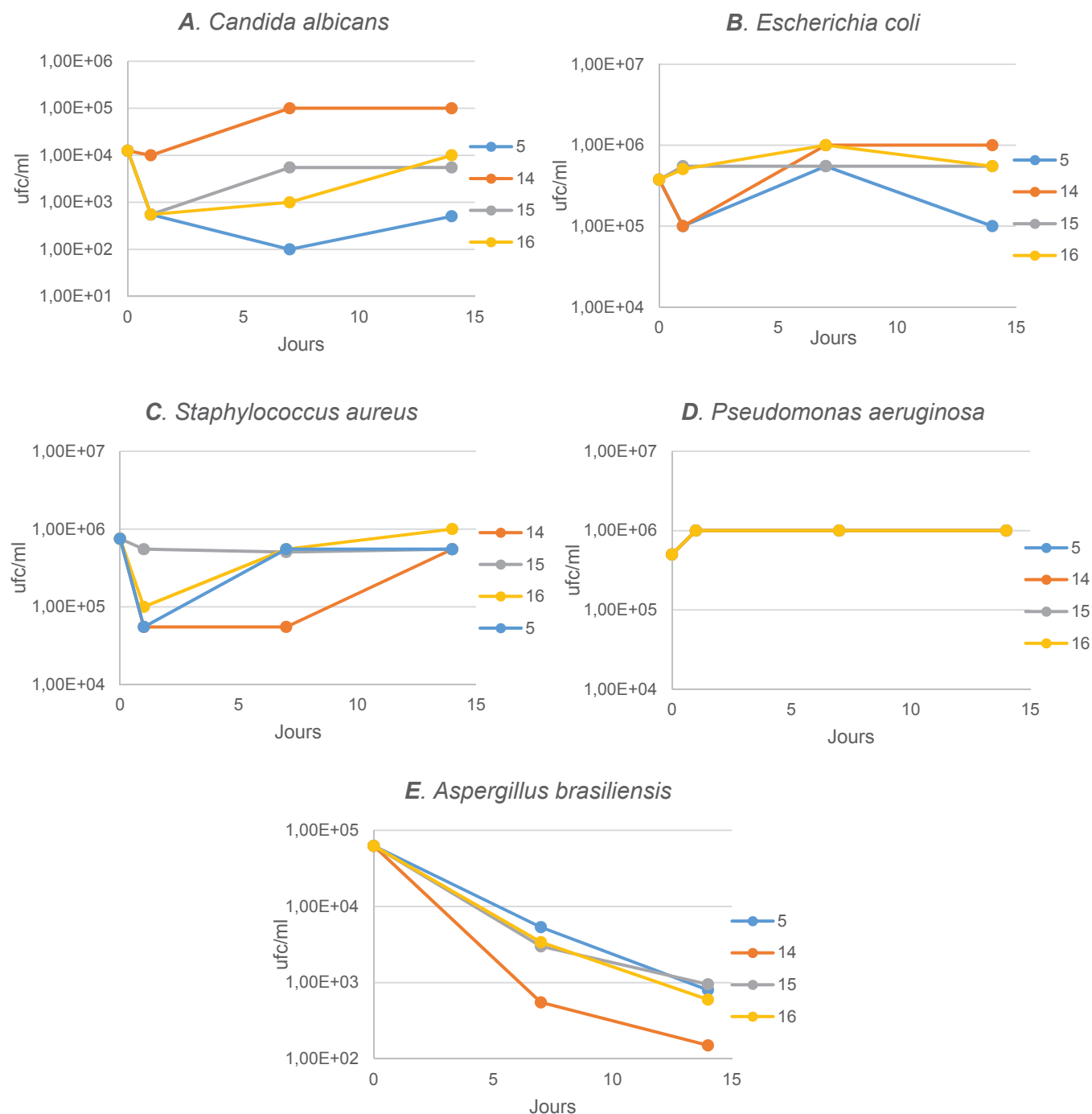
**Tableau 12 :** Tableau récapitulatif des rendements obtenus pour la synthèse d'O-glycoconjugués par DtGly.

Au terme de cette voie de synthèse en deux étapes une série de trois glycosyl-(S)-glycérol ont été obtenus avec un rendement global supérieur à 30 %. Les propriétés antibactériennes de ces différents composés ont pu alors être déterminées à partir de la méthode de Microchallenge test développée par l'entreprise GLYcoDiag.

## 2.2. Recherche de propriétés antimicrobiennes des glycoconjugués synthétisés

Les agents de préservation sont des composés capables de limiter la croissance de micro-organismes sur une longue durée et sont largement retrouvés dans la formulation de produits du quotidien (issus de l'industrie agro-alimentaire, cosmétique ou pharmaceutique). En tant qu'agents antimicrobiens leur utilisation est nécessaire pour limiter la prolifération d'organismes potentiellement pathogènes et ainsi garantir une sécurité pour le consommateur. Les qualités préservatives d'un composé sont déterminées par la méthode de challenge test qui consiste à évaluer le comportement d'un composé potentiellement biologiquement actif sur une variété de souches microbiennes. Dans le domaine de la cosmétique, la réglementation européenne actuelle (EN ISO11930) préconise l'utilisation des souches microbiennes *Pseudomonas aeruginosa* (Pa) ATCC 9027, *Staphylococcus aureus* (Sa) ATCC 6538, *Escherichia coli* (Ec) ATCC 8739, *Candida albicans* (Ca) ATCC 10231 et *Aspergillus brasiliensis* (Ab) ATCC 16404 pour tester les propriétés préservatives d'un composé cible. Ces essais ont été réalisés en partenariat avec l'entreprise française GLYcoDiag spécialisée dans la gestion de projet de glycobiologie et le développement d'outils de glycoanalyse. Son expertise est basée sur l'utilisation d'une librairie de lectines permettant la détermination de profil d'interactions (GLYcoPROFILE®) pour l'identification de glycanes complexes libres en solution ou intégrés sur des membranes. Elle propose également des services de microanalyses, notamment dans le domaine de la cosmétique, et a développé une méthodologie dérivée du challenge test permettant le criblage d'activités antimicrobiennes sur microplaque, aussi appelée microchallenge test. (El Abdellaoui, 2010)

La série de glycéroglycosides composée du 1-O-( $\beta$ -D-glucopyranosyl)-(S)-glycérol **14**, du 1-O-( $\beta$ -D-galactopyranosyl)-(S)-glycérol **15** et du 1-O-( $\beta$ -D-xylopyranosyl)-(S)-glycérol **16** synthétisé ainsi que le glucopyranoside d'hexyle **5** ont ainsi été testé par microchallenge test selon le protocole décrit par GLYCoDiag. Dans un premier temps les solutions de chaque échantillon ont été préparées en sérum physiologique pour constituer la gamme de concentrations suivantes : 12,5  $\mu\text{g/ml}$ , 25  $\mu\text{g/ml}$ , 50  $\mu\text{g/ml}$ , 100  $\mu\text{g/ml}$  et 500  $\mu\text{g/ml}$ . Par la suite, chaque échantillon a été contaminé par une quantité fixe, exprimée en unité formant les colonies (ufc), de chacune des souches décrites ci-dessus, suivi d'une incubation à température ambiante. Pour chaque souche un échantillon blanc a été réalisés et placé dans les mêmes conditions d'incubation que les échantillons test. Le taux résiduel de contamination a été mesuré par prélèvement après 24h, 7 jours et 14 jours d'incubation. Les résultats obtenus après 2 semaines d'incubation ont permis de tracer les courbes d'évolution de la population microbienne de chaque souche en fonction des 4 composés testés (Figure 52).



**Figure 52 :** Evolution de la population de *Candida albicans* (A), *Escherichia coli* (B), *Staphylococcus aureus* (C), *Pseudomonas aeruginosa* (D) et *Aspergillus brasiliensis* (E) en fonction des dérivés glycosylés 5, 14, 15 et 16.

Au terme de l'expérience, seul le 1-O-( $\beta$ -D-glucopyranosyl)-(S)-glycerol **14** a présenté une activité antifongique spécifique (E, Figure 52). En effet par comparaison avec les autres molécules testées, le composé **14** possède une activité intéressante puisqu'il est capable de réduire la population du champignon *Aspergillus brasiliensis* d'un facteur 1000 en 14 jours contrairement aux autres composés qui présentent une activité plus faible sur la même



période. A l'heure actuelle le développement de nouveaux antifongiques est un réel besoin face à la multiplication des phénomènes de résistances aux traitements disponibles. (Tebbetts, **2012**) Le composé **14** est donc un candidat potentiel pour une application antifongique spécifique. Néanmoins son spectre d'activité n'est pas assez étendu pour répondre aux exigences des normes européennes de la cosmétique pour rentrer dans une formulation de produit en tant qu'agent de préservation. Dans la Nature des glycosides glycérol sont fonctionnalisés par une ou plusieurs chaînes grasses. (Holzl, **2007**; Zhang, **2014**) L'addition de chaînes alkyles de longueur variable comme nouveau degré de diversité sur ces O-glycoconjugués synthétisés est une piste à explorer pour améliorer leurs propriétés antimicrobiennes.

#### IV. Conclusion

A partir du génome de la bactérie thermophile *D. thermophilum*, cinq gènes codant pour de potentielles glycosidases ont été sélectionnés. Parmi ces derniers, les trois gènes *dicth\_0289*, *dicth\_0359* et *dicth\_1692* ont pu être clonés dans le plasmide pET28-a(+) et transformés chez *E.coli* Rosetta-DE(3). Chaque protéine a été produite et purifiée pour obtenir une forme pure nécessaire à leur caractérisation. Dans un premier temps les spécificités de substrat de ces enzymes ont été déterminées en utilisant un panel de pNPsucre disponibles au laboratoire, conduisant alors à la caractérisation de la **DtMan** et de la **DtRha** présentant une spécificité exclusive pour le pNP- $\beta$ -D-Man et le pNP- $\alpha$ -L-rhamnose respectivement. La 3<sup>ème</sup> enzyme supposée être une  $\beta$ -D-glucosidase s'est révélée quant à elle très versatile vis-à-vis du substrat puisqu'elle est capable de reconnaître jusqu'à 5 motifs sucres (pNP- $\beta$ -D-Glc, pNP- $\beta$ -D-Gal, pNP- $\beta$ -D-Fuc, pNP- $\beta$ -D-Xyl et pNP- $\beta$ -D-Man) et que nous avons donc caractérisé comme la **DtGly** de *D. thermophilum*.

Par la suite l'ensemble des constantes cinétiques des enzymes ont été déterminées ainsi que les conditions optimales de pH et de température pour la réaction d'hydrolyse. L'activité des glycosidases a également été évaluée en présence de différents co-solvants qui pourraient être utilisés dans les réactions de glycosylation pour faciliter la solubilisation d'accepteurs de glycoside hydrophobes. A l'issue de cette étude la propriété de thermorésistance a pu être confirmée pour les trois enzymes avec un maximum d'activité d'hydrolyse enregistré pour des températures  $\geq 80$  °C. DtMan et DtRha présentent une relative indépendance vis-à-vis du pH puisqu'elles sont capables de catalyser le clivage de la liaison glycosidique sur des gammes de pH étendue (pH 4 à pH 7 pour DtRha et pH 5 à pH 7 pour DtMan) tandis que DtGly semble plus sensible au pH. Ces glycosidases ont démontré leur tolérance vis-à-vis de l'ajout de méthanol dans le milieu réactionnel puisque dans les trois cas plus de 5 % d'activité enzymatique sont conservés dans un mélange 50:50 méthanol/eau.

Des études de cristallogénèse ont été réalisées afin de résoudre les structures tridimensionnelles de nos différents biocatalyseurs. Dans le cas de la DtMan et de la DtRha les conditions optimales de cristallisation ont été déterminées et les différents cristaux ont pu être analysés aux RX sur les lignes Proxima 1 et 2 du Synchrotron Soleil à St Aubin (91). Dans

le cas de DtMan un modèle préliminaire a pu être résolu mettant en évidence l'organisation cristalline sous forme de tétramère de l'enzyme et les rôles des Glutamates E425 et E508 comme les résidus acide/base et nucléophile respectivement. En revanche malgré l'ensemble des conditions testées aucun cristal n'a pu être obtenu dans le cas de la DtGly, des essais sont toujours en cours pour tenter d'y parvenir.

Enfin le potentiel de transglycosylation des enzymes a été évalué sur la synthèse de glycosides d'alkyle. Cette étude a permis de mettre en évidence la glycosidase DtGly alors capable de catalyser la réaction de transglycosylation à partir de donneurs activés, pNPGlc et pNPGal, sur un panel d'alcools aliphatiques avec des rendements de 30 à 99 %. Cette enzyme répond également à notre volonté de proposer des biocatalyseurs éco-compatibles comme alternatives aux procédés chimiques puisque l'ensemble des composés synthétisés a été obtenu après 8h de réaction à 37 °C en milieu aqueux. En revanche aucune activité transglycosidase n'a été détectée pour DtMan et DtRha laissant supposer que les propriétés intrinsèques de ces enzymes étaient en faveur de l'activité d'hydrolyse exclusivement. Devant le potentiel de DtGly, sa valorisation a été envisagée pour développer une voie d'accès chimio-enzymatique en deux étapes aux glycéroglycosides, composés d'intérêt dans les domaines de la cosmétique et de la santé humaine. Trois composés ont pu ainsi être synthétisés avec un rendement global supérieur à 30 %. Ces derniers ont par la suite été testés biologiquement afin d'évaluer leurs propriétés antimicrobiennes par microchallenge test en partenariat avec l'entreprise GLYcoDiag. A l'issue du test le 1-O-( $\beta$ -D-glucopyranosyl)-(S)-glycérol **14** a présenté une activité intéressante antifongique sur la souche *Aspergillus brasiliensis*.

A partir de ces glycosidases natives entièrement caractérisées il est à présent possible de se tourner vers la production des mutants thioglycoligases correspondants et leur évaluation pour la synthèse de S-glycoconjugués.



## Bibliographie

- Bauer, M. W., Bylina, E. J., Swanson, R., and Kelly, R. M.; *J. Biol. Chem.* **(1996)**, 271, 23749-23755
- Birgisson, H., Hreggvidsson, G. O., Fridjónsson, O. H., Mort, A., Kristjánsson, J. K., and Mattiasson, B.; *Enzyme Microb. Technol.* **(2004)**, 34, 561-571
- Bissaro, B., Durand, J., Biarnés, X., Planas, A., Monsan, P., O'Donohue, M. J., and Fauré, R.; *ACS Catal.* **(2015)**, 4598-4611
- Bojarova, P., and Kren, V.; *Trends Biotechnol.* **(2009)**, 27, 199-209
- Butters, T. D.; *Curr. Opin. Chem. biol.* **(2007)**, 11, 412-418
- Cairns, J. R. K., and Esen, A.; *Cell. Mol. Life Sci.* **(2010)**, 67, 3389-3405
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B.; *Nucleic Acids Res.* **(2009)**, 37, D233-238
- Cateni, F., Bonivento, P., Procida, G., Zacchigna, M., Gabrielli Favretto, L., Scialino, G., and Banfi, E.; *Bioorg. Med. Chem.* **(2007)**, 15, 815-826
- Chirasuwan, N., Chaiklahan, R., Kittakoop, P., Chanasattru, W., Ruengjitchatchawalya, M., Tanticharoen, M., and Bunnag, B.; *Sci. Asia* **(2009)**, 35, 137-141
- Cobucci-Ponzano, B., Aurilia, V., Riccio, G., Henrissat, B., Coutinho, P. M., Strazzulli, A., Padula, A., Corsaro, M. M., Pieretti, G., Pocsfalvi, G., Fiume, I., Cannio, R., Rossi, M., and Moracci, M.; *J. Biol. Chem.* **(2010)**, 285, 20691-20703
- Coil, D. A., Badger, J. H., Forberger, H. C., Riggs, F., Madupu, R., Fedorova, N., Ward, N., Robb, F. T., and Eisen, J. A.; *Genome Announc.* **(2014)**, 2, 109-114
- Colombo, D., Scala, A., Taino, I. M., Torna, L., and Ronchetti, F.; *Bioorg. Med. Chem.* **(1996)**, 1187-1190
- Cui, Z., Maruyama, Y., Mikami, B., Hashimoto, W., and Murata, K.; *J. Mol. Biol.* **(2007)**, 374, 384-398
- Dhawan, S. K., J.; *Crit. Rev. Biotechnol.* **(2007)**, 27, 197-216
- Dion, M., Fourage, L., Hallet, J. N., and Colas, B.; *Glycoconjugate J.* **(1999)**, 16, 27-37
- Drew, M. G. B., Ennis, S. C., Gridley, J. J., Osborn, H. M. I., and Spackman, D. G.; *Tetrahedron* **(2001)**, 57, 7919-7937
- Duffaud, G. D., McCutchen, C. M., Leduc, P., Parker, K. N., and Kelly, R. M.; *Appl. Microbiol. Biotechnol.* **(1997)**, 63, 169-177
- El Abdellaoui, S., Destandau, E., Toribio, A., Elfakir, C., Lafosse, M., Renimel, I., Andre, P., Cancellieri, P., and Landemarre, L.; *Anal. Bioanal. Chem.* **(2010)**, 398, 1329-1338
- Feng, H. Y., Drone, J., Hoffmann, L., Tran, V., Tellier, C., Rabiller, C., and Dion, M.; *J. Biol. Chem.* **(2005)**, 280, 37088-37097
- Fujimoto, Z., Jackson, A., Michikaa, M., Maehara, T., Momma, M., Henrissat, B., Gilbert, H. J., and Kaneko, S.; *J. Biol. Chem.* **(2013)**, 288, 12376-12385
- Gantt, R. W., Peltier-Pain, P., and Thorson, J. S.; *Nat. Prod. Rep.* **(2011)**, 28, 1811-1853
- Goedel, C., Sawangwan, T., Mueller, M., Schwarz, A., and Nidetzky, B.; *Angew. Chem. Int. Ed.* **(2008)**, 47, 10086-10089
- Gomes, J., Terler, K., Kratzer, R., Kainz, E., and Steiner, W.; *Enzyme Microb. Technol.* **(2007)**, 40, 969-975
- Goyal, K., Selvakumar, P., and Hayashi, K.; *J. Mol. Catal. B: Enzym.* **(2001)**, 15, 45-53

- Hansson, T., and Adlercreutz, P.; *Biotechnol. Bioeng.* **(2001a)**, 75, 656-665
- Hansson, T., Andersson, M., and Adlercreutz, P.; *Enzyme Microb. Technol.* **(2001b)**, 29, 527-534
- Henrissat, B.; *Biochem. J.* **(1991)**, 280, 309-316
- Holzl, G., and Dormann, P.; *Prog. Lipid Res.* **(2007)**, 46, 225-243
- Kone, F. M., Le Behec, M., Sine, J. P., Dion, M., and Tellier, C.; *Protein Eng., Des. Sel.* **(2009)**, 22, 37-44
- Li, H., Voutilainen, S., Ojamo, H., and Turunen, O.; *Enzyme Microb. Technol.* **(2015)**, 70, 66-71
- Lindberg, B., Leontein, K., Lindquist, U., Svenson, S. B., Wrangsell, G., Dell, A., and Rogers, M.; *Carbohydr. Res.* **(1988)**, 174, 313-322
- Mathrani, I. M., and Ahring, B. K.; *Appl. Microbiol. Biotechnol.* **(1992)**, 38, 23-27
- Matsui, I., Sakai, Y., Matsui, E., Kikuchi, H., Kawarabayasi, Y., and Honda, K.; *FEBS Lett.* **(2000)**, 467, 195-200
- Maxam, A. M., and Gilbert, W.; *Proc. Natl. Acad. Sci. U.S.A.* **(1977)**, 74, 560-564
- Meynial-Salles, I. C., D.; *J. Biotechnol.* **(1996)**, 46, 1-14
- Mladenoska, I., Winkelhausen, E., and Kuzmanova, S.; *Food Technol. Biotechnol.* **(2008)**, 46, 311-316
- Nashiru, O., Zechel, D. L., Stoll, D., Mohammadzadeh, T., Warren, R. A. J., and Withers, S. G.; *Angew. Chem., Int. Ed.* **(2001)**, 40, 417-420
- O'Neill, E., Stevenson, C. E. M., Paterson, M. J., Rejzek, M., Chauvin, A. L., Lawson, D. M., and Field, R. A.; *Proteins* **(2015)**,
- Offen, W. A., Zechel, D. L., Withers, S. G., Gilbert, H. J., and Davies, G. J.; *Chem. Commun. (Cambridge, U.K.)* **(2009)**, 18, 2484-2486
- Park, S. H., Park, K. H., Oh, B. C., Alli, I., and Lee, B. H.; *New Biotechnol.* **(2011)**, 28, 640-648
- Rabausch, U., Ilmberger, N., and Streit, W. R.; *J. Biotechnol.* **(2014)**, 191, 38-45
- Reichardt, C.; *Solvents and Solvent Effects in Organic Chemistry* **(2003)**, 389-469
- Saiki, T., Kobayashi, Y., Kawagoe, K., and Beppu, T.; *Int. J. Syst. Bacteriol.* **(1985)**, 35, 253-259
- Sanger, F., and Coulson, A. R.; *FEBS Lett* **(1978)**, 87, 107-110
- Shi, R., Li, Z., Ye, Q., Xu, J., and Liu, Y.; *Biores. Technol.* **(2013)**, 142, 338-344
- Sooriyaarachchi, S., Ubhayasekera, W., Boos, W., and Mowbray, S. L.; *FEBS J.* **(2009)**, 276, 2116-2124
- Sugita, M., Sanai, Y., Itonori, S., and Hori, T.; *Biochim. Biophys. Acta* **(1988)**, 962, 159-165
- Tailford, L. E., Money, V. A., Smith, N. L., Dumon, C., Davies, G. J., and Gilbert, H. J.; *J. Biol. Chem.* **(2007)**, 282, 11291-11299
- Tebbets, B., Stewart, D., Lawry, S., Nett, J., Nantel, A., Andes, D., and Klein, B. S.; *PLoS One* **(2012)**, 7, 1-11
- Teze, D., Hendrickx, J., Czjzek, M., Ropartz, D., Sanejouand, Y. H., Tran, V., Tellier, C., and Dion, M.; *Protein Eng., Des. Sel.* **(2014)**, 27, 13-19
- Thibodeaux, C. J., Melancon, C. E., and Liu, H. W.; *Angew. Chem., Int. Ed. Engl.* **(2008)**, 47, 9814-9859

- Toone, E. J., Simon, E. S., Bednarski, M. D., and Whitesides, G. M.; *Tetrahedron* **(1989)**, *45*, 5365-5422
- Turner, P., Svensson, D., Adlercreutz, P., and Karlsson, E. N.; *J. Biotechnol.* **(2007)**, *130*, 67-74
- Vinogradov, E., Pertersen, B. O., and Duus, J. O.; *Carbohydr. Res.* **(2000)**, *325*, 216-221
- Wang, L. X., and Huang, W.; *Curr. Opin. Chem. biol.* **(2009)**, *13*, 592-600
- Wei, W., Qi, D., Zhao, H. Z., Lu, Z. X., Lv, F., and Bie, X.; *Food Chem.* **(2013)**, *141*, 3085-3092
- Weignerova, L., Marhol, P., Gerstorferova, D., and Kren, V.; *Biores. Technol.* **(2012)**, *115*, 222-227
- Zeikus, J. G.; *Enzyme Microb. Technol.* **(1979)**, *1*, 243-252
- Zhang, J., Li, C., Yu, G., and Guan, H.; *Mar. Drugs* **(2014)**, *12*, 3634-3659
- Zhang, M., Jiang, Z. Q., Li, L. T., and Katrolia, P.; *J. Mol. Catal. B: Enzym.* **(2009)**, *60*, 119-124
- Zou, Z. Z., Yu, H. L., Li, C. X., Zhou, X. W., Hayashi, C., Sun, J., Liu, B. H., Imanaka, T., and Xu, J. H.; *Biores. Technol.* **(2012)**, *118*, 425-430
- Zverlov, V. V., Hertel, C., Bronnenmeier, K., Hroch, A., Kellermann, J., and Schwarz, W.; *Mol. Microbiol.* **(2000)**, *35*, 173-179



**Chapitre 2**  
**Conception et évaluation de mutants**  
**thioglycoligases pour la synthèse de S-**  
**glycoconjugués**

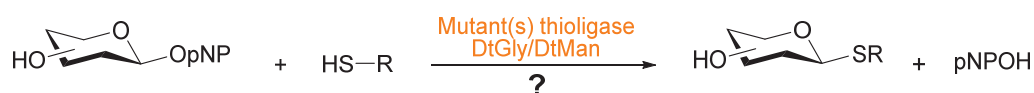






## Chapitre 2 – Conception et évaluation de mutants thioglycoligases pour la synthèse de S-glycoconjugués

L'objectif de notre projet repose sur la validation du concept de thioglycosylation enzymatique d'acides aminés en utilisant la technologie des thioglycoligases. Suite à l'étude réalisée sur notre banque de GH natives de *D. thermophilum*, notre choix s'est naturellement porté sur la seule représentante ayant démontré une activité importante en synthèse, la glycosidase versatile DtGly. A l'heure actuelle nos connaissances sur les thioglycoligases ne sont pas assez étendues pour répondre à une question fondamentale : l'activité de transglycosylation de l'enzyme native est-elle une condition nécessaire pour faire une bonne thioligase ? A ce jour aucune étude n'a été rapportée concernant cette problématique. Pour tenter d'y répondre nous nous sommes alors intéressés à la mannosidase DtMan qui n'a présenté jusqu'ici aucune activité de transglycosylation, en générant une banque de variants basée sur le concept de thioglycoligase (Schéma 36).



**Schéma 36** : Etude de l'activité thioligase de mutants de DtGly et DtMan.

Ce deuxième chapitre sera donc consacré dans sa majeure partie à la conception de mutants de DtGly par mutagenèse dirigée ainsi qu'à leur caractérisation en activité thioligase. A l'issue de ces étapes, le meilleur candidat retenu sera utilisé pour la synthèse de S-glycoconjugués, étape préliminaire à la glycosylation d'acides aminés, qui nous permettra de définir le motif type d'accepteur reconnu par le biocatalyseur. La seconde partie de ce chapitre, plus fondamentale, portera sur l'étude de 7 variants, type thioligase, de DtMan. Chaque variant sera caractérisé et comparé à l'enzyme en type sauvage de manière à évaluer l'influence de chaque mutation sur l'activité enzymatique. D'autre part, les conditions de cristallogenèse de DtMan *wt* ayant été optimisées nous nous sommes inspirés de ces conditions pour l'obtention de cristaux de variants de DtMan. La résolution et comparaison des structures de DtMan *wt* et mutée nous permettra de compléter l'étude de relation structure/activité de cette enzyme. Avant d'aborder ces deux sujets, quelques rappels sur le concept de thioglycoligase sont présentés ainsi que la stratégie suivie pour leur conception.

### I. Thioglycoligases : rappels et généralités

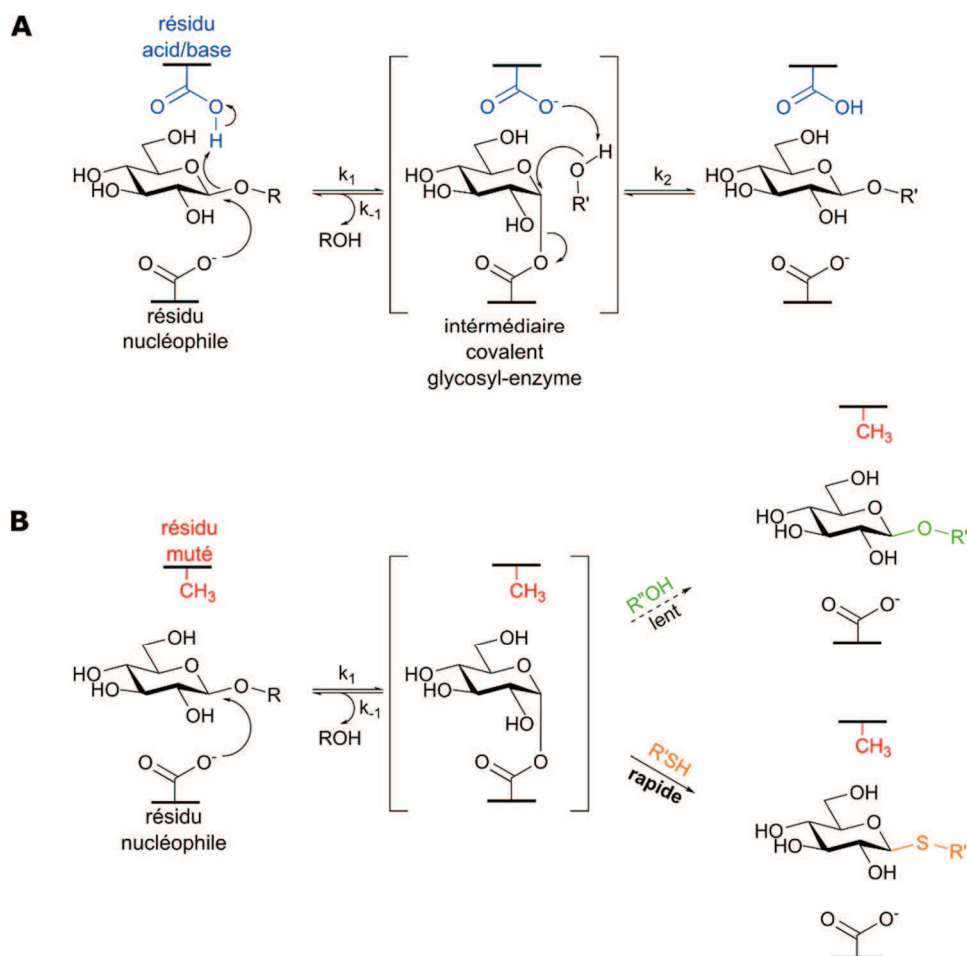
#### 1. Rappels

Les thioglycosides, glycosides dont l'atome d'oxygène en position anomérique a été remplacé par un atome de soufre, sont caractérisés par des propriétés physico-chimiques remarquables et notamment par une relative résistance face à l'hydrolyse chimique ou enzymatique. Ces composés ont fait l'objet de nombreuses études et se sont révélés être de précieux outils en synthèse d'une part, comme donneurs de glycosides, mais également en

tant qu'intermédiaires stables pour l'analyse cristallographique de nombreuses structures de protéines (Driguez, **2001**) ou encore comme inhibiteurs efficaces de glycosidases. (Wardrop, **2010**)

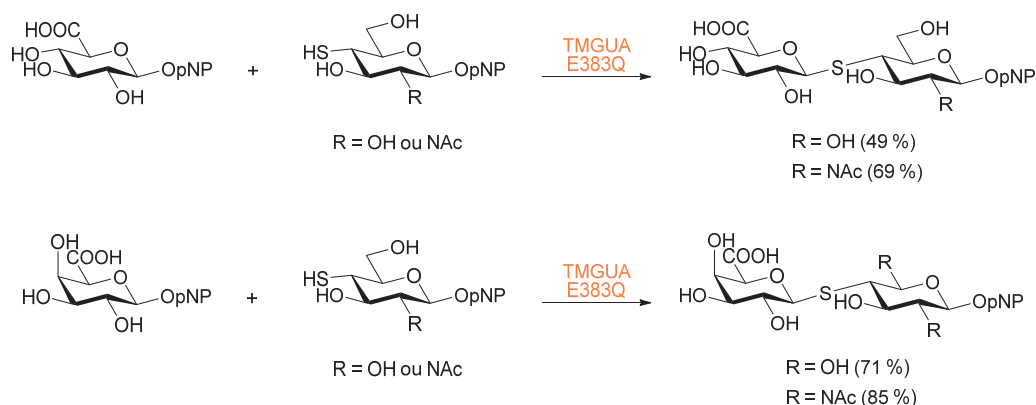
En chimie organique leur étude a favorisé le développement de nombreuses méthodologies de synthèse (substitution de type  $SN_2$ , (Fort, **2001**) addition de type Michael (Witczak, **1995**) ou plus récemment par couplage métallo-catalysé (Bruneau, **2015**)) qui requièrent pour la plupart le passage par de complexes étapes de protection/déprotection ainsi que le contrôle de la configuration du centre anomérique. Dans le même temps, les avancées réalisées dans le domaine des GH, couplées aux techniques de bio ingénieries, ont permis la génération de biocatalyseurs efficaces, les glycosynthases, pour la synthèse d'O-oligosaccharides. Fort de ces résultats prometteurs, l'équipe du Prof. Withers a choisi d'étendre ce potentiel en proposant la première génération de glycosidases mutées (Abg E170A d'*A. sp* et Man2A E249A de *C. fimi*) capables de créer des liaisons S-glycosidiques : les thioglycoligases. (Jahn, **2003**)

Ces enzymes sont générées à partir de glycosidases à rétention de configuration dont le résidu acide/base a été remplacé par un acide aminé catalytiquement inactif. Pour assurer l'activité de synthèse de tels mutants il est alors essentiel d'utiliser des donneurs et accepteurs de glycosides compatibles avec le mécanisme à double déplacement rencontré chez les glycosidases à rétention de configuration (Schéma 37). La première étape qui est caractérisée par la formation de l'intermédiaire glycosyl-enzyme n'est favorisée que pour des donneurs présentant des aglycones électrofuges, tels que les dinitrophényles, qui ne requièrent pas de catalyse acide pour initier leur départ. Lors de la seconde étape, l'absence d'activation basique dans le site actif rend la transglycosylation de nucléophiles de type alcools relativement lente, voire impossible. L'augmentation du turnover ( $k_{cat}$ ) de déglycosylation n'est donc assurée que par l'intervention de nucléophiles fort tels que les thiols, formant les thioglycosides correspondants.



**Schéma 37 : (A) mécanisme des glycosidases à rétention de configuration ; (B) mécanisme des thioglycoligases.**

D'autre part si les rôles du donneur et de l'accepteur de glycosides sont prépondérants, il apparaît que le choix de la mutation du résidu acide/base soit également d'une importance significative. En effet Mullegger *et coll.* ont réalisé le screening d'une banque de variants de la  $\beta$ -glucosidase Abg E170 d'*A. sp.*, produits par mutagenèse saturante, et ont mis en évidence le potentiel de thioglycosylation du mutant E170Q présentant une efficacité catalytique jusqu'à 100 fois supérieure au mutant E170A. (Mullegger, **2005**) En 2006 la même équipe a rapporté la synthèse de thiodisaccharides à partir des acides glucuronique et galacturonique avec de bons rendements (de 49 à 85 %) à partir des mutants Glutamine et Alanine de la  $\beta$ -glucuronidase de *T. maritima* respectivement (Schéma 38). (Mullegger, **2006**)



**Schéma 38** : Synthèse enzymatique de  $\beta$ -(1,4) thiodisaccharides par TMGUA E388Q et E388A.

A l'heure actuelle 8 thioglycoligases ont été décrites dans la littérature (Tableau 1, p. 36 - 37) pour la synthèse de liaisons thioglycosidiques de type  $\beta$ -(1,3) ou  $\beta$ -(1,4), à l'exception des thioglycoligases YicI D482 d'*E.coli* et MalA D416 de *S. solfataricus* permettant la synthèse de liaisons de type  $\alpha$ -(1,4) et  $\alpha$ -(1,6). (Kim, **2006**) Enfin le variant E173Q de l'Araf51 de *C. thermocellum* développé par Almendros *et coll.* représente le seul exemple de thioglycoligase utilisé pour la synthèse de thioglycosides à partir d'accepteurs non glycosylés de type thioarylique. (Almendros, **2011**; Almendros, **2013**)

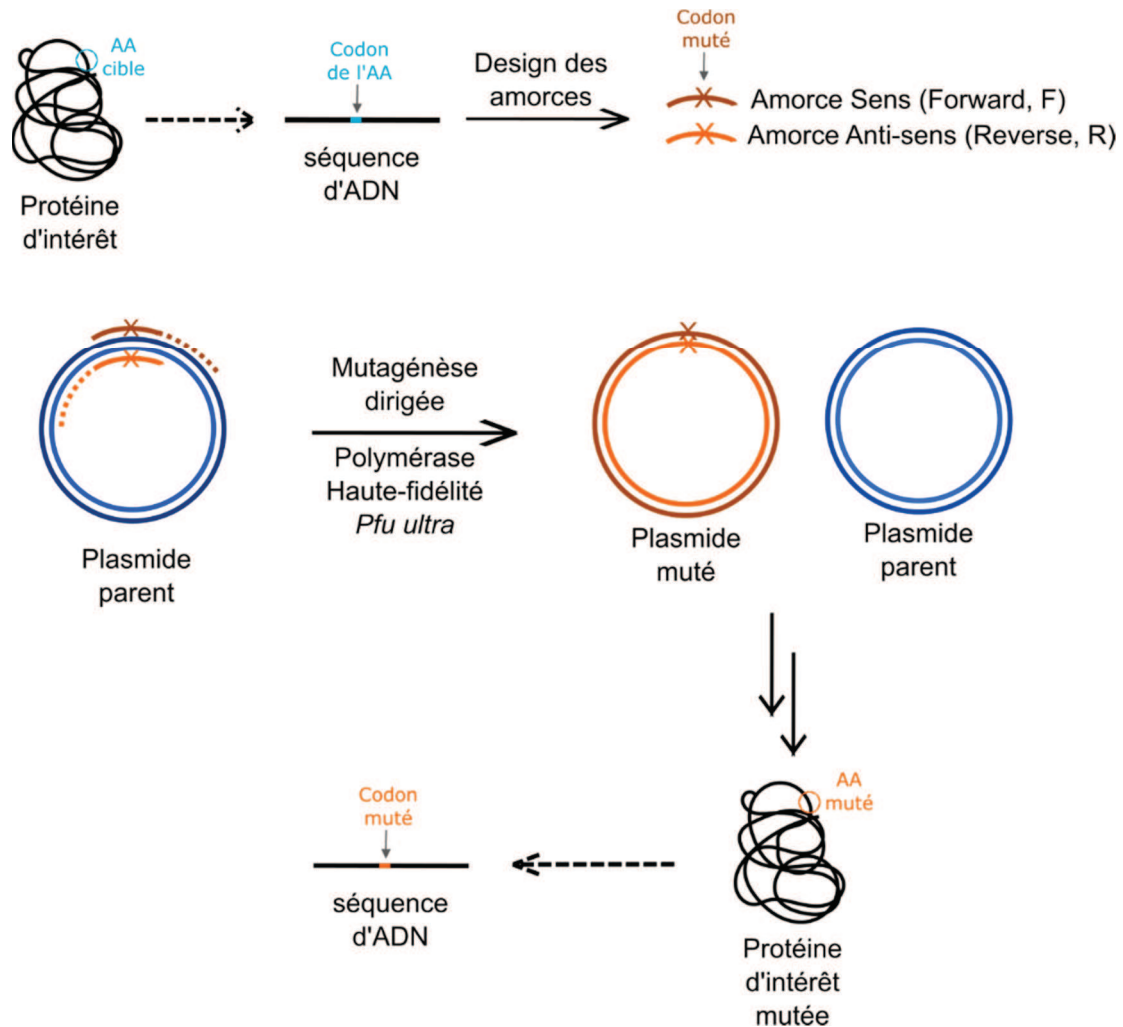
## 2. Création de points de mutation sélectifs par mutagenèse dirigée

Grâce à l'ensemble des données disponibles à ce jour sur les thioglycoligases nous avons pu orienter nos choix pour la génération de mutants thioglycoligases de *D. thermophilum*. Dans le cas de DtGly l'objectif visé est de proposer un catalyseur efficace permettant la création de liaison  $\beta$ -1-thioglycosidique. Dans cette optique, il nous a paru judicieux de se focaliser sur la génération des mutants les plus actifs décrits dans la littérature, à savoir les variants Glutamine et Alanine en lieu et place du résidu acide/base. Pour la mannosidase DtMan l'approche est différente puisqu'il s'agit ici de voir si la suppression de l'activité d'hydrolyse de l'enzyme permet d'améliorer son efficacité en synthèse. Ainsi il nous a semblé pertinent de faire varier le type de mutation et ainsi détecter l'effet potentiel du type de résidu sur l'activité enzymatique.

Afin d'obtenir les mutations désirées pour les deux enzymes, l'ensemble des mutants a été produit par mutagenèse dirigée. Cette technique développée par le Prof. Smith en 1979 consiste à incorporer un point de mutation fixe sur un site spécifique d'une séquence d'ADN parent. (Gillam, **1979**) Pour réaliser cette manipulation génétique plusieurs éléments sont nécessaires (Schéma 39) :

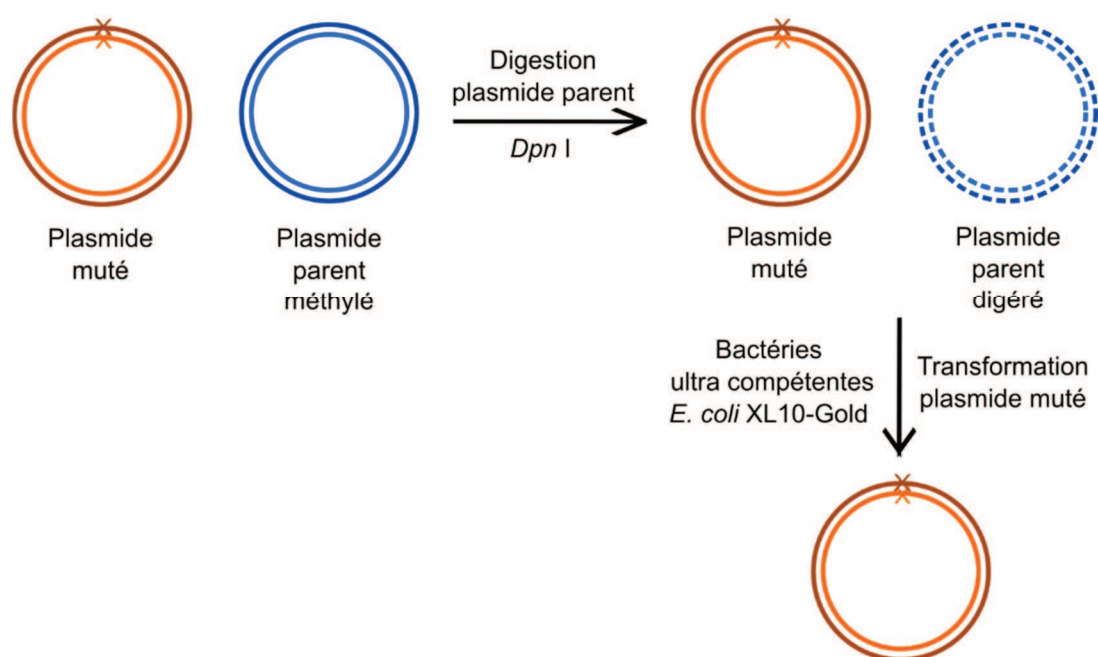
- i) connaître la position précise du résidu à muter sur la séquence en AA de la protéine cible,
- ii) l'ADN plasmidique « parent » contenant le gène de la protéine cible, dans le cas présent les plasmides recombinants de DtGly et DtMan,

- iii) les amorces sens et anti-sens qui serviront de point de départ pour la PCR. Ces courtes séquences d'acides nucléiques (40 à 60 nucléotides) sont construites de manière à porter en leur centre la mutation choisie, ce qui permet une hybridation optimale avec les 2 brins de l'ADN parent,
- iv) une ADN polymérase permettant de reproduire en grande quantité l'ADN parent avec la mutation désirée lors de la PCR. Cette enzyme se doit d'être de haute fidélité de manière à éviter l'introduction de mutations non-désirées.



**Schéma 39** : Stratégie de mutagenèse dirigée.

Dans le cas de nos 2 enzymes, la position des résidus catalytiques a été déterminée en alignant leur séquence peptidique respective avec des enzymes présentant une identité de séquence proche. Le design des amorces a pu dans un second temps être réalisé. Ces 2 étapes seront cependant décrites plus en détails dans chacune des parties consacrées à ces enzymes. Les mutations de DtGly et DtMan ont été réalisées à partir du kit commercial Quikchange II XL (Agilent) en suivant le protocole mis à disposition par le fournisseur. L'ADN parent de chaque enzyme a dans un premier temps été produit en grande quantité ( $\approx 50\text{-}80$  ng/ $\mu\text{l}$  d'ADN extrait et purifié) *via* le système de clonage *E. coli* DH5 $\alpha$ . Au terme de la réaction de PCR conduite en présence de la polymérase *Pfu ultra*, l'ADN muté amplifié ainsi que l'ADN parent sont présents dans le micro-tube de PCR. Une enzyme de restriction, *DpnI*, est par la suite utilisée comme agent discriminant de l'ADN fils et parent (Schéma 40). En effet tout ADN transformé chez *E. coli* est susceptible d'être méthylé, *DpnI* est une endonucléase spécifique d'un site de restriction méthylé et permet par conséquent de digérer spécifiquement l'ADN parent. Enfin l'ADN muté extrait et purifié est transformé dans les cellules ultra-compétentes XL10-Gold. Ces bactéries, développées par Agilent, assurent de meilleurs rendements de transformation que les bactéries compétentes pouvant être produites au laboratoire.



**Schéma 40** : Digestion sélective du plasmide parent par l'enzyme de restriction *DpnI* et transformation dans le vecteur de clonage ultra compétant *E. coli* XL10-Gold.

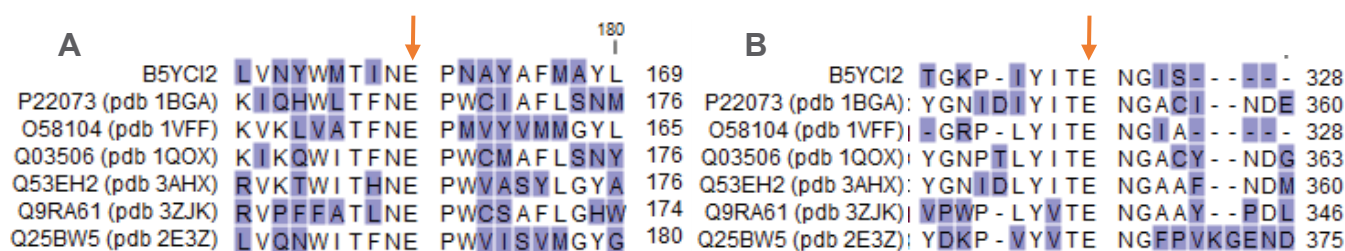


## II. Production des mutants thioglycoligases de DtGly : Application à la synthèse de thioglycosides

### 1. DtGly E159A et E159Q

#### 1.1. Conception et surexpression des mutants

La génération de mutants thioglycoligases passe par la mutation du résidu acide/base par un acide aminé catalytiquement inactif. Afin d'y parvenir il est nécessaire d'identifier les sites catalytiques de l'enzyme. En l'absence de cristaux, l'accès à la structure de DtGly est impossible, ainsi pour déterminer la position des résidus acide/base et nucléophile la séquence protéique de l'enzyme a été alignée avec les séquences de protéines proches en utilisant l'outil d'alignement Clustal Omega mis à disposition par l'Institut Européen de Bioinformatique (EBI, <http://www.ebi.ac.uk/Tools/msa/clustalo/>). A partir du résultat de l'alignement, présenté dans son intégralité en annexe (Annexe 05, p. 261), différentes régions de haute conservation en acides aminés ont pu être mises en évidence. L'étude des structures de protéines déjà décrites dans la littérature a permis d'identifier les zones correspondantes aux résidus catalytiques des enzymes (Figure 53). Ainsi le Glutamate E159 a pu être identifié comme le résidu acide/base potentiel de DtGly tandis que le Glutamate E325 jouerait le rôle du nucléophile.

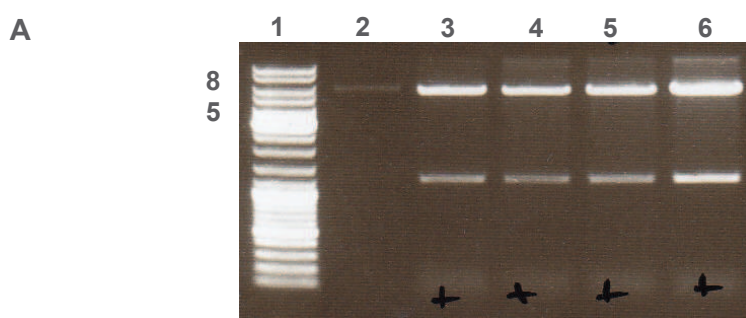


**Figure 53** : Alignement des séquences peptidiques de DtGly (Uniprot B5YCI2) et de glycosidases de GH1 de structure résolues (PDB) autour des résidus acide/base (A) et nucléophile (B). Les protéines sont identifiées selon leur code Uniprot ainsi que leur code PDB indiqué entre parenthèses. Les résidus non-homologues sont sur fond bleu, les numéros correspondent à la position du résidu terminal de chaque séquence.

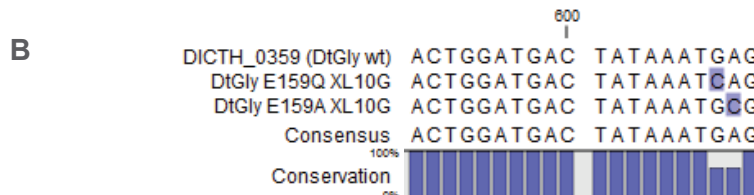
Dans le cas de DtGly, la production de mutants thioligases s'est limitée aux deux formes les plus rencontrées parmi les thioligases déjà décrites : les mutants Glutamine et Alanine. Ainsi ces 2 mutants ont été produits en remplaçant le résidu Glutamate (E) 159 par une Glutamine (Q) et une Alanine (A). Les amorces nécessaires à la génération de chaque mutant ont été construites grâce à l'outil de design d'amorces proposé par Agilent ([www.genomics.agilent.com](http://www.genomics.agilent.com)) en remplaçant le codon **GAG** (position 475 - 477 sur la séquence d'ADN de DtGly) par le codon **CAG** pour le mutant E159Q et **GCG** pour le variant E159A (cf. Partie expérimentale, p.199). Ainsi les plasmides mutés correspondant à DtGly E159Q et E159A ont pu être obtenus par mutagenèse dirigée à partir des amorces appropriées, produites par l'entreprise Eurofins (<http://www.eurofinsgenomics.eu/>), et transformés chez *E. coli* XL-10 Gold. Les bactéries transformées ont été déposées sur boîte de LB-agar contenant l'antibiotique Kan, condition discriminante des bactéries contenant le plasmide muté. Pour



chaque mutant, 2 colonies ont été isolées, mises en culture dans un milieu nutritif LB/kan et traitées pour extraire l'ADN recombinant. L'analyse sur gel d'agarose du produit de transformation ne permet pas ici de confirmer la présence de la mutation unique attendue. En effet aucune incorporation ou délétion d'acides aminés n'a été réalisée, la taille du plasmide muté est donc identique à celle du plasmide parent (**A**, Figure 54). Seul le résultat du séquençage des 4 clones sélectionnés et l'alignement des séquences d'ADN fils et parent (disponible en intégralité dans l'annexe 06, p.262) permet de confirmer la modification du codon GAG par le motif CAG du plasmide DtGly E159Q et par le motif GCG pour le plasmide DtGly E159A (**B**, Figure 54).

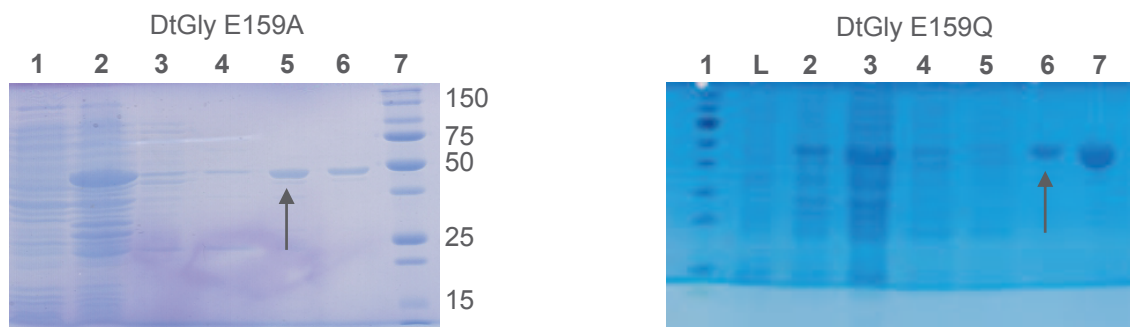


(1) Marqueur de taille (kpb) ;(2) contrôle, plasmide parent DtGly wt ;  
(3)&(4) clones positifs (+), plasmide muté DtGly E159A ;(5)&(6) clones positifs (+), plasmide muté DtGly E159Q.



**Figure 54** : Analyse sur gel d'agarose (**A**) et alignement des séquences (**B**) des produits de transformation DtGly wt, E159Q et E159A dans *E. coli* XL10 Gold. Les nucléotides non homologues sont indiqués sur fond bleu.

Les mutants E159A et E159Q ont été exprimés chez *E. coli* Rosetta-(DE3) en suivant le même protocole que celui suivi pour DtGly wt (cf. Chapitre 1, p.93). Au terme des étapes de lyse, clarification thermique et purification par chromatographie d'affinité, les deux protéines sont obtenues sous forme pure avec une taille de 53 kDa (Figure 55). A partir d'un litre de culture bactérienne, 2 mg du mutant DtGly E159A ont été obtenus contre 3 mg de DtGly E159Q.



**Figure 55 :** Analyse sur gel SDS-PAGE des mutants DtGly E159Q et E159A. (1) T<sub>0</sub> induction IPTG ; (2) T20h induction IPTG ; (3) surnageant de lyse ; (4) surnageant traitement thermique 70 °C ; (5) Purification chromatographie d'affinité ; (6) Purification chromatographie d'exclusion stérique (réalisée pour les essais de cristallographie) ; (7) Marqueur de taille (kDa). Les concentrations protéiques obtenues après purification sur colonne d'affinité sont indiquées entre parenthèses ; (L) Lysat cellulaire.

## 1.2. Caractérisation

La substitution du résidu catalytique E159 de DtGly par un acide aminé qui ne peut pas assurer l'activation basique d'une molécule d'eau entrant dans site actif a pour conséquence la réduction voire la perte de l'activité d'hydrolyse de l'enzyme ainsi modifiée. Afin de confirmer cette propriété, les paramètres cinétiques des deux variants E159A et E159Q ont été déterminés en condition d'hydrolyse à pH 6 et 37 °C. La littérature préconise l'utilisation de donneurs présentant des groupes partants électrofuges, tels que les dinitrophényles, cependant notre choix s'est porté sur le pNPGlc de manière à pouvoir comparer les valeurs obtenues avec celles du type sauvage DtGly *wt* (Tableau 13)

DtGly	$K_M$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$ (mM <sup>-1</sup> .s <sup>-1</sup> )	$\frac{\left[\frac{k_{cat}}{K_M}\right]_{wt}}{\left[\frac{k_{cat}}{K_M}\right]_{mutant}}$
<i>wt</i>	0,5 ± 0,04	31 ± 0,7	66	-
E159A	0,70 ± 0,08	0,9 ± 0,2	1,16	57
E159Q	0,20 ± 0,02	0,20 ± 0,00	1,0	66

**Tableau 13 :** Paramètres cinétiques d'hydrolyse de DtGly type sauvage et mutée.

A la lecture du tableau comparatif des constantes cinétiques de DtGly *wt*, E159A et E159Q plusieurs points sont à mettre en exergue :

- Pour les deux mutants générés l'activité d'hydrolyse a été considérablement réduite comme en attestent les valeurs du ratio  $[k_{cat}/K_M]_{wt} / [k_{cat}/K_M]_{mutant}$  obtenues. L'activité d'hydrolyse des mutants est de 60 à 70 fois plus faible que l'enzyme en type sauvage. Cette première constatation confirme donc l'implication du résidu E159 dans l'activité catalytique de DtGly, en résulte bien une perte de l'activité d'hydrolyse par mutation de cet acide aminé.

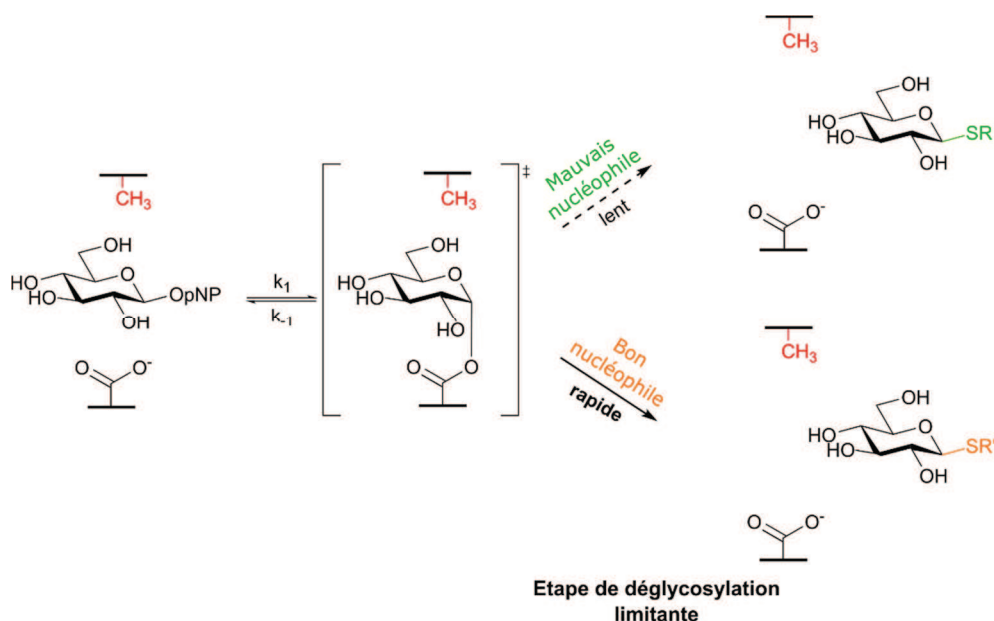
- Le variant DtGly E159Q se démarque par un  $K_M$  plus faible que celui de DtGly *wt*. Pour comprendre ce résultat il est nécessaire de considérer l'équation définissant le  $K_M$  comme  $K_M = \frac{k_{-1}+k_2}{k_1}$ . Ainsi, comme il a été décrit précédemment, en absence d'activation basique l'étape de déglycosylation caractérisée par la constante  $k_2$  devient l'étape limitante dans le mécanisme à double déplacement. Avec  $k_2 \ll k_1$  la valeur de  $K_M$  diminue et cette diminution est d'autant plus importante que la formation de l'intermédiaire glycosyl-enzyme est rapide.
- Dans le cas de DtGly E159A malgré la perte de l'activité d'hydrolyse nette le mutant présente une valeur de  $K_M$  (0,70 mM) supérieure à l'enzyme en type sauvage (0,50 mM). La formation de l'intermédiaire glycosyl-enzyme semble donc moins favorisée en présence du résidu Alanine que du résidu Glutamine dans le site actif.

Ainsi grâce à cette première étude, l'implication du résidu E159 dans l'activité catalytique de DtGly a pu être confirmée puisque sa mutation par la Glutamine ou l'Alanine permet de réduire considérablement l'activité hydrolytique de l'enzyme. D'autre part, les mutants E159A et E159Q, présentant des activités d'hydrolyse similaires, sont tous les deux de potentiel bons candidats thioglycoligase. En revanche la production du mutant E159Q semble être plus efficace chez *E. coli* que le mutant E159A (2,5 mg obtenus par litre de culture bactérienne contre 1,5 mg/l dans le cas de E159A). DtGly E159Q a donc été retenu comme candidat potentiel pour la synthèse de S-glycoconjugués.

## 2. Synthèse de S-glycoconjugués catalysée par DtGly E159Q

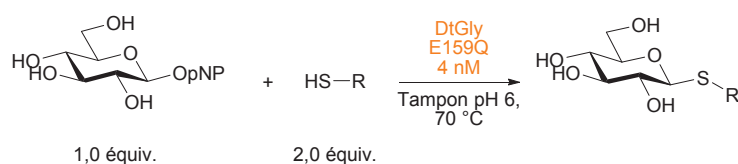
### 2.1. Etude préliminaire

Si la perte d'activité hydrolytique par mutation du résidu acide/base de DtGly en Glutamine a pu être démontrée, il s'agit à présent de déterminer la capacité de ce mutant à synthétiser des liaisons de type  $\beta$ -1-thioglycosidiques. Dans cette optique une première étude préliminaire a été réalisée en présence de différents types d'accepteurs présentant une fonction thiol et de pKa variés. L'objet de cette étude consiste à placer l'enzyme en présence du donneur de glycoside, le pNPGlc, et de déterminer le turnover de l'enzyme en présence des différents accepteurs sélectionnés. D'un point de vue mécanistique lors de la mutation d'une glycosidase en thioligase la cinétique de la première étape (formation de l'intermédiaire covalent) n'est pas affectée, comme en atteste les valeurs de  $K_M$  du même ordre de grandeur entre DtGly *wt* et ses 2 mutants (i.e.,  $\approx 0,5$  mM). L'activité enzymatique des thioligases est contrôlée par l'étape limitante de déglycosylation et peut par conséquent être modulée en fonction de la nucléophilie de l'accepteur (Schéma 41). Ainsi plus l'intermédiaire covalent sera clivé rapidement, et donc extrait du site actif, plus vite sera formé un nouvel intermédiaire covalent. En présence d'un nucléophile fort, la consommation du donneur de glycoside étant plus rapide, une nette augmentation du  $k_{cat}$  devrait être observée. A l'inverse, en présence d'un nucléophile faible, l'intermédiaire covalent formé dans le site actif sera clivé très lentement. Le turn-over de l'enzyme sera par conséquent du même ordre de grandeur que celui de l'enzyme mutée en activité d'hydrolyse.



**Schéma 41** : Activité thioligase modulée par la nucléophilie de l'accepteur de glycoside.

Les essais ont été réalisés à pH 6 et 70 °C en présence d'1,0 équivalent de donneur (soit 1 mM de pNPGlc) pour 2,0 équivalents d'accepteur (Schéma 42). Afin d'assurer la solubilité des motifs thiol, 5 % v/v de MeOH ont été introduits dans le milieu réactionnel. La température a été fixée à 70 °C dans le but de limiter le temps de réaction et donc de déterminer de manière rapide si l'enzyme présente une activité en thioglycosylation. A cette température DtGly E159Q peut réagir très rapidement en fonction du type d'accepteur introduit en solution. Afin de pouvoir suivre l'avancée de la réaction, l'enzyme a été introduite en très faible quantité 4 nM. D'autre part, l'activité thioligase de DtGly E1E9Q peut également être déterminée qualitativement par simple suivi visuel. L'échantillon contenant un bon nucléophile sera caractérisé par une couleur jaune vive, caractéristique du pNP libéré rapidement en solution, tandis que l'échantillon contenant un mauvais nucléophile restera incolore ou prendra une faible coloration jaune.



**Schéma 42** : Screening de l'activité thioglycoligase de DtGly E159Q.

Au terme de l'expérience, une nette différence d'activité de l'enzyme peut être observée en fonction du motif de l'accepteur introduit dans la réaction (tableau 14). A pH 6 Les thiols aliphatiques, caractérisés par les plus hautes valeurs de pKa, ne peuvent être sous la forme de thiolate qui est l'espèce nucléophile. En conséquence ils ne permettent pas d'accélérer l'étape limitante de déglycosylation ce qui est reflété par des valeurs de  $k_{cat}$  faibles, équivalentes à l'activité d'hydrolyse de DtGly E159Q à 70 °C. A l'inverse les thiols aromatiques, qui présentent de manière générale un proton acide (faible valeur de pKa), sont de très bons

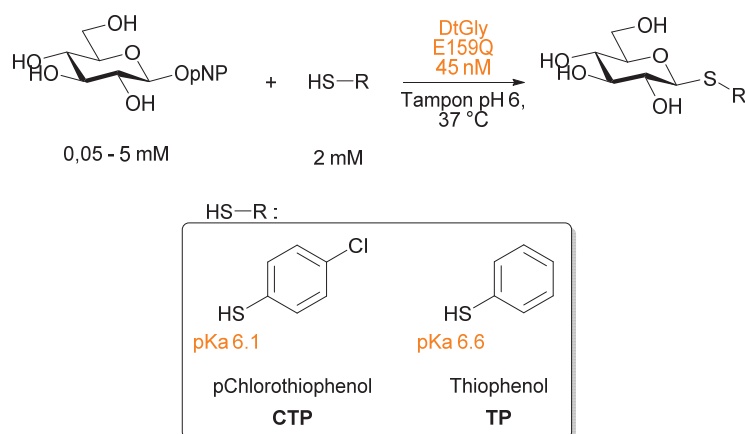
accepteurs puisqu'ils favorisent le déplacement de l'équilibre réactionnel de manière significative dans le sens de la formation du produit de thioglycosylation. Dans le cas du *para*-Chlorothiophénol le turnover de DtGly E159Q a été multiplié par 10 par rapport à celui de l'hydrolyse.

Accepteur	pKa	$k_{cat}$ (s <sup>-1</sup> )
H <sub>2</sub> O	7	6,85
pChlorothiophénol	6,1	60
Thiophénol	6,6	17
nBuSH	10,7	6,9
4-Mercaptoglycérol	9,6	8,1

**Tableau 14 :** Etude préliminaire de l'activité du mutant E159Q (4 nM) en présence de différents accepteurs à 70 °C.

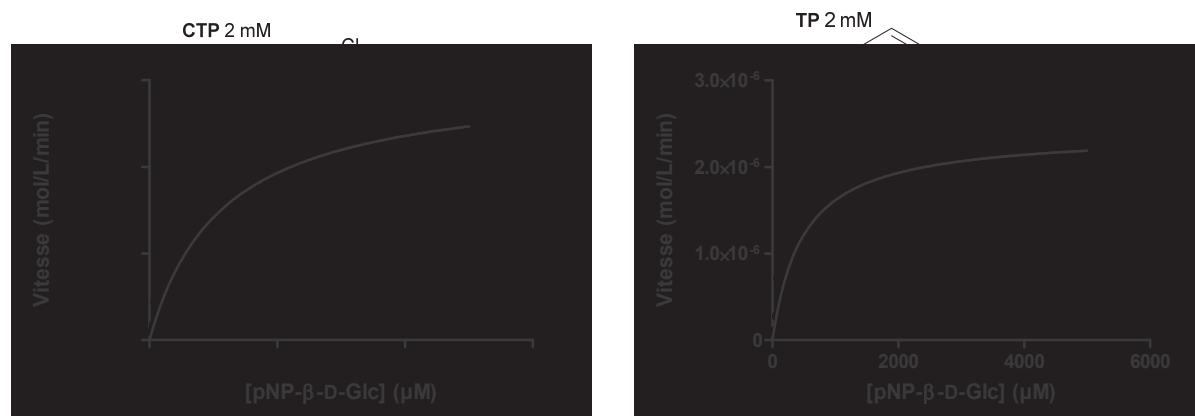
Ces résultats sont en accord avec la littérature puisqu'en 2011 Almendros *et coll.* ont utilisé le potentiel de thioglycoligase de l'arabinofuranosidase de *C. thermocellum* Araf51 E173A sur une série d'accepteurs thiols et ont présenté d'excellents rendements de réaction dans le cas de thiols aromatiques contre une totale inactivité en présence de thiols aliphatiques.(Almendros, **2011**) Ainsi l'activité thioglycoligase de DtGly E159Q a été validée mais n'est uniquement favorisée qu'en présence d'accepteurs de type thioarylique.

Dans un second temps les constantes cinétiques de l'enzyme vis-à-vis du donneur pNPGlc ont été déterminées en présence des 2 accepteurs positifs au test préliminaire, le pChlorothiophénol (CTP) et le Thiophénol (TP) (Schéma 43). Les essais ont été réalisés avec DtGly E159Q 45 nM dans un tampon acide citrique/phosphate 20 mM pH 6 et 37 °C en faisant uniquement varier la concentration du donneur pour une concentration fixe de l'accepteur. L'activité enzymatique a été déterminée par suivi du relargage du pNP à  $\lambda = 405$  nm.



**Schéma 43** : Suivi cinétique de l'activité d'hydrolyse de DtGly E159Q en présence de CTP et TP comme accepteur de glycoside.

En présence du CTP et du TP les constantes de Michaelis-Menten (Figure 56) calculées sont en accord avec celles déterminées précédemment. En effet, le turn-over de l'enzyme est très nettement amélioré en présence des deux accepteurs. L'enzyme catalyse la formation des thioglycosides correspondants 22 fois plus vite qu'elle ne forme le produit d'hydrolyse (Tableau 15). L'étape de déglycosylation n'étant plus l'étape limitante,  $k_2 \gg k_1$ , cela a pour conséquence une augmentation du  $K_M$  de l'enzyme vis-à-vis de son substrat comme en attestent les valeurs mesurées (1,2 et 0,5 mM en présence de CTP et TP respectivement).



**Figure 56** : Courbes de Michaelis-Menten de DtGly E159Q en présence d'une concentration fixe de CTP ou TP.

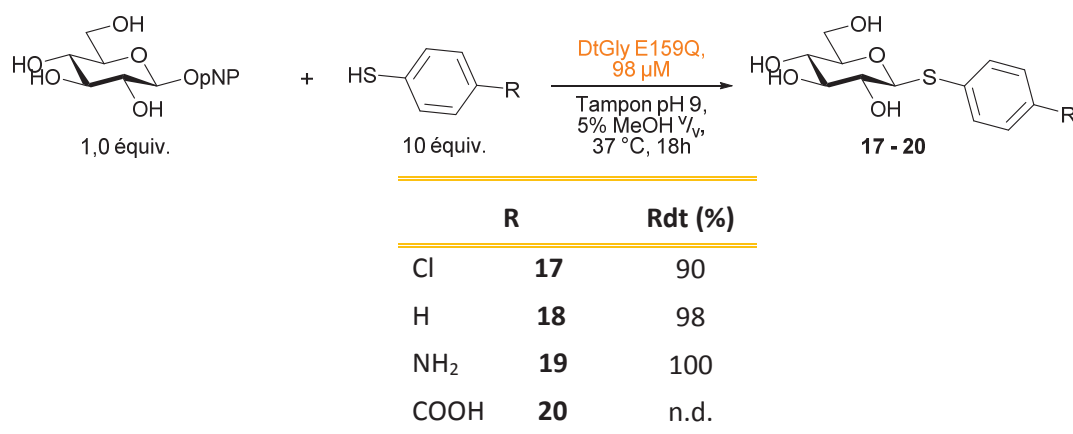
Accepteur	$K_M$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $mM^{-1}.s^{-1}$ )
H <sub>2</sub> O	$0,2 \pm 0,02$	$0,2 \pm 0,002$	1,0
CTP	$1,2 \pm 0,08$	$26 \pm 0,6$	22
TP	$0,5 \pm 0,07$	$10 \pm 0,4$	21

**Tableau 15** : Comparaison des constantes cinétiques de DtGly E159Q (45 nM) en activité d'hydrolyse et de thioglycosylation.

D'autre part en considérant ces résultats il est intéressant de noter que le pNPOH reste un bon groupe partant (malgré son pKa élevé 7,65) puisque les valeurs de  $k_{cat}$  du mutant en présence du CTP ou TP sont comparables au  $k_{cat}$  de l'enzyme en type sauvage (i.e., DtGly wt  $k_{cat} = 31 \text{ s}^{-1}$ ,  $T = 37 \text{ °C}$ ). Le pNPGlc a donc été choisi comme donneur de glycoside lors de la synthèse de dérivés thioglycosylés par DtGly E159Q.

## 2.2. Synthèse de thioglycoconjugués, relation structure/activité

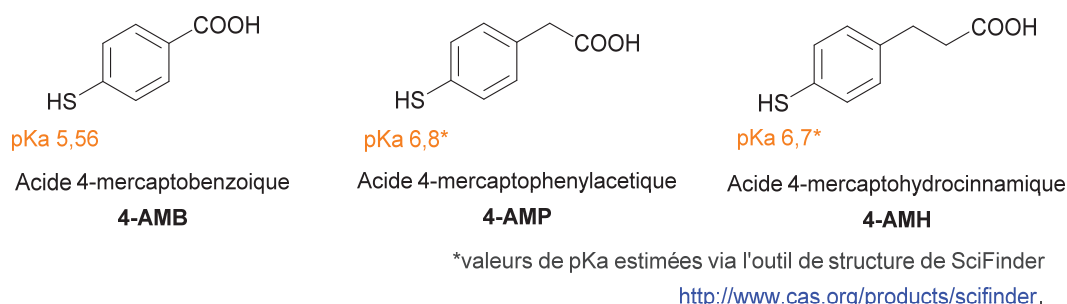
Avant d'aborder la thioglycosylation enzymatique d'acides aminés (AA) par DtGly E159Q, la relation structure/activité de l'enzyme vis-à-vis d'un panel d'accepteurs a été étudiée. Pour mimer les groupements fonctionnels des acides aminés, la 4-Mercaptoaniline a été choisie pour s'assurer de la compatibilité de l'enzyme avec la fonction amine et le 4-AMB (Acide 4-mercaptobenzoïque) pour représenter la fonction acide carboxylique. Les réactions de thioglycosylation ont été réalisées dans un tampon Tris 20 mM pH 9 et 37 °C, pour limiter la dégradation thermique des différents réactifs, et en présence d'1,0 équivalent de donneur pour 10,0 équivalents d'accepteur (Schéma 44).



**Schéma 44** : Réaction de thioglycosylation catalysée par DtGly E159Q.

Les excellents rendements de réaction obtenus à 37 °C pour les composés **17**, **18** et **19** confirment la capacité de DtGly E159Q à supporter la réaction de transfert du motif sucre sur la fonction thiol de l'accepteur. Par ailleurs, l'analyse RMN des différents produits synthétisés et le calcul des constantes de couplage des protons anomériques ( $J_{1,2} = 10 \text{ Hz}$ ) confirme la configuration  $\beta$  de la liaison thioglycosidique créée et correspondent aux données décrites dans la littérature. (Samoshin, **2014**) En revanche aucun produit de réaction n'a pu être détecté dans le cas du 4-AMB malgré son large excès devant le donneur de glycoside et les 18h d'incubation à 37 °C. Afin de vérifier si la proximité de la fonction acide carboxylique avec le cycle aromatique n'avait pas d'incidence sur la réaction, 2 autres accepteurs ont été testés en réaction de thioglycosylation l'acide 4-mercatophénylacétique et l'acide 4-mercaptohydrocinnamique (Figure 57).

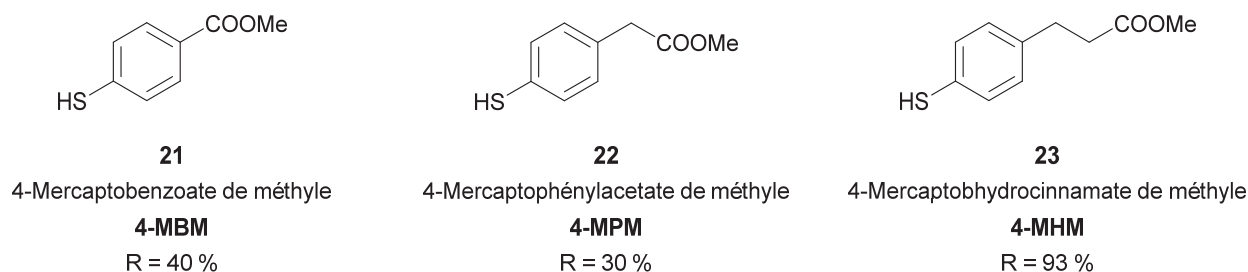




**Figure 57** : Accepteurs de glycosides présentant une fonction acide carboxylique.

Malheureusement dans les deux cas aucun produit de réaction n'a pu être détecté, la position de la fonction acide carboxylique par rapport au cycle aromatique ne semble pas être le paramètre influençant l'issue de la réaction. Le manque de réactivité présenté ici n'est sûrement pas dû à une mauvaise nucléophilie des formes thiolates de ces 3 molécules, leur pKa respectifs étant du même ordre de grandeur que ceux du CTP et TP (ie, pKa  $\approx$  6,5). Une hypothèse peut être avancée concernant une potentielle répulsion de charge entre le site actif de l'enzyme et l'accepteur. En effet dans chaque cas, la fonction acide carboxylique (pKa  $\approx$  4) est sous la forme carboxylate. Si l'entrée du site actif de l'enzyme est composée de résidus polaires chargés négativement des répulsions d'ordre électrostatique peuvent avoir lieu avec la molécule d'accepteur approchant le site actif et limitant alors sa réaction sur l'intermédiaire glycosyl-enzyme.

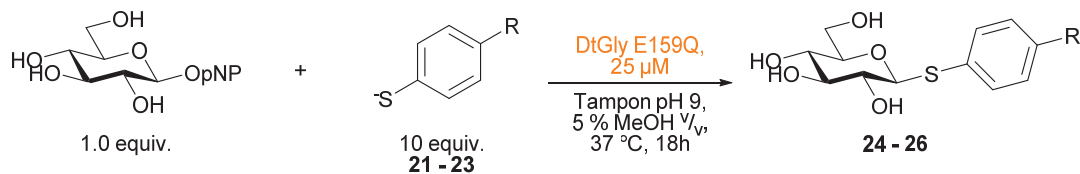
Bien que ces premiers résultats confirment le potentiel pour la synthèse de thioglycosides de DtGly E159Q, le fait que l'enzyme ne soit pas compatible avec des molécules présentant une fonction acide carboxylique constitue un obstacle pour la poursuite de notre projet. Afin de contourner ce problème et confirmer notre hypothèse nous nous sommes orientés vers l'estérification de cette fonction  $\text{-COOH}$  de manière à s'affranchir de la répulsion électrostatique supposée entre l'accepteur et le site actif de l'enzyme. Pour cette étude les réactions d'estérifications du 4-AMB, 4-AMP et 4-AMH ont été réalisées dans le méthanol en présence d'une quantité catalytique d' $\text{H}_2\text{SO}_4$  selon les conditions reportées dans la littérature et ont conduit aux 3 esters correspondants **21**, **22** et **23** (Figure 58). (Sasaki, 2011)



**Figure 58** : Estérification du 4-MBA, 4-MPA et 4-MHA.

Les réactions de thioglycosylation en présence des accepteurs nouvellement formés ont été conduites dans les mêmes conditions de réaction que celles décrites précédemment (Schéma 45).



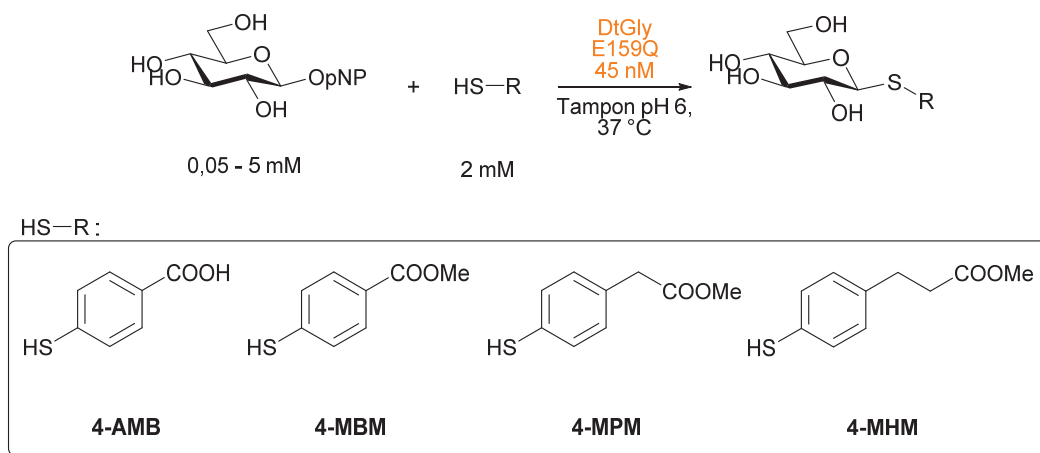


R	Produit (R %)
<b>21</b> COOMe	<b>24</b> (90)
<b>22</b> CH <sub>2</sub> COOMe	<b>25</b> (78)
<b>23</b> (CH <sub>2</sub> ) <sub>2</sub> COOMe	<b>26</b> (87)

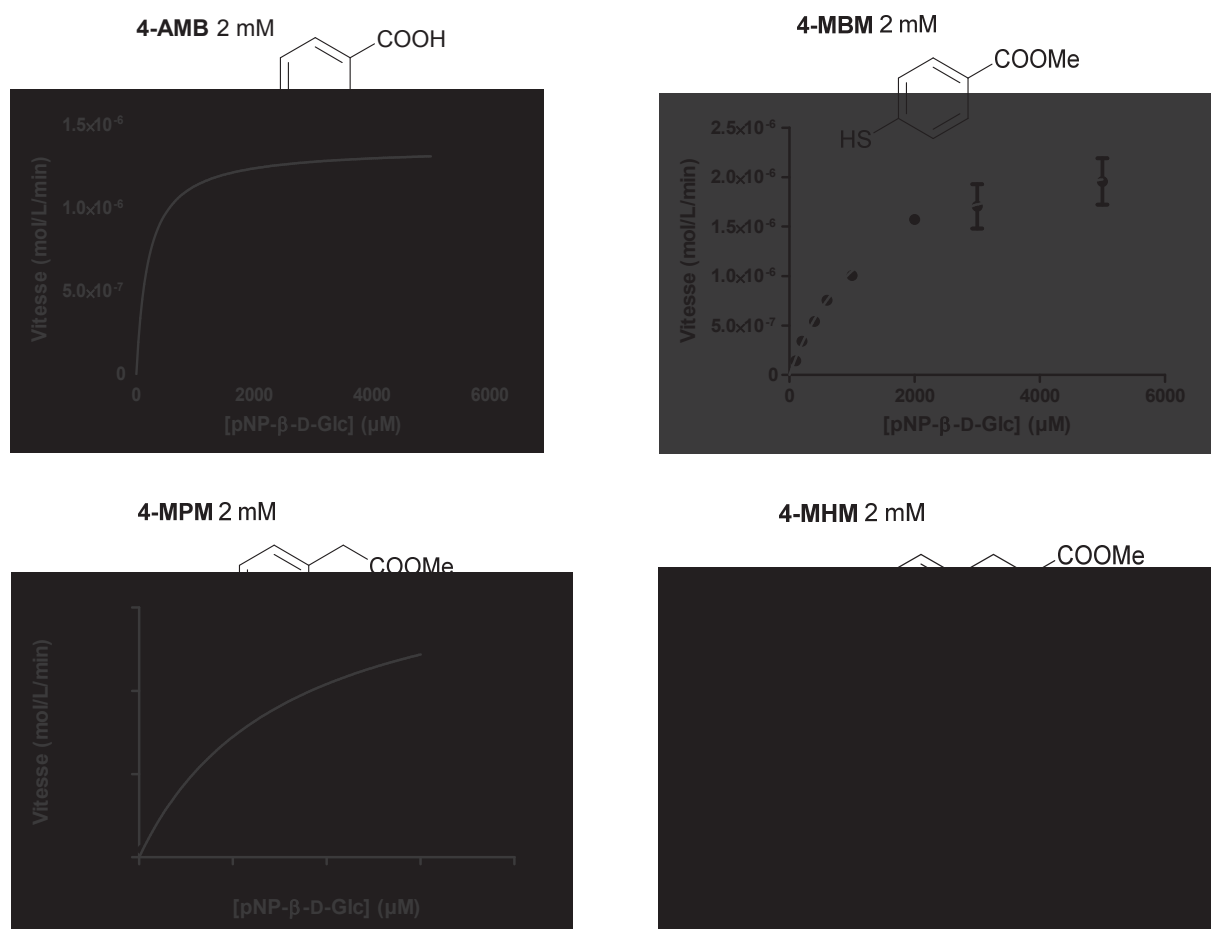
**Schéma 45** : Influence de la fonction ester de méthyle sur la réaction de thioglycosylation de DtGly E159Q.

Au vu des résultats présentés ci-dessus il apparaît clairement que la protection de la fonction acide carboxylique sous forme d'ester de méthyle permet une meilleure reconnaissance de l'accepteur par le site actif de l'enzyme. Les rendements obtenus pour les composés **24**, **25** et **26** sont comparables, il n'y a donc aucune influence vis-à-vis de l'éloignement de la fonction ester de méthyle sur la nucléophilie du groupement thiol.

Les constantes cinétiques de DtGly E159Q ont été mesurées en présence du 4-MBM, 4-MPM, 4-MHM ainsi qu'en présence de 4-AMB (contrôle). Les essais ont été réalisés en suivant le protocole classiquement utilisé, l'activité enzymatique du mutant est mesurée sur une gamme de concentration de pNPGlc pour une concentration fixe d'accepteur (Schéma 46 et Figure 59).



**Schéma 46** : Suivi cinétique d'hydrolyse du pNPGlc par DtGly E159Q en présence du 4-AMB (contrôle négatif) et des 3 accepteurs estérifiés.



**Figure 59 :** Courbe de Michaëlis-Menten de DtGly E159Q en présence d'une concentration fixe d'accepteur.

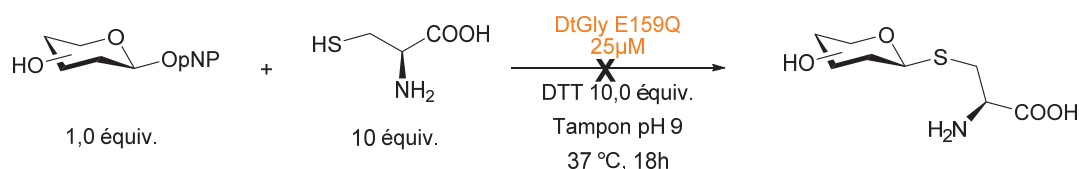
La détermination des constantes cinétiques de DtGly en présence des 3 accepteurs estérifiés confirme l'influence de la présence de charge sur l'accepteur de glycoside (Tableau 16). Le mutant E159Q présente une meilleure activité catalytique en présence du 4-MBM, 4-MPM et 4-MHM avec des valeurs de  $k_{cat}$  jusqu' à 6 fois supérieure qu'en présence du 4-AMB portant une fonction acide carboxylique non protégée.

Accepteur	$K_M$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$ (mM <sup>-1</sup> .s <sup>-1</sup> )	$\frac{k_{cat}^{COOH}}{k_{cat}^{COOMe}}$
4-AMB	0,2 ± 0,02	0,42 ± 0,01	2	-
4-MBM	1,4 ± 0,2	0,95 ± 0,05	0,7	2
4-MPM	1,6 ± 0,2	1,4 ± 0,07	0,9	3
4-MHM	2,3 ± 0,3	2,7 ± 0,2	1,2	6

**Tableau 16 :** Comparaison des constantes cinétiques d'hydrolyse du pNPGlc de DtGly E159Q en présence d'accepteurs.  $k_{cat}^{COOH}$  =  $k_{cat}$  de DtGly en présence d'un accepteur possédant une fonction acide carboxylique ;  $k_{cat}^{COOMe}$  =  $k_{cat}$  DtGly en présence d'un accepteur portant une fonction ester.

L'accepteur mieux reconnu par l'enzyme semble accéder plus facilement dans le site actif pour attaquer l'intermédiaire covalent comme en atteste l'augmentation d'au moins un facteur 7 du  $K_M$  de DtGly avec les différents esters thioaryliques ( $k_2 \gg k_1$ ) par rapport au 4-AMB. D'autre part, il est à noter que l'efficacité catalytique est dépendante de la position de la fonction ester sur la molécule d'accepteur. En effet les valeurs de  $k_{cat}$  augmentent à mesure que la fonction ester est éloignée du cycle aromatique. Dans la poche catalytique de l'enzyme des AA encombrants (tels que le Tryptophane ou la Tyrosine) peuvent être impliqués dans la stabilisation du substrat. Une gêne stérique peut donc avoir lieu entre ces AA et la molécule d'accepteur portant une fonction ester également volumineuse. Dans le cas du 4-MHM le groupement ester est le plus éloigné (chaîne alkyle à 2 carbone), la gêne stérique semble moins importante ce qui est caractérisé par la plus grande valeur de  $k_{cat}$  (2,7 s<sup>-1</sup>). Cette hypothèse d'encombrement stérique peut être appuyée en considérant les constantes cinétiques de DtGly E159Q en présence de CTP ( $K_M$  = 1,2 mM et  $k_{cat}$  = 26 s<sup>-1</sup>). Le CTP porte un atome de chlore beaucoup moins volumineux qu'un ester et peut donc facilement accéder à l'intermédiaire glycosyl-enzyme en résulte donc un turn-over de l'étape de déglycosylation jusqu'à 27 fois plus élevé que dans le cas du 4-MBM, 4-MPM et 4-MHM. Ces 3 dernières molécules sont de moins bons accepteurs de glycosides que le CTP. Néanmoins comme il a été décrit précédemment, les réactions de thioglycosylation qui ont été réalisées sur un temps d'incubation long (18h) ont permis de s'affranchir de ce facteur limitant de gêne stérique en obtenant les 3 thioglycosides attendus et ce avec des rendements quasi-quantitatifs.

Enfin un dernier essai a été réalisé en présence de Cystéine, seul acide aminé naturel portant une fonction thiol, en suivant les mêmes conditions que celles décrites précédemment. Un agent réducteur, le dithiothréitol (DTT) a été ajouté dans le milieu réactionnel pour éviter la formation de pont disulfure entre deux molécules d'accepteur (Schéma 47). Malheureusement, et de manière prévisible, aucun produit de réaction n'a été détecté au terme des 18h d'incubation. En effet la présence de la fonction acide carboxylique ainsi que la faible nucléophilie de la fonction thiol (pKa = 8,53) ne favorise pas l'attaque de la Cystéine sur l'intermédiaire covalent glycosyl-enzyme pour former le thioglycoside correspondant.



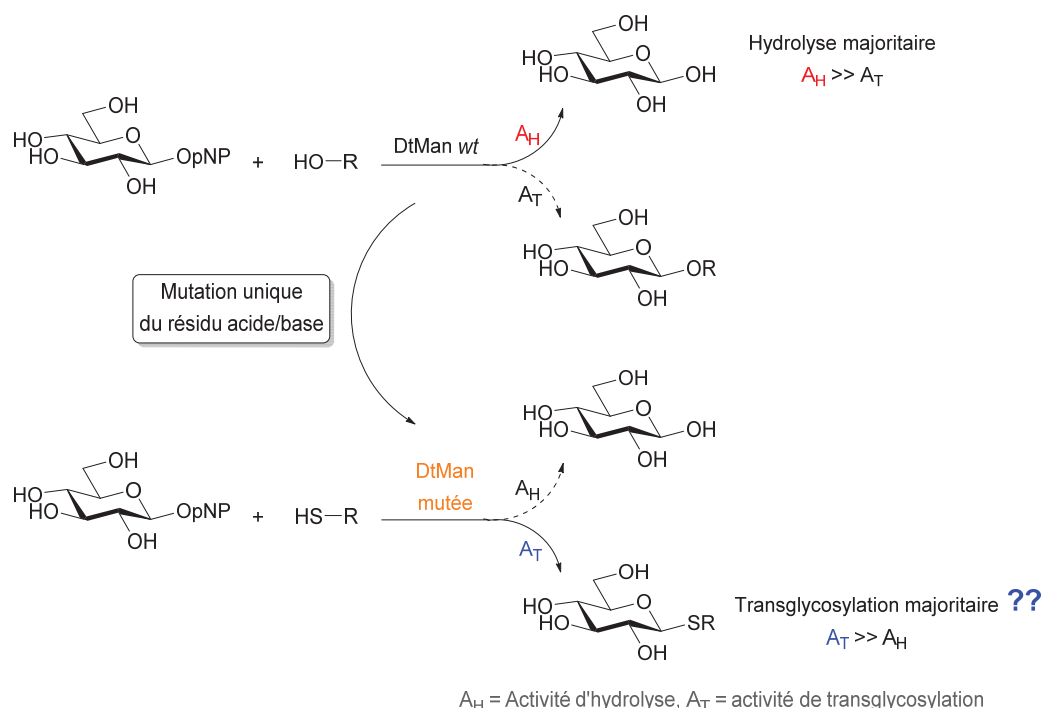
**Schéma 47** : Essai de thioglycosylation de la Cystéine catalysée par DtGly E159Q dans un tampon Tris 20 mM pH 9 et 37 °C pendant 18h.

En conclusion sur cette recherche de relation structure/activité préliminaire à la synthèse enzymatique d'acides aminés glycosylés, il apparaît que seuls les accepteurs présentant un thiophénol, fonctionnalisé ou non, soient compatibles avec l'activité de thioglycoligase de DtGly E159Q. La nucléophilie des thiols aliphatiques n'étant pas assez forte ( $pK_a > 7$ ) pour déplacer l'équilibre de la réaction vers la formation de la liaison thioglycosidique. Le premier essai de thioglycosylation enzymatique d'acide aminé réalisé sur le seul représentant naturel portant une fonction thiol, la Cystéine, s'est révélé infructueux. En d'autre terme la synthèse d'un acide aminé non-naturel de structure thioarylique devra donc être réalisée pour tenter de valider le concept défendu dans le cadre de cette thèse. Afin de limiter les répulsions d'ordre électrostatique entre le site actif de l'enzyme cet acide aminé, la fonction acide carboxylique de ce dernier devra être protégée sous la forme d'un ester.

### III. Etude de l'effet de mutations sur l'activité de synthèse de DtMan

#### 1. Etude de variants thioglycoligases de DtMan

Au cours du précédent chapitre consacré à la caractérisation de glycosidases type sauvage de *D. thermophilum*, l'étude portée sur la  $\beta$ -mannosidase DtMan a révélé une forte activité hydrolytique représentée par des valeurs élevées de  $k_{cat}/K_M$  (9 932 et 37 000 et  $\text{mM}^{-1} \cdot \text{min}^{-1}$  à 37 °C et 70 °C respectivement) bien supérieure à celles de DtRha et DtGly (cf. Chapitre 1 Tableau 7 p.87 et Tableau 10 p.96 respectivement). Son utilisation en tant que transglycosidase s'est alors avérée infructueuse puisqu'aucun produit de synthèse n'a pu être détecté laissant supposer que les propriétés intrinsèques de l'enzyme n'étaient pas compatibles avec la réaction de transglycosylation. En effet en présence d'un accepteur, autre qu'une molécule d'eau, l'issue du clivage de l'intermédiaire glycosyl-enzyme est dépendante de la sélectivité de l'enzyme vis-à-vis de l'accepteur. Si celle-ci est en faveur de l'alcool la formation du produit de transglycosylation sera favorisée, dans le cas contraire le produit d'hydrolyse sera majoritaire. Ainsi suite à ces résultats, il paraît crucial de s'interroger sur la potentielle modulation du ratio hydrolyse/transglycosylation ( $A_H/A_T$ ) de DtMan par modification de son activité hydrolytique (Schéma 48). Cette problématique commune aux glycosidases suscite un fort intérêt parmi la communauté scientifique puisqu'elle constitue un frein au développement de ces biocatalyseurs pour la synthèse d'oligosaccharides.



**Schéma 48 :** Modulation du ratio  $A_H/A_T$  de DtMan par mutagenèse dirigée.

### 1.1. Conception, et production de la banque de mutant DtMan E425X

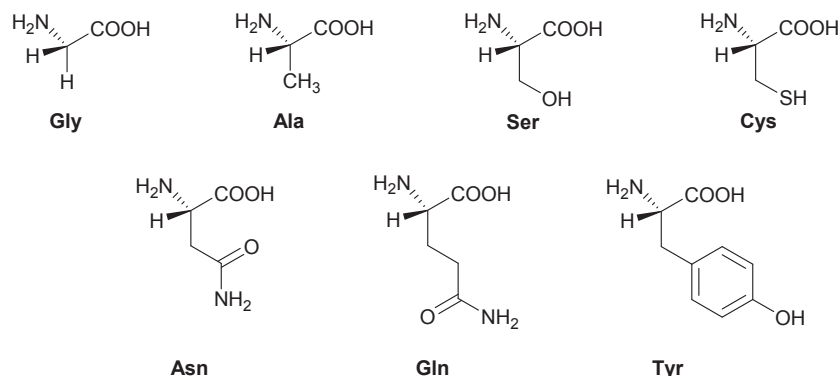
A partir de l'alignement de la séquence protéique de DtMan avec glycosidases de séquences proches, les sites E425 et E508 ont pu être identifiés comme les potentiels résidus acide/base et nucléophile respectifs de la poche catalytique (Figure 60).

A			B		
B5YAN4	NE	---	NHW	GY YAK	---
P00723 (pdb 3OB8)	NE	ACY	---	GR NHK	AMY KL
Q8KRF6 (pdb 1YQ2)	NE	SGT	---	GS NLA	AMAAW
P00722 (pdb 1DP0)	NE	SGH	---	GA NHD	ALYRW
P08236 (pdb 1BHG)	NE	PASHLESA		GY YLK	MMVIAH
Q8AAK6 (pdb 2JE8)	NE	ILEALKYW		GF EKK	---
					F
					476
					434
					526
					569
					539
					556
					548
					573

**Figure 60 :** Alignement des séquences peptidiques de DtMan (Uniprot B5YAN4) et de glycosidases de GH2 de structure résolues (PDB) autour des résidus acide/base (A) et nucléophile (B). Les protéines sont identifiées selon leur code Uniprot et leur code PDB indiqué entre parenthèses. Les résidus non-homologues sont indiqués sur fond bleu, les numéros correspondent à la position du résidu terminal de chaque séquence. L'intégralité de l'alignement de séquence est disponible dans en annexe (Annexe 07, p. 264).

Contrairement à DtGly dont l'application à la synthèse de thioglycosides s'est limitée à la génération des 2 mutants Alanine et Glutamine, correspondant aux thioglycoligases les plus actives décrites dans la littérature, l'objet de recherche porté sur DtMan, plus fondamental, a été orienté sur la conception d'une banque de 7 variants thioglycoligases. De manière à

permettre d'évaluer l'impact de chaque mutation sur le ratio hydrolyse/synthèse de l'enzyme, les acides aminés ont été sélectionnés avec une large variété de fonction mais également en tenant compte des phénomènes de gêne stérique ou de répulsion électrostatique pouvant interférer avec la bonne reconnaissance du substrat dans le site actif (Figure 61).



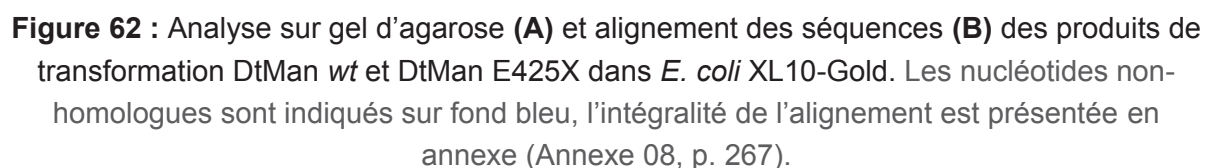
**Figure 61** : Acides aminés (série L) sélectionnés pour la génération des variants thioglycoligase de DtMan.

Les amorces nécessaires à la génération de chaque mutant, ont été construites grâce à l'outil de design d'amorces proposé par Agilent ([www.genomics.agilent.com](http://www.genomics.agilent.com)), en remplaçant le codon **GAG** (position sur la séquence d'ADN de DtGly) par les codons associés aux différents AA sélectionnés pour remplacer le résidu E425 (Tableau 17).

Résidu	<b>E425 (wt)</b>	E425A	E425Q	E425N	E425C	E425S	E425Y	E425G
Codon associé	<b>GAG</b>	GCG	CAG	AAC	TGC	TCA	TAT	GGA

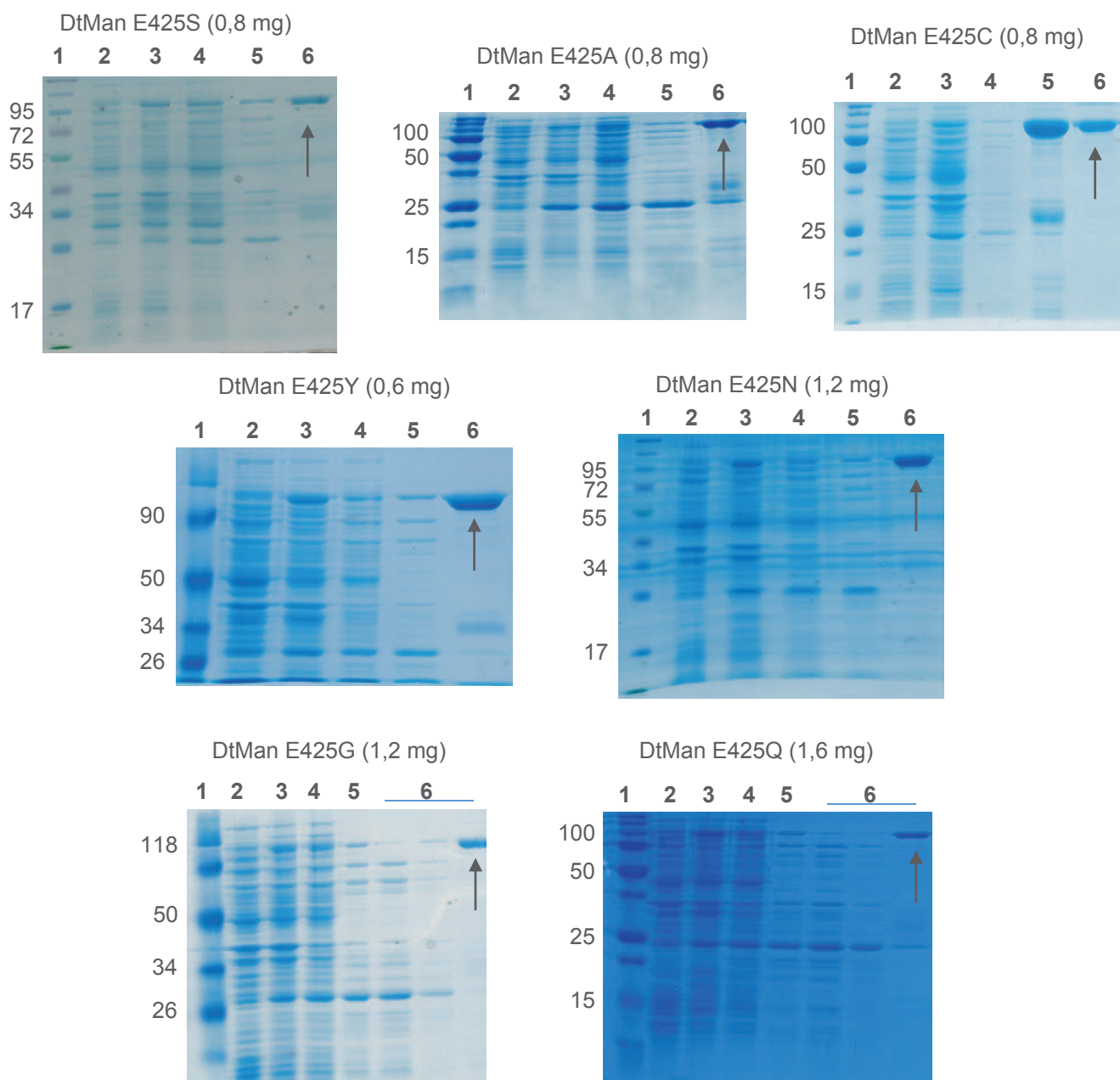
**Tableau 17** : Modifications du codon associé au résidu E159 de la séquence protéique de DtMan pour la génération de chaque mutant DtMan E425X. L'intégralité des amorces construites sont décrites dans la partie expérimentale, p.200.

De manière similaire à DtGly, les mutants DtMan E425X ont été obtenus par mutagenèse dirigée. Les plasmides mutés de chaque variant ont été obtenus par PCR à partir des amorces appropriées (sens et anti-sens pour encadrer les 2 brins d'ADN) et transformés dans les bactéries ultra compétentes *E. coli* XL10-Gold. Après croissance bactérienne sur boîte de LB/agar, Kan, plusieurs colonies de chaque mutant ont été sélectionnées et inoculées dans un milieu nutritif LB/Kan. L'ADN plasmidique muté de chaque clone a été extrait, purifié et analysé sur gel d'agarose de manière à identifier les clones potentiellement positifs contenant le plasmide de DtMan mutée (**A**, Figure 62). Pour chaque variant, un clone a été sélectionné et séquencé (Eurofins) pour confirmer la présence de la bonne mutation. A partir du résultat de séquençage de chaque clone, l'alignement des séquences d'ADN des mutants et de DtMan wt a pu être réalisé et a confirmé la présence de l'unique mutation recherchée (**B**, Figure 62).



Les différents mutants DtMan E425X ont été exprimés chez *E. coli* Rosetta-(DE3) en suivant le même protocole que celui suivi pour DtMan *wt* (cf. Chapitre 1, p.72). Au terme des étapes de lyse, clarification thermique et purification par chromatographie d'affinité, les 7 protéines recombinantes sont obtenues sous forme pure avec une taille de 98 kDa (Figure 63).





**Figure 63 :** Analyse sur gel SDS-PAGE des mutants DtMan E425X (98kDa).

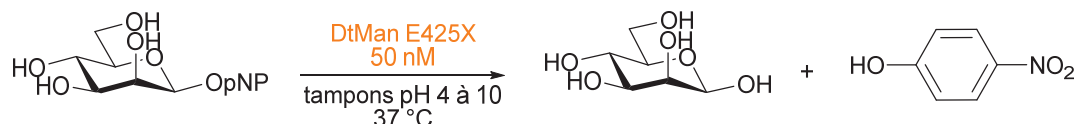
(1) Marqueur de taille (kDa) ; (2) T<sub>0</sub> induction IPTG ; (3) T20h induction IPTG ; (4) surnageant de lyse ; (5) surnageant clarification thermique 70 °C ; (6) Purification chromatographie d'affinité (bande unique ≈ 98 kDa). Les rendements de production des différents mutants pour un litre de culture bactérienne sont indiqués entre parenthèses.

## 1.2. Caractérisation

L'introduction d'une mutation ponctuelle sur la séquence d'une protéine, et notamment dans le site actif, engendre des modifications sur l'activité catalytique de cette dernière mais également sur ses propriétés physico-chimiques. Le caractère thermostable a été conservé pour l'ensemble des variants DtMan E425X, comme en atteste l'analyse sur gel présentée ci-dessus. En effet les protéines n'ont pas été dégradées lors de l'étape de clarification thermique

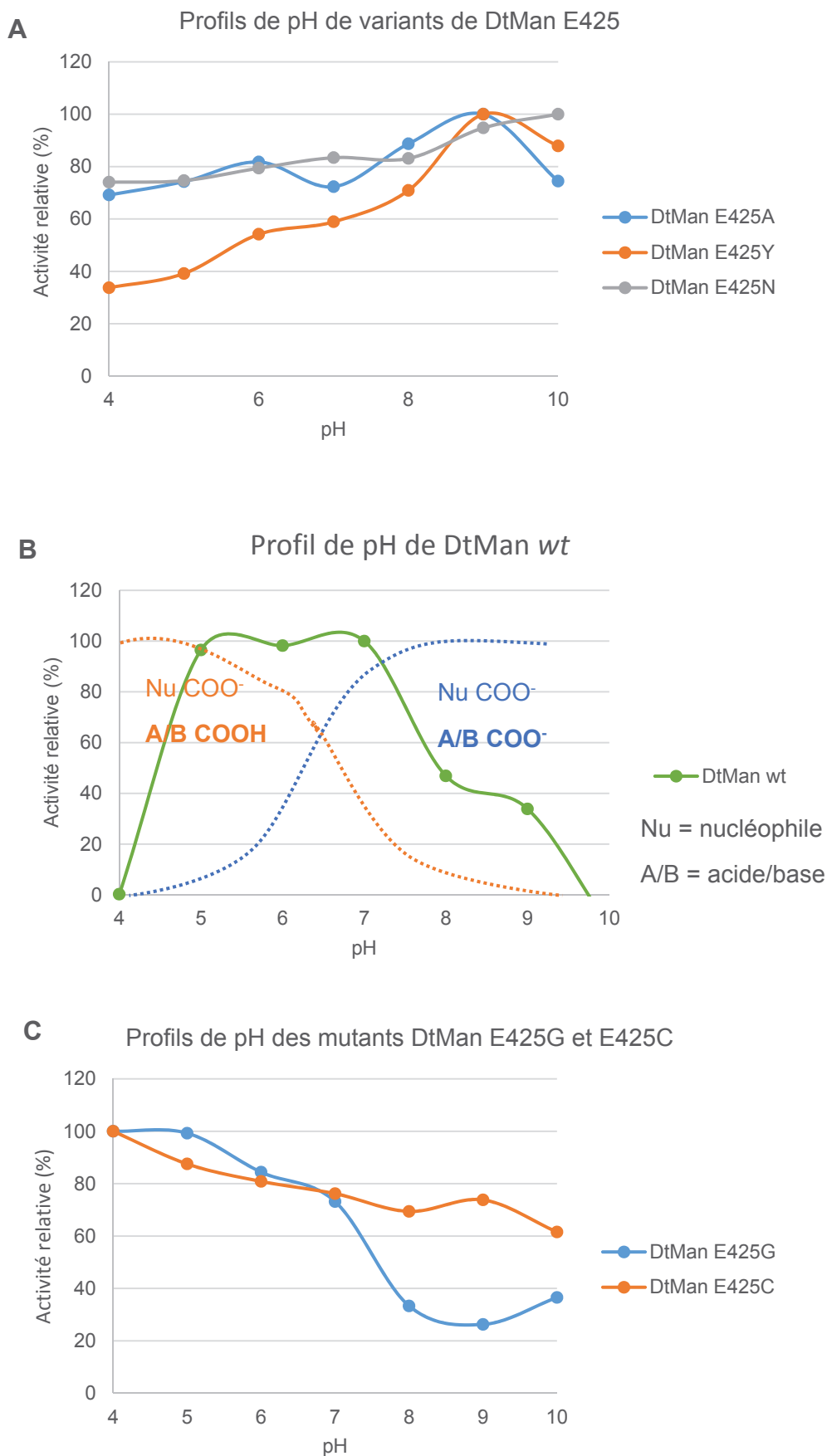


(ie, 15 min à 70 °C) puisque la bande correspondante est observée sur gel SDS-Page (Figure 63, pistes 5). En revanche la génération de mutants thioligase à de manière générale une incidence sur le pH optimal d'activité enzymatique. L'influence du type de mutation sur le pH optimal de DtMan E425X a donc été étudiée. Les essais ont été réalisés dans les mêmes conditions que DtMan wt, en suivant l'activité d'hydrolyse de l'enzyme sur une gamme de pH allant de 4 à 10 (Schéma 49).



**Schéma 49 :** Etude de l'influence du pH sur l'activité d'hydrolyse des mutants DtMan E425X.

De manière générale les profils de pH obtenus pour les différents mutants sont sous la forme d'un plateau et montrent une indépendance de l'activité enzymatique vis-à-vis du pH (**A**, Figure 64). Chez les glycosidases natives l'activité catalytique est à l'inverse très dépendante du pH puisqu'elle dépend de l'état d'ionisation des résidus nucléophile et acide/base. (Mullegger, **2005**) Lors de la première phase du mécanisme, l'activité enzymatique sera favorisée pour des valeurs acides de pH, en conservant le nucléophile sous forme déprotonée ainsi que le résidu acide/base sous forme protonée alors nécessaire pour l'assistance du départ du groupe partant. Lors de la 2<sup>ème</sup> phase du mécanisme, l'activité enzymatique est en revanche favorisée pour des valeurs basiques de pH qui assurent la présence du résidu acide/base sous forme déprotonée essentielle à l'activation d'un nucléophile entrant dans le site actif. Par conséquent le profil de pH des glycosidases natives est sous forme de cloche, comme la combinaison des 2 caractères tantôt acide, tantôt basique inhérents au mécanisme de ces enzymes (**B**, Figure 64). En remplaçant le résidu acide/base par un AA catalytiquement inactif, cette dépendance vis-à-vis du pH est levée, puisque le nucléophile est sous la forme déprotonée à partir de faible valeur de pH, l'activité enzymatique des mutants est quasi équivalente entre pH 4 et pH 9. A l'exception des mutants Glycine et Cystéine (**C**, Figure 64) le maximum d'activité des mutants est atteint à pH basique (pH 9 ou 10) ce qui peut être également en accord avec la présence de molécules d'eau majoritairement sous la forme HO<sup>-</sup> et donc capables de réagir sur l'intermédiaire covalent.



**Figure 64** : Comparaison des profils de pH de DtMan *wt* et mutée.

Enfin les constantes cinétiques ont été calculées pour le pNPMan en respectant les mêmes conditions d'hydrolyse pour les 7 mutants : tampon Imidazole.HCl 20 mM pH 7 et 70 °C. Ces paramètres ont été fixés de manière à comparer l'activité catalytique des différents mutants avec le maximum de l'activité de l'enzyme en type sauvage.

En observant les résultats de ces deux études, regroupés dans le tableau 18, il apparaît clairement que la suppression du résidu acide/base de DtMan, empêche l'activation des molécules d'eau présentes dans le site actif et par conséquent réduit l'activité enzymatique.

L'étape de déglycosylation ( $k_2$ ) ainsi ralentie provoque une chute du  $k_{cat}$  ( $k_{cat} = \frac{V_{max}}{[E]} \approx k_2$ ) des mutants qui ne sont plus alors capables de réaliser l'acte catalytique avec une même fréquence que le type sauvage ( $k_{cat} \approx 300 \text{ s}^{-1}$ ). De plus, la valeur de la constante  $k_2$  diminuant ceci a une répercussion directe sur les valeurs de  $K_M$  des mutants globalement plus faibles que celle de DtMan *wt*. En revanche parmi les 7 candidats étudiés, les variants Glycine et Cystéine font figure d'exception avec le maximum d'activité d'hydrolyse observé à pH acide et dans le cas de la mutation Glycine une valeur de  $K_M$  supérieure à celle de l'enzyme en type sauvage.

DtMan	$K_M$ (mM)	$k_{cat}$ ( $\text{s}^{-1}$ )	$k_{cat}/K_M$ ( $\text{mM}^{-1}.\text{s}^{-1}$ )	pH Opt
<b>wt</b>	<b>0,48 ± 0,03</b>	<b>297 ± 33</b>	<b>622</b>	<b>7</b>
E425A	0,090 ± 0,01	0,75 ± 0,03	8	9
E425Q	0,043 ± 0,07	2,8 ± 0,2	65	9
E425N	0,16 ± 0,02	3,4 ± 0,2	22	10
E425C	0,25 ± 0,07	8,2 ± 1,3	33	4
E425S	0,06 ± 0,07	0,67 ± 0,03	11	10
E425Y	0,2 ± 0,02	1,3 ± 0,7	6	9
E425G	0,57 ± 0,1	38,1 ± 4,1	67	4

**Tableau 18 :** Etude comparative des constantes cinétiques et  $\text{pH}_{\text{Opt}}$  des variants de DtMan E425X à pH 7 et 70 °C.

- **DtMan E425G :**

La Glycine est le plus petit acide aminé parmi les 20 retrouvés dans la nature. Par conséquence sa substitution au résidu Glutamate peut engendrer une modification de la structure du site catalytique de l'enzyme et ainsi empêcher la bonne fixation du substrat, résultant une hausse de la valeur de  $K_M$  (0.57 mM) par rapport au  $K_M$  (0.47 mM) de l'enzyme en type sauvage. Par ailleurs E425G est le mutant catalysant le plus efficacement la réaction d'hydrolyse ( $k_{cat} = 38 \text{ s}^{-1}$ ). Il est alors envisageable que la modification structurale du site catalytique de l'enzyme permette l'action d'un résidu Glu ou Asp à proximité de la mutation permettant ainsi l'activation basique des molécules d'eau et donc la formation du produit d'hydrolyse. Ceci serait en accord avec le maximum d'activité mesuré à pH 4 et non pH 9 comme les autres mutants. Une mauvaise fixation ou orientation du substrat dans le site actif

peut limiter l'action du résidu nucléophile nécessaire au départ du groupe partant. La formation de l'intermédiaire glycosyl-enzyme serait donc favorisée par l'assistance acide des protons du milieu réactionnel en accélérant le départ du groupe électrofuge (pNP). Malgré tout, seule la résolution de la structure tridimensionnelle du variant E425G permettrait d'apporter les données essentielles à la confirmation de cette hypothèse.

- **DtMan E425C :**

La Cystéine est l'équivalent de l'acide aminé Sérine, la fonction alcool étant remplacée par une fonction thiol. L'oxygène et le soufre possédant des propriétés similaires, les mutants Sérine et Cystéine présentent une activité catalytique du même ordre de grandeur mais avec des valeurs de pH<sub>Opt</sub> différentes (pH 10 et pH 4 respectivement). Cette observation dans le cas du mutant E425C fait référence au mutant Glycine pour lequel l'hypothèse de la modification structurelle de la poche catalytique avait été proposée. Le soufre est une entité très réactive, sensible à l'oxydation et capable de réaliser très facilement des ponts disulfures si un autre atome de soufre est à proximité. Sur la séquence protéique de DtMan, deux résidus Cystéines (C387 et C421) sont à proximité du résidu muté E425 et pourraient donc permettre la formation d'un pont disulfure. Pour confirmer cette hypothèse une seconde étude a été réalisée sur ce mutant en présence d'un agent réducteur, le dithiothréitol (DTT), dans les mêmes conditions d'hydrolyse (Tableau 19).

DtMan E425C	$K_M$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_M$ (mM <sup>-1</sup> .s <sup>-1</sup> )	pH Opt
Sans DTT	0,25 ± 0,07	8,16 ± 1,3	33	4
Avec DTT	0,19 ± 0,02	9,53 ± 0,3	50	9

**Tableau 19 :** Constantes cinétiques de DtMan E425C en absence et en présence d'agent réducteur.

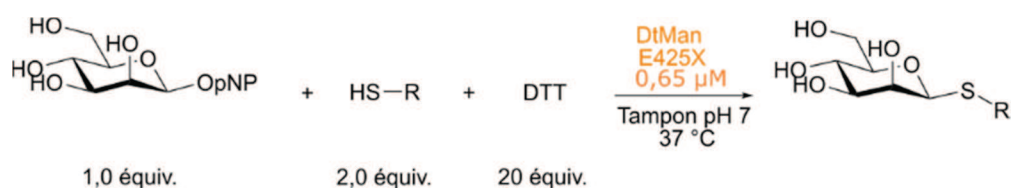
En présence de DTT, l'activité hydrolytique est du même ordre de grandeur que les autres mutants déjà caractérisés et avec un maximum d'activité atteint à pH basique également. En absence de DTT les conditions optimales d'hydrolyse sont atteintes à pH 4 et une légère augmentation du  $K_M$  est observée. Ces analyses seraient donc en accord avec la formation d'un pont disulfure au sein du site actif. En effet, la gêne stérique ainsi occasionnée empêche la bonne orientation du substrat dans le site actif ce qui implique à la fois une augmentation du  $K_M$  de l'enzyme et une diminution du pH<sub>Opt</sub> de cette dernière (la formation de l'intermédiaire glycosyl-enzyme étant assistée par les protons du milieu réactionnel). En présence d'un agent réducteur, la formation du pont disulfure devient impossible, le substrat peut s'orienter et se fixer convenablement dans le site actif de l'enzyme qui retrouve alors des constantes cinétiques comparables aux autres variants de DtMan. Afin de répondre à ces hypothèses une étude cristallographique de DtMan E425C a été réalisée en présence et en absence d'agent réducteur (voir section 2, p.143).

En activité d'hydrolyse le type de mutation du résidu acide/base n'a aucune influence puisque les constantes cinétiques mesurées pour chacun des mutants sont du même ordre de

grandeur, à l'exception du mutant Glycine. Pour réaliser une étude représentative en thioglycosylation, les essais ont été effectués à partir des mutants Alanine, Tyrosine et Asparagine présentant des activités hydrolytiques résiduelles croissantes ( $k_{\text{cat}} = 0,75 \text{ s}^{-1}$  ;  $1,3 \text{ s}^{-1}$  et  $3,4 \text{ s}^{-1}$  respectivement).

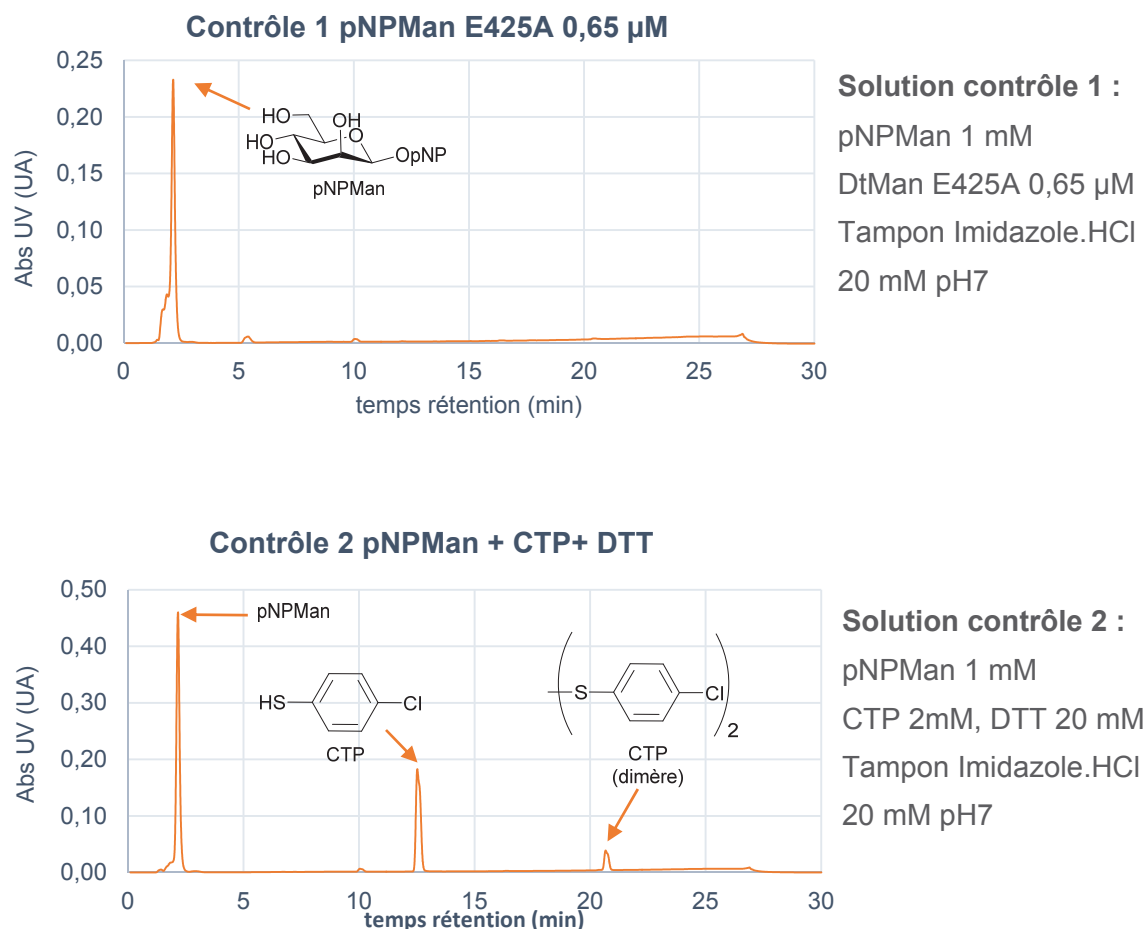
### 1.3. Essais de thioglycosylation

Les réactions de thioglycosylation ont été réalisées à partir d'1,0 équivalent de donneur (1 mM pNPMan) pour 2,0 équivalents d'accepteur (CTP) et 20 équivalents de DTT, agent réducteur empêchant la formation de pont disulfure, dans un tampon Imidazole.HCl pH 7 à 37 °C. Afin de comparer l'activité des 3 mutants (DtMan E425A, E425Y et E425S) ils ont été introduits dans les mêmes concentrations (0,65  $\mu\text{M}$ ) (Schéma 50).



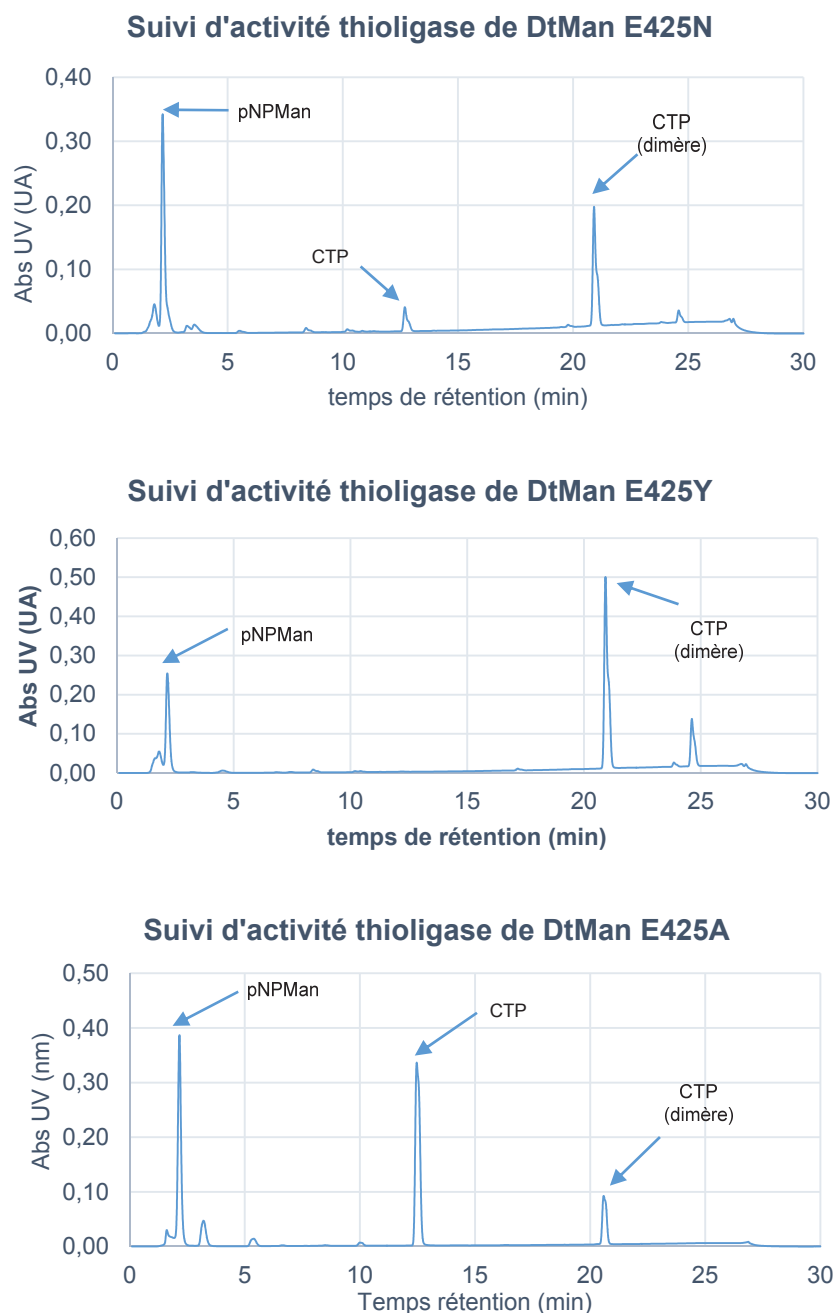
**Schéma 50** : Réaction de thioglycosylation des variants A, Y et N de DtMan E425.

La formation de produits de thioglycosylation a été dans un premier temps suivie par chromatographie sur couche mince (CCM) après 2h et 4h d'incubation à 37 °C. Malheureusement dans aucun des cas, la formation d'un nouveau composé n'a pu être observée sur CCM. L'activité thioligase des différents mutants étant potentiellement faible, les réactions ont alors été laissées à incuber pendant 18h. D'autre part les produits de départ et de thioglycosylation étant tous UV-actif, il est apparu plus judicieux d'analyser la composition de chaque milieu réactionnel par HPLC (colonne C18 ; gradient ACN/H<sub>2</sub>O). Dans un premier temps, 2 solutions contrôle ont été injectées de manière à identifier les pics correspondants aux réactifs de départ. Le chromatogramme UV de la solution contrôle **1** n'est composé que d'un pic unique à  $t_R \approx 3 \text{ min}$  correspondant au pNPMan. Le chromatogramme de la solution contrôle **2** présente trois pics pouvant être attribués au pNPMan ( $t_R \approx 3 \text{ min}$ ), au CTP sous la forme de monomère ( $t_R \approx 13 \text{ min}$ ) et au CTP sous forme de dimère ( $t_R \approx 21 \text{ min}$ ), molécule la plus apolaire donc présentant le temps de rétention le plus élevé (Figure 65).



**Figure 65 :** Chromatogrammes UV des solutions contrôle pNPMMan DtMan E425A 0,65  $\mu$ M **(A)** et pNPMMan + CTP + DTT **(B)** après 2h d'incubation à 37 °C.

Les échantillons prélevés des réactions de thioglycosylation de DtMan 425A, DtMan E425Y et E425S ont par la suite été injectés sur HPLC (Figure 66). Sur les chromatogrammes obtenus, 3 pics apparaissent de façon majoritaire avec des temps de rétention ( $t_R$ ) de 3 min, 13 min et 21 min et correspondent, comme décrit précédemment au donneur, accepteur et accepteur sous forme de disulfure respectivement. Après 18h d'incubation à 37 °C aucun produit de thioglycosylation n'a pu être détecté, la réaction de thioglycosylation n'a pas eu lieu pour les 3 mutants testés. Le même suivi a été réalisé en augmentant la concentration de l'enzyme dans le milieu d'un facteur 10 mais aucune amélioration n'a été détectée.



**Figure 66** : Chromatogrammes HPLC de suivi de thioglycosylation de variants DtMan E425X après 18h d'incubation à 37 °C.

L'absence de réactivité peut alors être due à l'accepteur qui n'est pas capable de déplacer l'équilibre de la réaction dans le sens de la formation du produit, ou aux propriétés intrinsèques de l'enzyme qui ne favorise pas l'attaque de l'accepteur sur l'intermédiaire glycosyl-enzyme. La qualité d'accepteur du CTP a été confirmée dans le cas du mutant thioligase de DtGly où l'efficacité de l'enzyme en thioglycosylation est comparable à l'activité d'hydrolyse de DtGly *wt* ( $k_{cat} = 26 \text{ s}^{-1}$  et  $k_{cat} = 66 \text{ s}^{-1}$  respectivement).

Parallèlement le potentiel de thioligase du mutant E425Q a été étudié en suivant le protocole de screening d'accepteur réalisé sur DtGly (p.123). L'objectif recherché ici est de montrer l'augmentation nette du  $k_{cat}$  de l'enzyme pour une quantité croissante de CTP. Les réactions ont été conduites à pH 7 et 70 °C avec 2, 4, 5 et 10 équivalents de CTP par rapport au pNPMAN, en présence de 9 nM de biocatalyseur. Contrairement à ce qui était attendu, les résultats, présentés dans le tableau 20, ne montrent aucune modification de l'activité enzymatique en présence de CTP quelle que soit sa concentration dans le milieu réactionnel. Le mutant Glutamine ne présente donc aucune activité de thioglycosylation.

DtMan E425Q	$k_{cat}$ ( $s^{-1}$ )
Hydrolyse	$2,8 \pm 0,2$
2 CTP	3,1
4 CTP	3, 1
5 CTP	2,7
10 CTP	2,5

**Tableau 20 :** Mesure de l'efficacité de DtMan E425Q en présence de concentrations croissantes de CTP. L'expérience a été réalisée sur un essai unique.

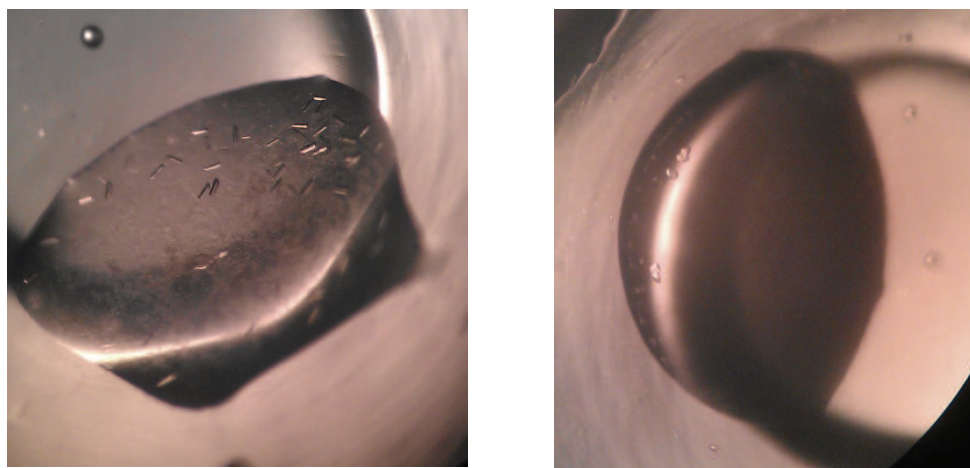
Ainsi sur les 4 mutants de DtMan E425 testés, aucun n'a présenté l'activité de thioglycosylation recherchée. Ces résultats sont comparables à ceux obtenus en transglycosylation dans le 1<sup>er</sup> chapitre. La problématique de modulation du ratio synthèse/hydrolyse de DtMan n'est donc pas réalisable par la simple suppression de l'activité hydrolytique, mais doit tenir compte de l'ensemble des interactions enzyme-substrat et enzyme-accepteur. Ces mutants, incapables d'hydrolyser le pNPMAN peuvent en revanche être utilisés pour obtenir des co-cristaux de protéines qui permettront de connaître les différentes interactions existantes entre l'enzyme et son substrat. L'étude approfondie des structures tridimensionnelles de DtMan *wt* et mutée, couplée aux techniques d'évolution dirigée des glycosidases demeurent les seules solutions envisageables pour faire de cette super-hydrolase une transglycosidase.

## 2. Etude cristallographique DtMan E425C

Lors de l'étude en activité d'hydrolyse des mutants DtMan E425X un comportement intéressant a été observé pour le variant E425C dont la valeur de pH optimal mesurée était dépendante de l'ajout ou non d'un agent réducteur (pH 9 et pH 4 respectivement). Cette observation nous a permis d'émettre l'hypothèse de la formation d'un pont disulfure au sein du site actif de l'enzyme en l'absence de DTT. Le changement conformationnel résultant serait en accord avec l'activité maximale d'hydrolyse relevée à pH acide alors que de manière générale la perte du résidu acide/base pour les autres mutants entraîne un déplacement du pH optimal vers des valeurs basiques de pH. Sur la séquence protéique de DtMan, et après alignement structurel avec d'autres  $\beta$ Man, deux Cystéines (C387 et C421) sont à proximité du résidu acide/base et peuvent être à l'origine de ce pont disulfure. Pour confirmer cette hypothèse une étude de cristallogénèse de DtMan E425C a été réalisée en se basant sur les

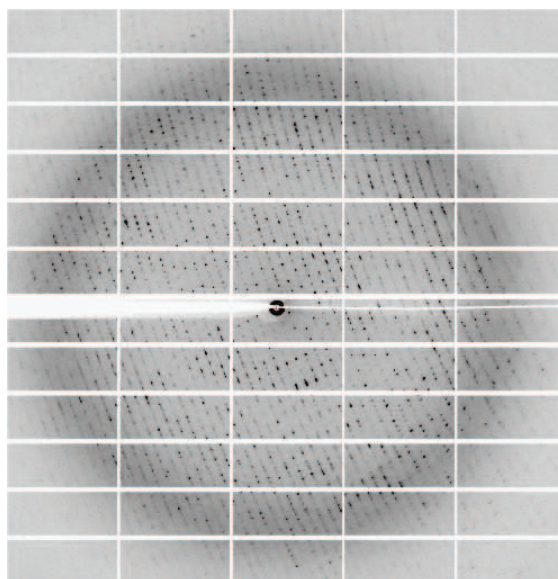


mêmes conditions de cristallisation que celles obtenues pour le type sauvage de DtMan (tampon NaHepes 100 mM pH 7,5, 30 % PEG 400 et  $\text{MgCl}_2$  200Mm, Chapitre 1, p. 79). Deux plaques ont dans un premier temps été préparées à partir de deux concentrations de protéine (10 et 15 mg/ml) en présence et en absence de DTT (100 mM). 12 compositions de solutions ont été préparées contenant 31 à 34 % de PEG 400 et 0,1 à 0,3 M de  $\text{MgCl}_2$ . Après incubation de ces deux plaques à température ambiante de petits cristaux ont pu être isolés en présence ou absence de DTT et de pNPMAN. Ils ont par la suite été utilisés pour obtenir de nouvelles plaques de cristallisation avec les mêmes solutions de cristallisation mais en utilisant la méthode d'ensemencement avec trois concentrations de DtMan E425C (5, 10 et 15 mg/ml). Plusieurs cristaux ont pu être visualisés et montrent l'influence de la présence de DTT dans la solution de cristallisation. En présence de l'agent réducteur, la protéine cristallise sous la forme de bâtonnet (**A**, Figure 67), tandis qu'en absence de celui-ci des cristaux de forme cubique sont observés (**B**).



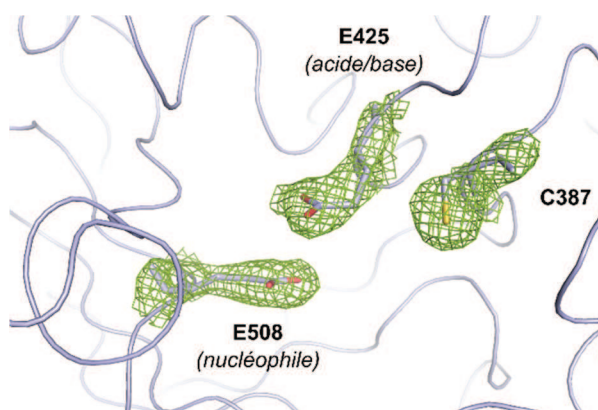
**Figure 67** : Cristaux de DtMan E425C en présence de DTT (**A**) et en absence de DTT (**B**).

Après cryoconservation (PEG 400 + huile) et congélation les cristaux ont été analysés sur la ligne Proxima 2 du Synchrotron SOLEIL. Une image de diffraction résolution de 3,5 Å a pu être obtenue pour un cristal sans DTT (Figure 68). Malgré une résolution faible les données ont été collectées pour obtenir un modèle général du mutant et observer la présence éventuelle d'un pont disulfure.



**Figure 68 :** Image de diffraction de DtMan E425C (3,5 Å).

Au moment de l'écriture de ce manuscrit, la résolution de la structure de DtMan E425C n'était pas complète, néanmoins une analogie avec le modèle préliminaire obtenu pour DtMan *wt* a pu être réalisée pour confirmer la possibilité de formation d'une liaison S-S. La structure du site actif de l'enzyme type sauvage étant bien défini il a été possible d'identifier chaque acide aminé le constituant. En effet plus la résolution d'une structure protéique est élevée (faible valeur en Å) meilleure sera la définition de la densité électronique autour de chaque résidu de la protéine. Dans le cas de DtMan la densité électronique du résidu voisin à l'acide/base E425 a pu être attribuée à la Cystéine C387 (Figure 69). En mutant le résidu E425 par une Cystéine, la proximité avec le résidu C387 n'est pas modifiée et compte-tenu de la réactivité de l'atome de soufre, à cette distance la formation d'un pont disulfure est possible.



**Figure 69 :** Représentation de la densité électronique entre les résidus C387, E425 et E508 de DtMan *wt*.

Grâce aux différents essais de cristallisation nous avons pu mettre en évidence que la présence ou non d'un agent réducteur semble avoir une influence sur la conformation de DtMan E425C, et donc sur son activité catalytique, puisque deux types de cristaux ont pu être

observés. D'autre part la résolution en parallèle de la structure de DtMan *wt* a confirmé la possibilité de formation d'un pont disulfure pour le mutant E425C compte-tenu de la proximité de la Cystéine 387 avec le résidu acide/base.

#### IV. Conclusion

Pour ce second chapitre consacré à la conception de mutants thioglycoligase de GH de *D. thermophilum*, plusieurs aspects ont été abordés :

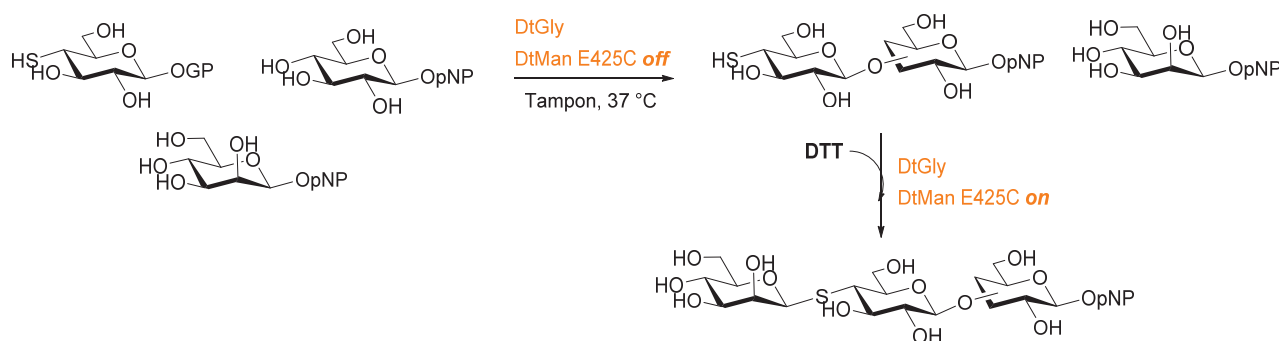
- Dans le cas de DtGly, l'objectif était de sélectionner le meilleur candidat thioligase pour l'application à la synthèse d'acides aminés glycosylés. Après identification des résidus acide/base E159 et nucléophile E324 de l'enzyme, deux mutants thioligases ont pu être obtenus par mutagenèse dirigée et exprimés chez *E. coli* : DtGly E159A et E159Q. La perte de l'activité hydrolytique due à la mutation du résidu acide/base a pu être confirmée chez les variants de DtGly qui semble tout deux de bons candidats thioligases. Néanmoins le rendement de production du mutant E159Q étant plus efficace ce dernier a été retenu pour le reste de l'étude. L'activité thioligase de l'enzyme a été criblée sur un panel de thiols aliphatiques et aromatiques de pKa varié. A l'issue de ce premier test, il a pu être démontré qu'en présence d'accepteurs de pKa < 7 l'enzyme retrouvait une activité catalytique comparable à l'activité hydrolytique de l'enzyme sauvage. Par la suite, ces résultats prometteurs nous ont poussés à étudier la relation structure/activité de DtGly E159Q, sur un panel d'accepteurs thioaryliques, afin d'anticiper la réaction de glucosylation d'acides aminés. La synthèse efficace des thioglycosides correspondants a révélé la flexibilité de l'enzyme vis-à-vis des fonctions portées par l'accepteur. Malgré tout, des difficultés ont été rencontrées dans le cas du 4-AMB (acide 4-mercaptobenzoïque), caractérisé par un groupement carboxylate à pH basique, suggérant une répulsion d'ordre électrostatique entre l'accepteur et la poche catalytique de l'enzyme. Cette hypothèse a été validée en retrouvant l'activité thioligase de l'enzyme pour les accepteurs dont la fonction acide carboxylique avait été protégée sous la forme d'un ester de méthyle.

Cette première étude a permis de mettre en évidence le potentiel du mutant DtGly E159Q pour la synthèse de thioglycosides et pourra donc être appliqué pour la synthèse enzymatique d'acides aminés glycosylés. D'autre part en tenant compte de l'étude de relation structure/activité et de l'essai infructueux de thioglycosylation réalisé sur la L-Cystéine, il est apparu évident que la synthèse d'un acide aminé non naturel était nécessaire puisqu'aucun, parmi les 20 présents dans la nature, ne possède à la fois une fonction thiol et un noyau aromatique.

- Dans le cas de DtMan l'approche suivie, plus fondamentale, a consisté à évaluer l'effet de mutations sur le contrôle du ratio hydrolyse/transglycosylation de l'enzyme qui ne présente en type sauvage aucune activité transglycosidase. Afin d'augmenter son potentiel de synthèse, il a été envisagé de supprimer l'activité hydrolytique de l'enzyme par mutation du résidu acide/base (E425) pour permettre le déplacement de l'équilibre réactionnel en présence d'un nucléophile fort. Dans cette optique une banque de 7 mutants de DtMan E425 a été

générée et exprimée chez *E. coli*. Leur caractérisation a permis de mettre en évidence la perte globale de l'activité hydrolytique. Quatre mutants ont alors été sélectionnés pour les essais de thioglycosylation en présence de l'espèce nucléophile *para*-Chlorothiophénol. Cependant au terme des différentes expériences aucune activité thioligase n'a pu être mise en évidence comme escompté. Cette étude permet de comprendre que le ratio hydrolyse/synthèse d'une enzyme ne peut pas être contrôlé par simple modification de son site actif et que d'autre part que toute glycosidase ne peut être convertie en thioglycoligase par simple mutation de son résidu acide/base. Une aptitude en transglycosylation de l'enzyme en type sauvage, et donc une bonne reconnaissance de l'accepteur, semble être un prérequis nécessaire pour l'élaboration de mutants thioglycoligases. Ainsi malgré la connaissance des mécanismes, d'apparence simple, des glycosidases, il n'existe aucune règle prédéfinie permettant de moduler à façon l'activité catalytique de ces enzymes.

- L'étude de cristallogénèse de DtMan E425C couplée au modèle préliminaire de DtMan *wt* a permis de mieux comprendre le comportement atypique de ce mutant en fonction de l'ajout ou non d'un agent réducteur. En effet la proximité directe du résidu Cystéine C387 avec le résidu E425 permet dans le cas du mutant E425C la formation du liaison S-S. Ce pont disulfure ainsi crée produit une modification de la conformation de la protéine, comme en atteste les différentes formes de cristaux obtenus (bâton ou cubique) et par conséquent a une répercussion sur l'activité catalytique de l'enzyme. Contrairement aux autres mutants caractérisés par un pH Opt  $\geq 9$ , DtMan E425C présente une activité maximale d'hydrolyse à pH acide. Compte-tenu de cette propriété, si l'enzyme avait présenté une activité thioligase il aurait pu être intéressant d'utiliser ce pont disulfure comme un interrupteur moléculaire de l'activité thioligase pour la synthèse one-pot d'oligothiosaccharides. Dans un milieu réactionnel contenant une transglycosidase (par exemple DtGly) et DtMan E425C il serait possible de réaliser la synthèse d'un disaccharide par transglycosylation et l'ajout successif de DTT permettrait la glycosylation d'une 3<sup>ème</sup> unité sucre *via* une liaison S-glycosidique par DtMan E425C (Schéma 51).



**Schéma 51** : Application potentielle du mutant DtMan E425C.



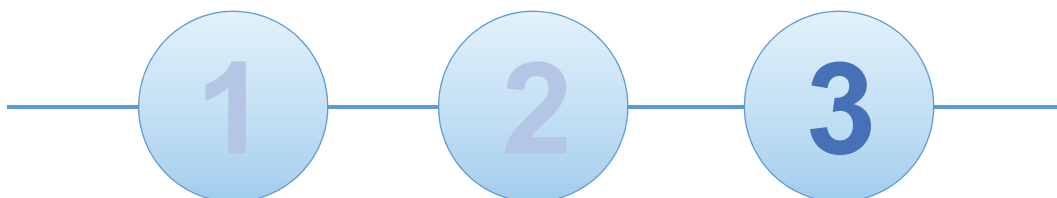
## Bibliographie

- Almendros, M., Danalev, D., Francois-Heude, M., Loyer, P., Legentil, L., Nugier-Chauvin, C., Daniellou, R., and Ferrieres, V.; *Org. Biomol. Chem.* **(2011)**, 9, 8371-8378
- Almendros, M., François-Heude, M., Legentil, L., Nugier-Chauvin, C., Daniellou, R., and Ferrières, V.; *ARKIVOC* **(2013)**, 2, 123-132
- Bruneau, A., Roche, M., Hamze, A., Brion, J. D., Alami, M., and Messaoudi, S.; *Chem. Eur. J.* **(2015)**, 21, 8375-8379
- Driguez, H.; *Chembiochem* **(2001)**, 2, 311-318
- Fort, S., Varrot, A., Schulein, M., Cottaz, S., Driguez, H., and Davies, G. J.; *Chembiochem* **(2001)**, 2, 319-325
- Gillam, S., and Smith, M.; *Genes* **(1979)**, 8, 81-97
- Jahn, M., Marles, J., Warren, R. A. J., and Withers, S. G.; *Angew. Chem., Int. Ed.* **(2003)**, 42, 352-354
- Kim, Y. W., Lovering, A. L., Chen, H. M., Kantner, T., McIntosh, L. P., Strynadka, N. C. J., and Withers, S. G.; *J. Am. Chem. Soc.* **(2006)**, 128, 2202-2203
- Mullegger, J., Chen, H. M., Chan, W. Y., Reid, S. P., Jahn, M., Warren, R. A. J., Salleh, H. M., and Withers, S. G.; *Chembiochem* **(2006)**, 7, 1028-1030
- Mullegger, J., Jahn, M., Chen, H. M., Warren, R. A. J., and Withers, S. G.; *Protein Eng., Des. Sel.* **(2005)**, 18, 33-40
- Samoshin, A. V., Dotsenko, I. A., Samoshina, N. M., Franz, A. H., and Samoshin, V. V.; *Int. J. Carbohydr. Chem.* **(2014)**, 2014, 1-8
- Sasaki, S., Kitamura, S., Negoro, N., Suzuki, M., Tsujihata, Y., Suzuki, N., Santou, T., Kanzaki, N., Harada, M., Tanaka, Y., Kobayashi, M., Tada, N., Funami, M., Tanaka, T., Yamamoto, Y., Fukatsu, K., Yasuma, T., and Momose, Y.; *J. Med. Chem.* **(2011)**, 54, 1365-1378
- Wardrop, D. J., and Waidyarachchi, S. L.; *Nat. Prod. Rep.* **(2010)**, 27, 1431-1468
- Witczak, Z. J., Sun, J. M., and Mielguj, R.; *Bioorg. Med. Chem. Lett.* **(1995)**, 5, 2169-2174



## **Chapitre 3**

### **Validation du concept de thioglycosylation enzymatique d'un acide aminé**

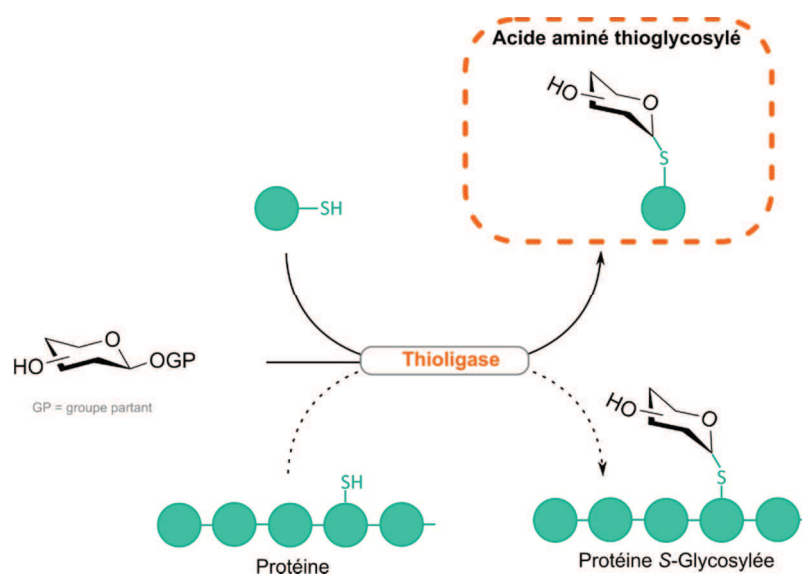






## Chapitre 3 – Validation du concept de thioglycosylation enzymatique d'un acide aminé

Suite au développement et à la caractérisation d'une banque de glycosidases natives thermorésistantes de *D. thermophilum*, deux candidates potentielles, DtGly et DtMan, ont été retenues pour la production de mutants thioligases. Par mutagenèse dirigée 9 enzymes mutées (dont 2 pour DtGly et 7 pour DtMan) ont pu être obtenues en remplaçant le résidu acide/base par un acide aminé catalytiquement inactif. Après avoir mis en évidence la perte générale de l'activité d'hydrolyse suite à la mutation, une étude a été effectuée pour déterminer leur activité en thioglycosylation et a permis de sélectionner le meilleur candidat thioligase, DtGly E159Q. Ce biocatalyseur s'est en effet révélé efficace pour la synthèse de thioglucosides à partir d'accepteurs de type thioarylique. La dernière phase de ce projet est désormais atteinte et vise à valider le concept innovant de S-glycosylation enzymatique d'acides aminés à partir d'une thioglycoligase (Schéma 52). En démontrant la faisabilité de ce type de couplage le potentiel des thioligases pourrait alors être étendu à la production de thioglycoprotéines, où aucune méthodologie de synthèse par voie enzymatique n'a été décrite à ce jour.



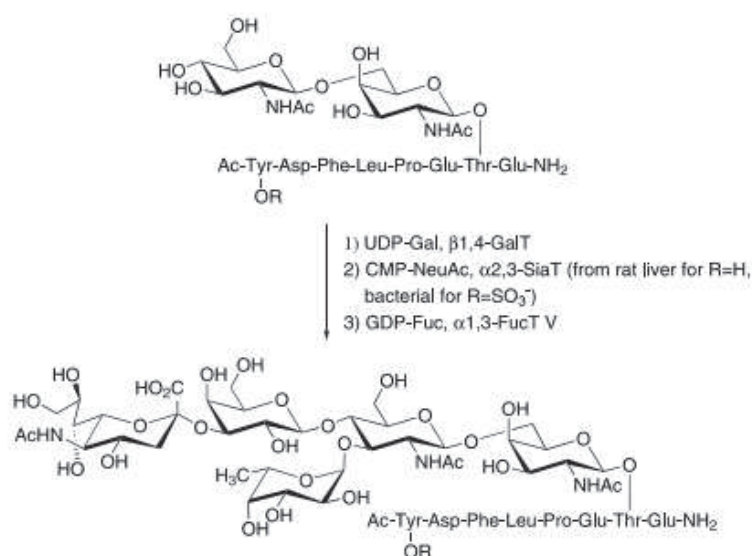
**Schéma 52 :** Validation du concept de thioglycosylation enzymatique d'acides aminés, vers la synthèse de protéines S-glycosylées.

### I. Introduction à la glycosylation enzymatique de protéines

La glycosylation est une des modifications post-traductionnelles rencontrées chez les protéines. Considérée comme l'une des plus complexes son étude suscite un vif intérêt dans le monde scientifique tant les relations établies entre les glycanes et les protéines sont étroites. (Varki, **1993**) D'un point de vue chimique la diversité de cette transformation est due à la possibilité d'assembler une grande variété de monosaccharides pour former des glycanes de longueurs, séquences ou de configurations anomériques ( $\alpha$ ,  $\beta$ ) variées. Ces

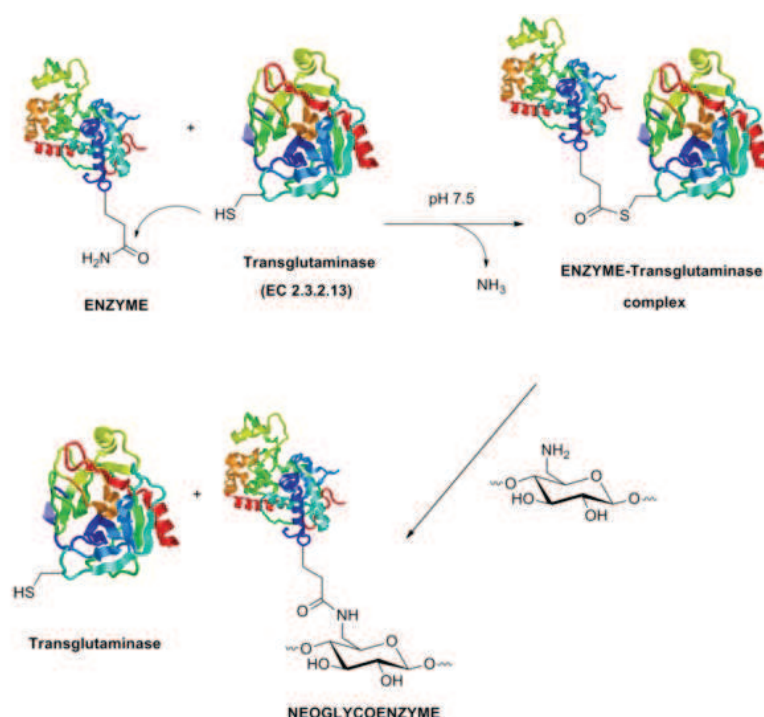
oligosaccharides peuvent alors être fixées sur différents sites de glycosylation de la protéine et subir de nouvelles fonctionnalisations (phosphorylation, acétylation, méthylation,...). Enfin d'un point de vue biologique, la composition des glycanes est dépendante du type de cellules ou de tissus où la modification post-traductionnelle a lieu, rajoutant ainsi un degré de diversité supplémentaire à la glycosylation des protéines.(Thayer, **2006**) La présence d'oligosaccharides sur les séquences peptidiques est associée à de nombreux processus biologiques, du correct repliement des protéines à leur sécrétion, ainsi que dans les phénomènes d'adhésion cellulaire, de transduction du signal ou encore dans la réponse immunitaire. Par conséquent, des erreurs de glycosylation peuvent engendrer l'apparition de pathologies diverses. Les protéines sont souvent exprimées sous la forme de mélanges hétérogènes ce qui complique la purification et l'isolation de formes homogènes pourtant nécessaires pour la compréhension des relations structure/activité. Pour répondre à ce challenge, de nombreuses méthodologies ont été développées pour permettre la synthèse spécifique *in-vitro* ou *in-vivo* de protéines glycosylées et *in fine* produire de nouveaux agents thérapeutiques.(Gamblin, **2009**)

La glycosylation est le résultat de l'action concomitante d'une large variété d'enzymes, les GT et GH, et par conséquent la reproduction *in-vitro* d'une telle machinerie enzymatique est impossible. Néanmoins de nombreux exemples ont été décrits prouvant l'intérêt de ces biocatalyseurs pour la synthèse ou la modification d'oligosaccharides complexes. L'utilisation de ces outils enzymatiques a alors été étendue pour la production de *N*- ou *O*-glycopeptides et glycoprotéines. Ainsi les GT peuvent être employées de manière séquentielle pour permettre le transfert d'unités sucre sur des glycanes simples. En suivant cette stratégie, la synthèse chimio-enzymatique d'un fragment de la P-sélectine ligand-1 (P-selectin glycoprotein-ligand-1, PSGL-1) a été rapportée à partir d'un *O*-glycopeptide synthétique, produit par SPPS, dont le disaccharide a été allongé de trois unités sucres par des glycosyltransférases (Schéma 53).(Koeller, **2000a**; Koeller, **2000b**)



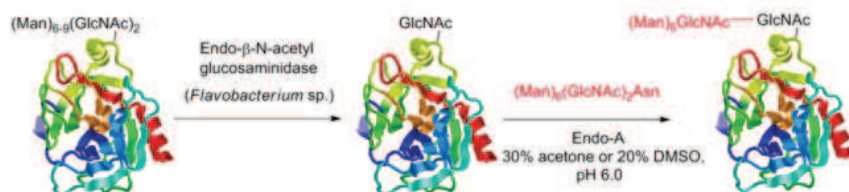
**Schéma 53** : Elaboration séquentielle d'un fragment peptidique du PSGL-1.(Thayer, **2006**)

Les GT permettent également le couplage direct d'une unité sucre sur une protéine, par exemple en 2007 Ramakrishnan *et coll.* ont montré que la polypeptide-*R-N*-acetylgalactosaminyltransférase pouvait être utilisée pour catalyser le transfert du motif GalNAc vers le résidu Thréonine ou Sérine d'un peptide *via* une liaison *N*-glycosidique. (Ramakrishnan, **2007**) Les transglutaminases sont aussi rencontrées pour réaliser ce type de couplage, l'enzyme réagit de manière spécifique sur le résidu Glutamine de la protéine à modifier conduisant à la formation d'un complexe protéine-transglutaminase. La présence dans le milieu réactionnel d'un donneur de glycoside portant une fonction amine primaire, nucléophile, permet le clivage de ce complexe pour former la *N*-glycoprotéine correspondante (Schéma 54). (Villalonga, **2014**)



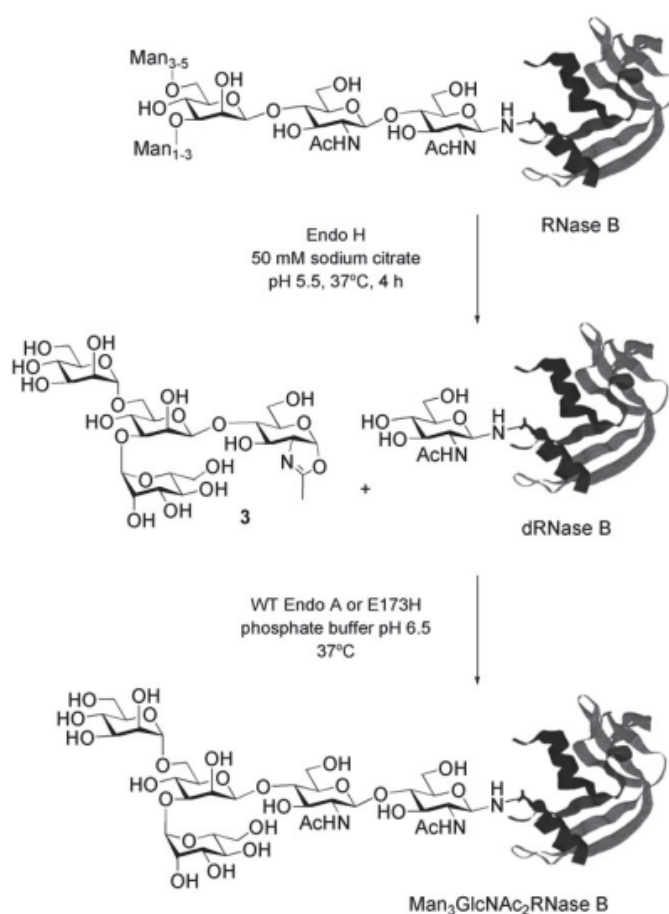
**Schéma 54** : Synthèse de *N*-glycoprotéines catalysée par les transglutaminases.

Les glycoside hydrolases présentent également un intérêt pour la production de protéines glycosylées. Grâce à leur activité d'hydrolyse elles permettent la déglycosylation partielle de glycoprotéines naturelles qui peuvent par la suite être modifiées par réaction de transglycosylation à partir de GT, comme décrit précédemment, ou de GH natives ou mutées. Parmi ces dernières, les endo- $\beta$ -*N*-acétylglucoaminidases (ENGases) ont suscité un fort intérêt. Ces enzymes sont naturellement responsables de la libération de *N*-glycanes par clivage de la liaison  $\beta$ -(1,4) du motif *N,N'*-diacétylchitobiose de cœur des *N*-glycoprotéines. (Li, **2005**) Néanmoins certaines représentantes, telles que l'Endo-A de *Arthrobacter protophormiae* (Takegawa, **1991**) ou l'Endo-M de *Mucor hiemalis* (Yamamoto, **1994**) sont capables du transfert de chaînes intactes plus ou moins longues de *N*-glycanes sur des séquences peptidiques présentant un motif GlcNAc (Schéma 55).



**Schéma 55** : Synthèse de *N*-glycoprotéines par réaction de transglycosylation catalysée par l'Endo-A d'*A. protophormiae*.(Villalonga, **2014**)

Comme pour l'ensemble des GH, les rendements de transglycosylation des ENGases sont souvent limités par la formation du produit d'hydrolyse. Pour résoudre ce problème, en 2005 l'équipe du Prof. Wang a démontré que l'activité de synthèse des ENGases pouvait être optimisée en utilisant des donneurs de glycosides activés présentant un groupement oxazoline. En suivant cette méthodologie l'équipe a rapporté la synthèse de fragments peptidiques tri- et pentaglycosylés dérivés de la protéine transmembranaire gp 41 du virus HIV-1 à partir de l'Endo-A avec des rendements de 73 et 75 % respectivement.(Li, **2005**) Enfin une autre stratégie d'optimisation des ENGases repose sur la génération de mutants glycosynthases complètement dépourvus d'activité d'hydrolyse. Yamamoto *et coll.* ont rapporté pour la première fois la production de mutants glycosynthase de l'Endo-M en remplaçant le résidu N175 par mutagenèse dirigée. Le criblage des différents variants obtenus a permis d'identifier l'Endo-M N175A capable de transférer la séquence Man<sub>9</sub>GlcNAc-oxazoline sans hydrolyser le glycopeptide naturel résultant. La même approche a été suivie pour produire des mutants glycosynthase de l'Endo-A. L'équipe du Prof. Fairbanks a notamment rapporté le potentiel en transglycosylation du variant Endo-A E173H alors capable de synthétiser une forme homogène de la ribonucléase B avec une faible activité d'hydrolyse résiduelle (Schéma 56).(Heidecke, **2008**)

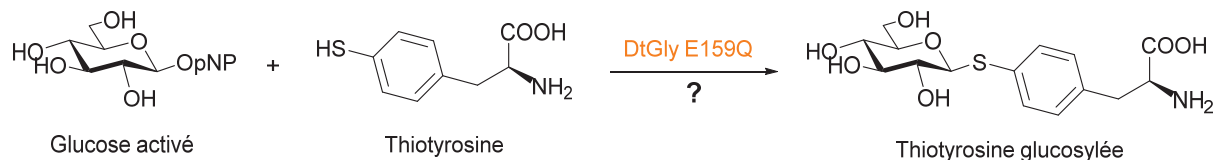


**Schéma 56 :** Synthèse enzymatique du glycoforme homogène  $\text{Man}_3\text{GlcNAc}_2$  de la ribonucléase B via des ENGases natives et mutées. (Heidecke, 2008)

Les GT et GH sont donc des outils prometteurs et efficaces pour la synthèse de formes homogènes de glycoprotéines. Néanmoins à l'heure actuelle les efforts se sont essentiellement focalisés sur l'emploi d'enzymes capables de catalyser la formation de liaisons *N*- et *O*-glycosidiques. Rares sont les exemples dans la littérature traitant de l'utilisation de ces biocatalyseurs la création de liaisons *S*-glycosidiques malgré l'intérêt des thioglycoprotéines. En raison, les *S*-GT naturelles n'ont été que très récemment découvertes, trois représentantes sont à ce jour décrites parmi lesquelles la ThuS capable de reproduire *in-vitro* le thioglycopeptide Thurandacine.(Wang, **2014**) Chez les glycosidases, le développement du concept de thioglycoligases en 2003 par l'équipe du Prof. Withers a permis de prouver l'intérêt de ce type d'enzyme pour la synthèse de thioglyconjugués, cependant leur application n'a encore jamais été décrite pour le couplage direct d'une unité sucre sur un résidu protéique *via* un atome de soufre.

Dans cette optique nous nous sommes appliqués à explorer le potentiel de thioglycosylation de notre mutant DtGly E159Q pour la synthèse d'un acide aminé S-glycosylé. Cependant l'essai infructueux réalisé sur la Cystéine, seul acide aminé naturel présentant une fonction thiol, ainsi que les études de relation structure/activité ont révélé la nécessité d'utiliser des

accepteurs présentant un motif thioarylique avec une fonction thiol de  $pK_a < 7$ . Nous nous sommes alors orientés vers la synthèse de la Thiotyrosine comme modèle pour la glycosylation enzymatique d'acides aminés catalysée par la thioligase DtGly E159Q. Ce dernier chapitre sera donc consacré à la synthèse de l'acide aminé non-naturel (AAn) Thiotyrosine ainsi qu'à son utilisation en tant qu'accepteur dans la réaction de thioglycosylation enzymatique catalysée par DtGly E159Q (Schéma 57).

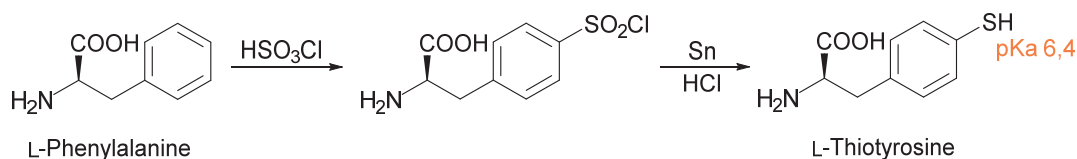


**Schéma 57** : Glycosylation enzymatique de la Thiotyrosine par DtGly E159Q.

## II. Synthèse de l'accepteur de glycoside, la Thiotyrosine

### 1. Synthèse de l'acide aminé non-naturel

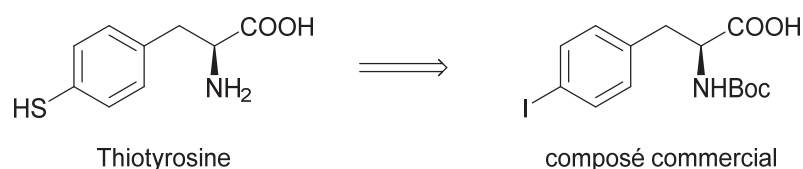
Historiquement la synthèse de la Thiotyrosine (Ttyr) a été décrite pour la première fois en 1983 par Hescher *et coll.* en deux étapes à partir de la L-Phénylalanine. (Hescher, **1983**) Cette méthode, reprise quelques années plus tard par l'équipe du Prof. DeGrado, repose sur la chlorosulfonation de l'acide aminé commercial suivi de sa réduction par un mélange Sn.HCl (Schéma 58). (Lu, **1997**) Néanmoins, malgré la simplicité de la synthèse proposée, cette dernière n'a pas été retenue dans le cadre de ce projet de par la relative toxicité de l'étain ainsi que les mesures de sécurité à prendre lors de sa manipulation.



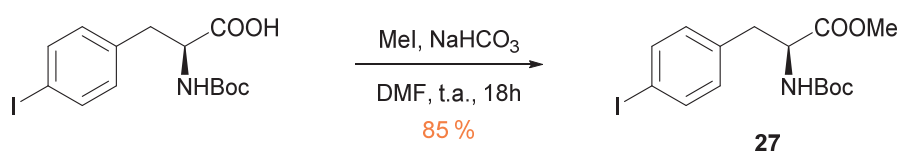
**Schéma 58** : 1<sup>ère</sup> synthèse en 2 étapes de la Thiotyrosine. (Hescher, **1983**; Lu, **1997**)

Par la suite diverses méthodes ont été décrites permettant la fonctionnalisation d'acides aminés aryliques par couplages catalysés *via* des métaux de transition et ce avec de bons rendements tout en assurant la régiosélectivité de la réaction. (Forbes, **2012**; Meyer, **2010**; Rajagopalan, **1996**) C'est donc en s'inspirant de ces méthodes que la synthèse de la Ttyr a été envisagée, à partir d'un dérivé halogéné commercial de la L-Phénylalanine (Schéma 59).



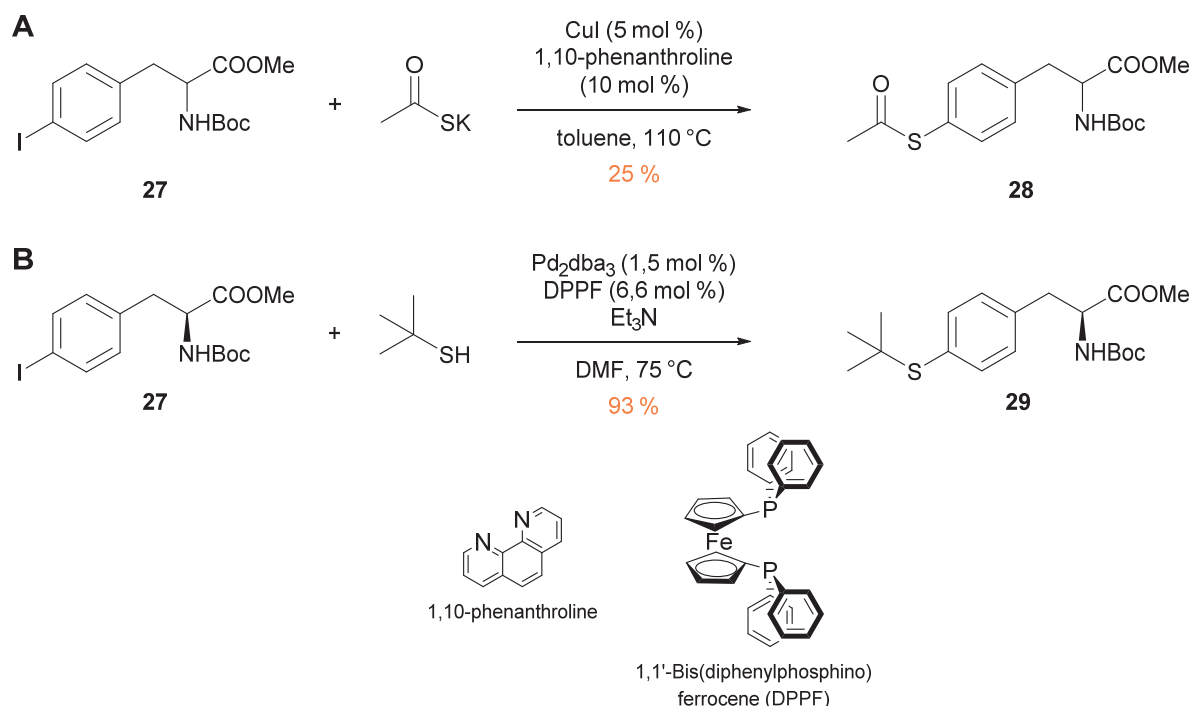
**Schéma 59** : Schéma de rétrosynthèse de la Thiotyrosine.

De manière générale la littérature fait référence à la synthèse d'AAns à partir de dérivés protégés sur les fonctions amine et acide carboxylique. Le groupe Boc est classiquement utilisé pour protéger la fonction amine, tandis que la fonction acide carboxylique est masquée par estérification. Ainsi la première étape de la voie de synthèse proposée ici a consisté à la protection de la fonction acide carboxylique du dérivé de la L-Phénylalanine en milieu basique en présence de Iodomethane (Schéma 60). (Meyer, **2010**) Le composé **27** résultant a été obtenu avec un excellent rendement de 85 %. La mesure du pouvoir rotatoire (**27**,  $[\alpha]^{20}_D = +103$  ;  $c = 1$ , EtOAc) comparée à celui du produit de départ ( $[\alpha]^{20}_D = +22,5$  ;  $c = 1$ , EtOAc) a permis de confirmer la conservation du caractère dextrogyre.

**Schéma 60** : Estérification du dérivé halogéné commercial de la Tyrosine.

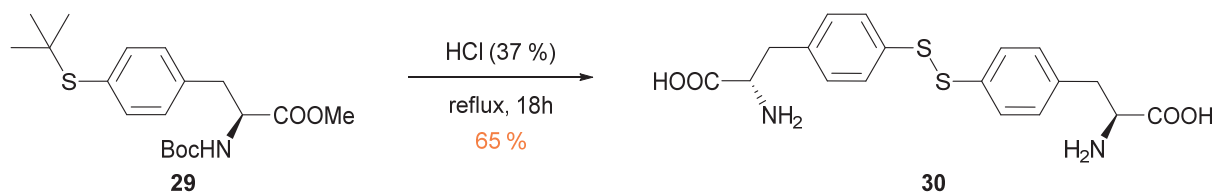
Par la suite, l'incorporation de l'atome de soufre a été réalisée en suivant deux méthodes de couplage, l'une utilisant le Cu(I) (**A**, Schéma 61) et la seconde le Pd(0) (**B**), de manière à sélectionner la plus efficace pour le reste de la synthèse. Dans le premier cas, la réaction de couplage s'effectue à partir de l'acide thioacétique en présence du système catalytique CuI/1,10-phenanthroline et d'une base forte (DIPEA) dans le toluène à reflux. En 2013, Soria-Castro *et coll.* ont proposé une méthode de synthèse efficace de thioacétates aryliques en présence de Cu(I) à partir de différents sels de thiolates. Leur utilisation ne nécessite donc pas l'ajout d'une base forte (telle que la DIPEA), la fonction thiol étant déjà sous forme thiolate, et conduit à la formation des produits correspondant avec de très bons rendements. (Soria-Castro, **2013**) Dans cette optique le thioacétate de potassium a été retenu pour le couplage sur le composé **27** en présence du système CuI/1,10-phenanthroline sous atmosphère inerte afin d'éviter l'empoisonnement du catalyseur par l'oxygène de l'air. Le composé **28** résultant a été obtenu avec un faible rendement de 25 % mais ceci reste en accord avec les résultats décrits dans la littérature. En effet, Forbes *et coll.* ont rapporté la synthèse de dérivés thioacétate de la Phénylalanine à partir de peptides courts en solution ou greffés sur support solide avec des rendements de 29 % et 96 % respectivement. (Forbes, **2012**) Le couplage au Cuivre(I) est donc plus adaptée à la modification d'acides aminés dans les procédés de SPSS qu'à celle d'acides aminés libres en solution ce qui est notre cas.





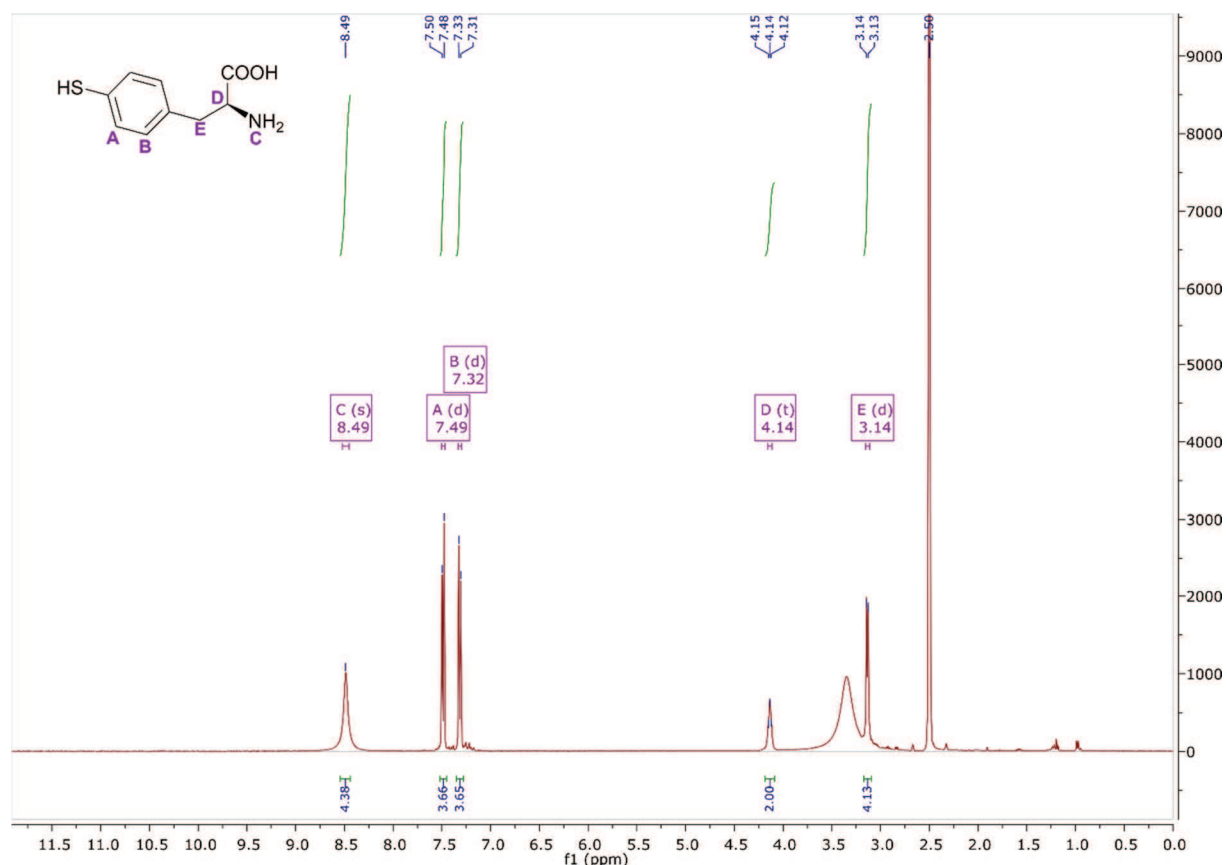
**Schéma 61** : Création de la liaison C-S par couplage métallo-catalysé. **(A)** catalyse au Cuivre (I) et **(B)** catalyse au Palladium (0).

La seconde méthode testée repose sur le couplage d'un thiol *via* le système  $\text{Pd}_2\text{dba}_3$  / DPPF en présence de triéthylamine ( $\text{Et}_3\text{N}$ ). A partir des deux solutions  $\text{Pd}_2\text{dba}_3$  / DPPF et **27** /  $\text{Et}_3\text{N}$  réunies dans le DMF, préalablement dégazé, le *tert*-butylthiol est ajouté sous atmosphère inerte comme décrit dans la littérature. (Rudolf, **2013**) Après 1h de réaction à 75 °C le produit de départ est totalement consommé, après purification le composé **29** est obtenu avec un excellent rendement de 93 %. Ainsi ce 2<sup>ème</sup> système catalytique s'avère plus approprié à la fonctionnalisation de la L-Phénylalanine libre en solution (i.e. non supportée). La déprotection générale du composé **29** a été conduite en milieu acide concentré à reflux pendant une nuit selon les conditions déjà décrites dans la littérature (Schéma 62). (Rudolf, **2013**)



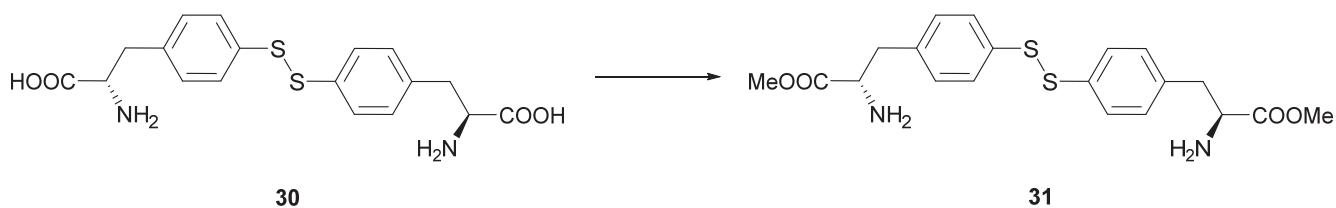
**Schéma 62**: Déprotection par traitement acide de la Thiotyrosine.

Au terme de ces trois étapes (protection, couplage, déprotection), la Ttyr (**30**) a pu être obtenue avec un bon rendement global de 59 %. Les analyses RMN et HRMS ont permis de mettre en évidence la forme disulfure de la Thiotyrosine. En effet sur le spectre RMN (Figure 70) aucun pic n'a pu être attribué au proton de la fonction thiol, tandis que le spectre de masse révèle un pic majoritaire à  $[M + H]^+ = 393$  m/z (ESI+) qui correspond à la masse de la forme dimérique de la Thiotyrosine ( $M_{\text{monomère}} = 197$  g.mol<sup>-1</sup>).



**Figure 70** : Spectre RMN de la Thiotyrosine dans le DMSO.

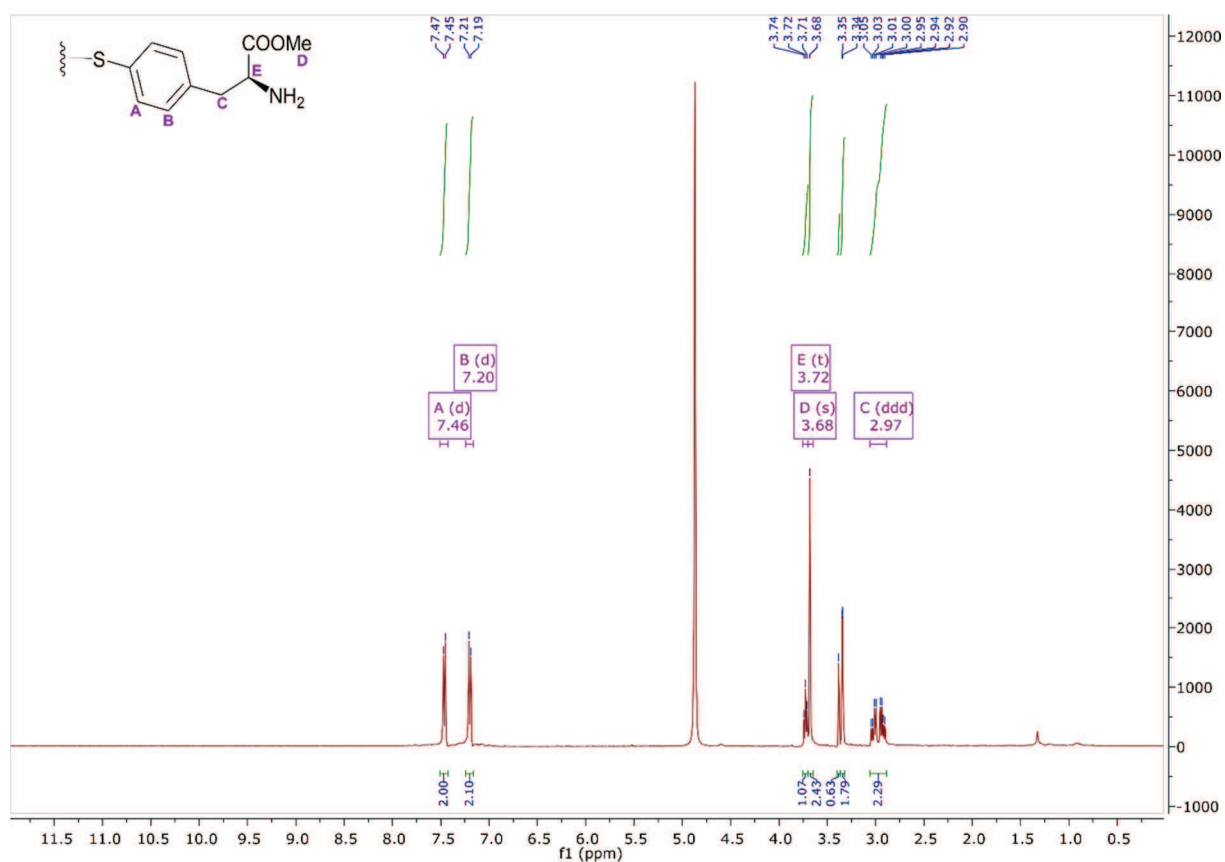
Comme décrit dans le chapitre 2, les réactions de thioglycosylation catalysées par DtGly E159Q en présence d'accepteurs portant une fonction  $-COOH$  ne sont favorisées que lorsque cette dernière est protégée sous forme d'ester. Ainsi l'estérification de Ttyr a été envisagée dans un premier temps en présence de Iodométhane en milieu basique (entrée 1, Schéma 63). Malheureusement aucun produit n'a pu être détecté à l'issue de la réaction. Deux autres alternatives ont alors été suivies. La première a nécessité l'emploi de  $SOCl_2$  dans le MeOH (entrée 2) mais une fois de plus aucun produit de réaction n'a pu être isolé. En suivant le protocole d'estérification dans le MeOH en présence d'une quantité catalytique de  $H_2SO_4$  (décrit dans le chapitre 2, p. 127) la formation de la Ttyr estérifiée **31** a pu être observée en CCM et par RMN du produit brut mais après purification seuls 4 mg (2,4 % de rendement) de produit ont été isolés (entrée 3).



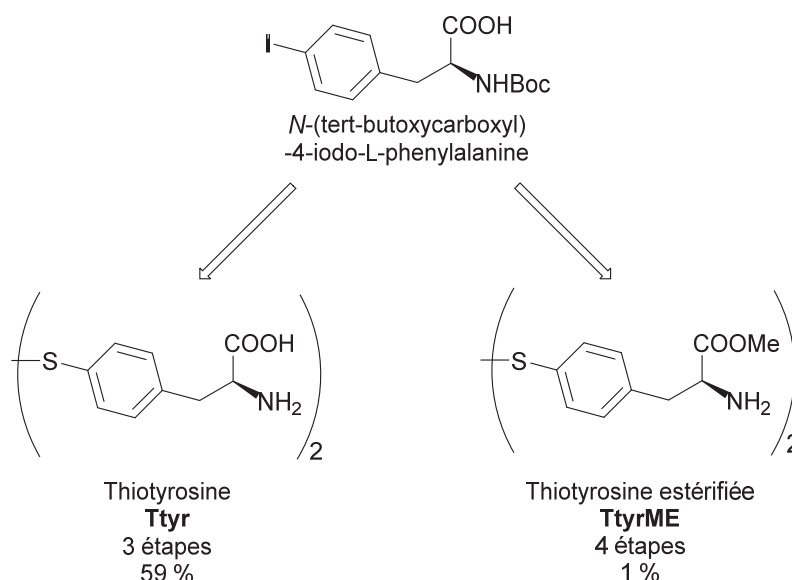
Entrée	Réactifs	Conditions	<b>31</b> Rdt (%)
<b>1</b>	MeI/NaHCO <sub>3</sub>	DMF, ta, 18h	n.d.
<b>2</b>	SOCl <sub>2</sub> /MeOH	reflux, 4h	n.d.
<b>3</b>	MeOH/H <sub>2</sub> SO <sub>4</sub> cat.	reflux, 18h	2,4

Schéma 63 : Essais d'estérification de la Thiotyrosine.

La Thiotyrosine estérifiée (**TtyrME**), de poids moléculaire  $M = 211 \text{ g.mol}^{-1}$ , a été identifiée par RMN néanmoins l'analyse HRMS a révélé une fois de plus la présence de la molécule sous la forme du dimère avec une masse  $[M+H]^+ = 421 \text{ m/z}$  (Figure 71).

Figure 71 : Spectre RMN de TtyrME **31** dans le MeOD.

Au bilan la Thiotyrosine, Ttyr **30**, a été synthétisée en suivant une voie de synthèse en trois étapes avec un rendement global de 59 % tandis que la forme protégée sous la forme d'un ester a été préparée en quatre étapes avec un faible rendement final d' 1 % (Schéma 64).



**Schéma 64** : Bilan des synthèses de la Thiotyrosine et de la Thiotyrosine estérifiée.

Avant d'envisager les essais de thioglycosylation de DtGly E159Q sur l'AAn des études préliminaires ont été réalisées afin d'évaluer le comportement de TtyrME en solution et ainsi déterminer les conditions les plus favorables au couplage glycosidique. Pour cela, les différents essais ont été réalisés à partir de la Ttyr non protégée, la forme estérifiée disponible en faible quantité a été réservée pour les couplages enzymatiques.

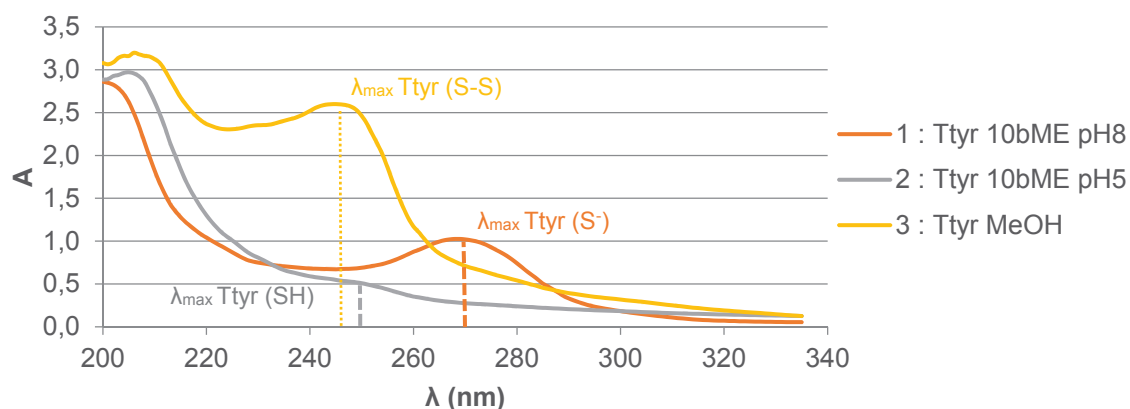
## 2. Etude de la Thiotyrosine en solution

La réaction de thioglycosylation enzymatique n'est favorisée qu'en présence d'accepteur nucléophile fort, c'est-à-dire une entité présentant une fonction thiol facilement déprotonable (proton acide) et qui se trouvera sous la forme de thiolate majoritairement en solution. L'objectif de cette étude est donc d'établir les conditions de réaction garantissant la présence de TtyrME en solution et surtout de l'espèce thiolate. Compte-tenu de la faible quantité de TtyrME disponible (i.e., 4 mg), les essais ont été réalisés avec la Thiotyrosine non protégée, Ttyr. Classiquement la rupture de liaisons disulfures peut être réalisée en ajoutant un agent réducteur, tel que le  $\beta$ -Mercaptoéthanol ( $\beta$ ME), dans le milieu réactionnel. Ainsi l'ajout de ce type de composé dans une solution contenant Ttyr sous forme de dimère devrait permettre de cliver la liaison S-S. L'influence du pH sur la formation de l'espèce thiolate a été étudiée, en plaçant la Thiotyrosine dans des milieux tamponnés à pH 5 et pH 8 en présence de  $\beta$ ME. Une solution contrôle a été préparée en solubilisant Ttyr dans du méthanol. Les trois solutions ont été réalisées à partir d'une solution mère de Ttyr (20 mM) dans le MeOH (Tableau 21).

Solution 1	Solution 2	Contrôle
Ttyr 1mM, $\beta$ ME 10 mM	Ttyr 1mM, $\beta$ ME 10 mM	Ttyr 1 mM
Tampon phosphate 20 mM pH 8	Tampon phosphate 20 mM pH 5	MeOH

**Tableau 21** : Composition des solutions 1 et 2 de Ttyr en milieu tamponné à pH 5 et pH 8 et de la solution contrôle.

Pour suivre le comportement de la Thiotyrosine suivant ces différentes conditions, les propriétés UV actives de la molécule ont été exploitées. Ainsi après incubation pendant 45 min à 37 °C, le spectre d'absorbance UV-visible de chaque échantillon a été enregistré (Figure 72). La comparaison des trois spectres permet de mettre en évidence la forte dépendance de Ttyr vis-à-vis du pH. En effet à pH 5 l'AAn présente un profil UV-visible comparable à celui du contrôle (dimère de Ttyr) avec 2 pics d'absorbance relevé à  $\lambda_{\max} = 208$  et 251 nm. En revanche à pH basique la molécule présente un pic intense d'absorbance à  $\lambda_{\max} \approx 273$  nm ce qui est caractéristique de la présence d'une nouvelle espèce en solution. En effet à pH basique ( $> pK_{a_{SH}} = 6,4$ ) et en présence de  $\beta$ ME, la forme thiolate Ttyr<sub>S-</sub> est favorisée. La conjugaison entre les trois doublets non-liants du soufre et le système  $\pi$  du noyau aromatique confère une propriété UV-visible unique à la molécule. Dans le cas de Ttyr<sub>SH</sub> et Ttyr<sub>S-S</sub>, cette délocalisation électronique n'est pas possible, ou reste limitée aux deux formes mésomères du cycle aromatique, les deux molécules présentent donc une activité UV-visible similaire. Cette observation est conforme aux données déjà décrites dans la littérature puisqu'en 2012 Forbes *et coll.* ont rapporté l'influence du pH sur la Thiotyrosine en solution. A pH acide la forme thiol est favorisée avec un  $\lambda_{\max} = 249$  nm tandis qu'à pH basique la forme thiolate est présente en solution avec un  $\lambda_{\max} = 276$  nm. (Forbes, 2012) Les résultats présentés au terme de cette étude sont convaincants puisque la forme thiolate de Ttyr peut être assurée en solution à pH basique en présence de 10 équivalents  $\beta$ ME.



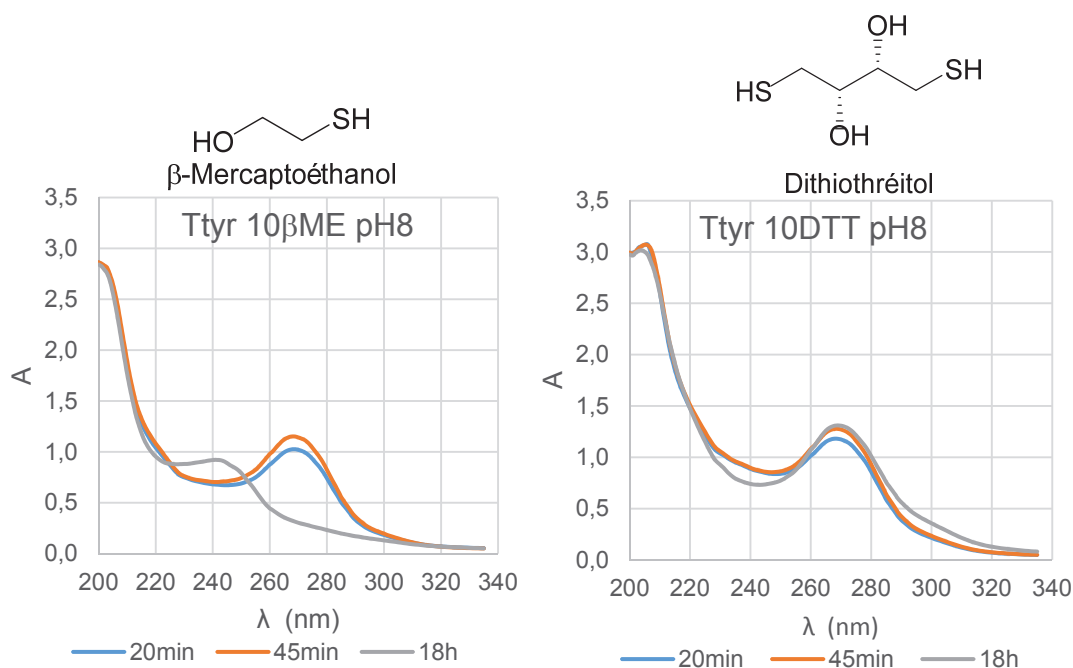
**Figure 72** : Spectres d'absorbance de la Thiotyrosine en solution après 45 min d'incubation à 37 °C.

Enfin une dernière étude a été réalisée afin d'évaluer la stabilité de l'agent réducteur en solution au cours du temps. En effet les synthèses de thioglucosides catalysées par DtGly E159Q ont été réalisées sur des temps longs de réaction (i.e. 18h, Chapitre 2) pour assurer la consommation totale du produit de départ. Il est donc nécessaire que la forme monomérique Ttyr soit présente tout au long de la réaction enzymatique. Pour cela deux agents réducteurs ont été testés, le  $\beta$ ME et le dithiothréitol (DTT), et introduits en fort excès dans la solution pour assurer la conversion totale du dimère Ttyr<sub>S-S</sub> sous la forme de thiolate (ratio 1:10, Ttyr / agent réducteur). Les solutions ont été préparées de la même manière que précédemment à partir d'une solution mère de Ttyr à 20 mM dans le méthanol (Tableau 22).

Solution 1	Solution 2
Ttyr 1mM, $\beta$ ME 10 mM	Ttyr 1mM, DTT 10 mM
Tampon phosphate 20 mM pH 8	Tampon phosphate 20 mM pH 8

**Tableau 22** : Composition des solutions 1 et 2 pour l'étude de stabilité du  $\beta$ ME et du DTT en solution.

La présence de l'espèce thiolate en solution a été suivie par mesure des spectres d'absorbance UV-visible des deux solutions après  $t = 20$  min, 45 min et 18 h d'incubation à 37 °C sous agitation (Figure 73).



**Figure 73** : Spectres d'absorbance de la Thiotyrosine en solution à pH basique en présence de  $\beta$ ME ou de DTT.

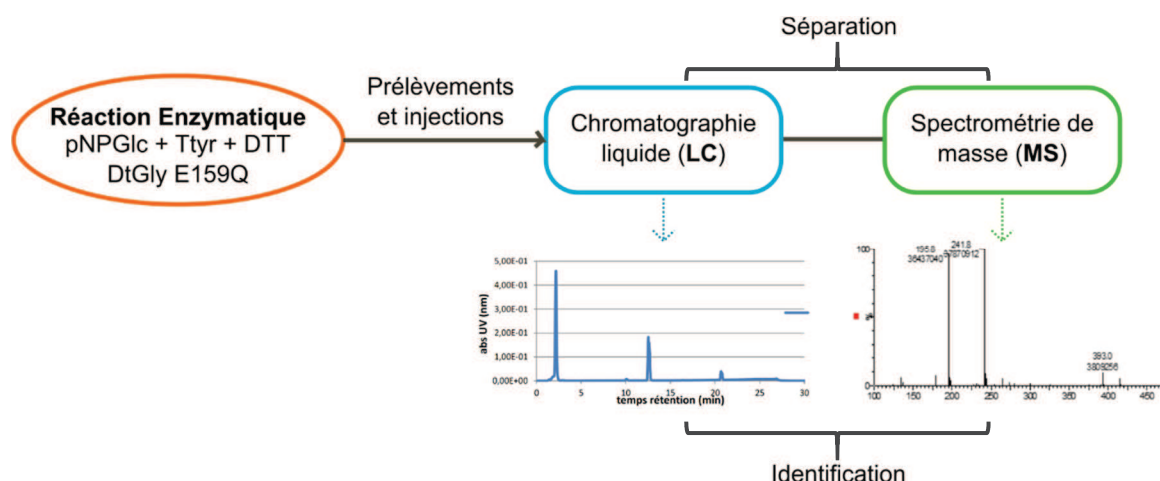
Le pic d'absorbance du thiolate Ttyr<sub>S-</sub> ( $\lambda_{\text{max}} = 273$  nm) est observé dans les 2 solutions après 20 min et 45 min d'incubation à pH basique. En revanche dans le cas du  $\beta$ -mercaptoéthanol la formation de l'espèce thiolate semble ne plus être assurée après 18 h d'incubation à 37 °C.

Le pic d'absorbance mesuré à  $\lambda_{\max} = 243$  nm est caractéristique du dimère de Ttyr à pH 8 comme décrit précédemment. Pour des temps d'incubation trop longs le  $\beta$ ME, composé volatil, n'est plus présent en solution. En absence d'agent réducteur, la formation de la liaison disulfure est de nouveau possible. En présence de DTT la forme déprotonée Ttyr<sub>S-</sub> est observée même après 18h de réaction, cet agent réducteur plus stable en solution a donc été choisi pour les essais de thioglycosylation de la Thiotyrosine.

### III. Couplage glycosidique de la Thiotyrosine catalysé par DtGly E159Q

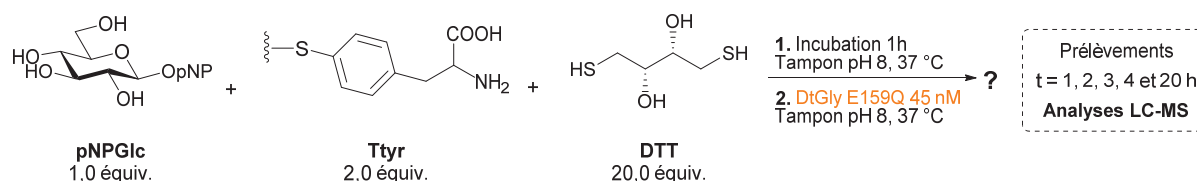
#### 1. Optimisation des conditions et de suivi de la réaction enzymatique

A partir de l'ensemble des données collectées précédemment les essais de glycosylation de la Thiotyrosine catalysés par DtGly E159Q ont pu être envisagés. Pour assurer la présence de l'espèce monomérique et thiolate de Ttyr la réaction doit se tenir à pH basique (pH  $\geq$  8) en présence de DTT (ratio 1:10 accepteur / agent réducteur). Enfin, une période d'incubation ( $\approx$  1h) du mélange donneur / accepteur / agent réducteur devra être respectée pour assurer la conversion du dimère Ttyr<sub>S-S</sub> en forme Ttyr<sub>S-</sub> avant d'initier la réaction de thioglycosylation par ajout du biocatalyseur. D'autre part, compte-tenu de la faible quantité disponible de la forme ester de l'acide aminé, le volume réactionnel des échantillons a été limité à 1 ml. Enfin le suivi de l'avancement de la réaction est réalisé par LC-MS à partir de prélèvements réalisés à différents temps d'incubation. La méthode d'analyse, sensible, permet en effet de suivre la disparition des produits de départ et l'apparition même en faible quantité du produit de thioglycosylation, tout trois UV-actifs. Dans un premier temps le passage en chromatographie en phase liquide (LC) permet de séparer les différents composés de l'échantillon en fonction de l'affinité phase stationnaire / phase mobile. Par la suite, le couplage avec la spectrométrie de masse (MS) permet d'identifier chacun de ces composés en fonction des ions de masse  $m/z$  détectés (Schéma 65). L'ionisation d'une molécule est dépendante du type de fonction(s) qu'elle présente : par exemple une molécule portant un groupement acide carboxylique (facilement déprotonable) aura une meilleure ionisation en mode négatif tandis qu'une molécule portant une fonction amine aura tendance à capter un proton et donc donnera une meilleure réponse en mode positif. Pour anticiper ce phénomène et assurer une bonne détection des différents composés du milieu réactionnel, les analyses en spectrométrie de masse ont été réalisées selon les deux modes d'ionisation, ESI (+) et ESI (-).



**Schéma 65** : Suivi de la réaction de thioglycosylation enzymatique par LC-MS.

Pour optimiser les conditions de détection et de séparation des différents constituants du milieu réactionnel un premier essai de thioglycosylation enzymatique a été réalisé à partir de la Ttyr non protégée disponible en plus grande quantité que la TtyrME (i.e., 4 mg). Les conditions de réaction mettent en jeu un rapport donneur / accepteur / réducteur de 1:2:20 (soit 1 mM de pNPGlc / 2 mM de Ttyr / 20 mM de DTT) dans un tampon phosphate 20 mM pH 8. Ces quantités ont été introduites à partir de solutions mère de pNPGlc (10 mM dans H<sub>2</sub>O), Ttyr (40 mM dans MeOH) et DTT (100 mM dans H<sub>2</sub>O). Le milieu réactionnel est laissé incuber 1h à 37 °C et sous agitation avant l'ajout de DtGly E159Q 45 nM (Schéma 66).



**Schéma 66** : Essai de thioglycosylation de la Thiotyrosine.

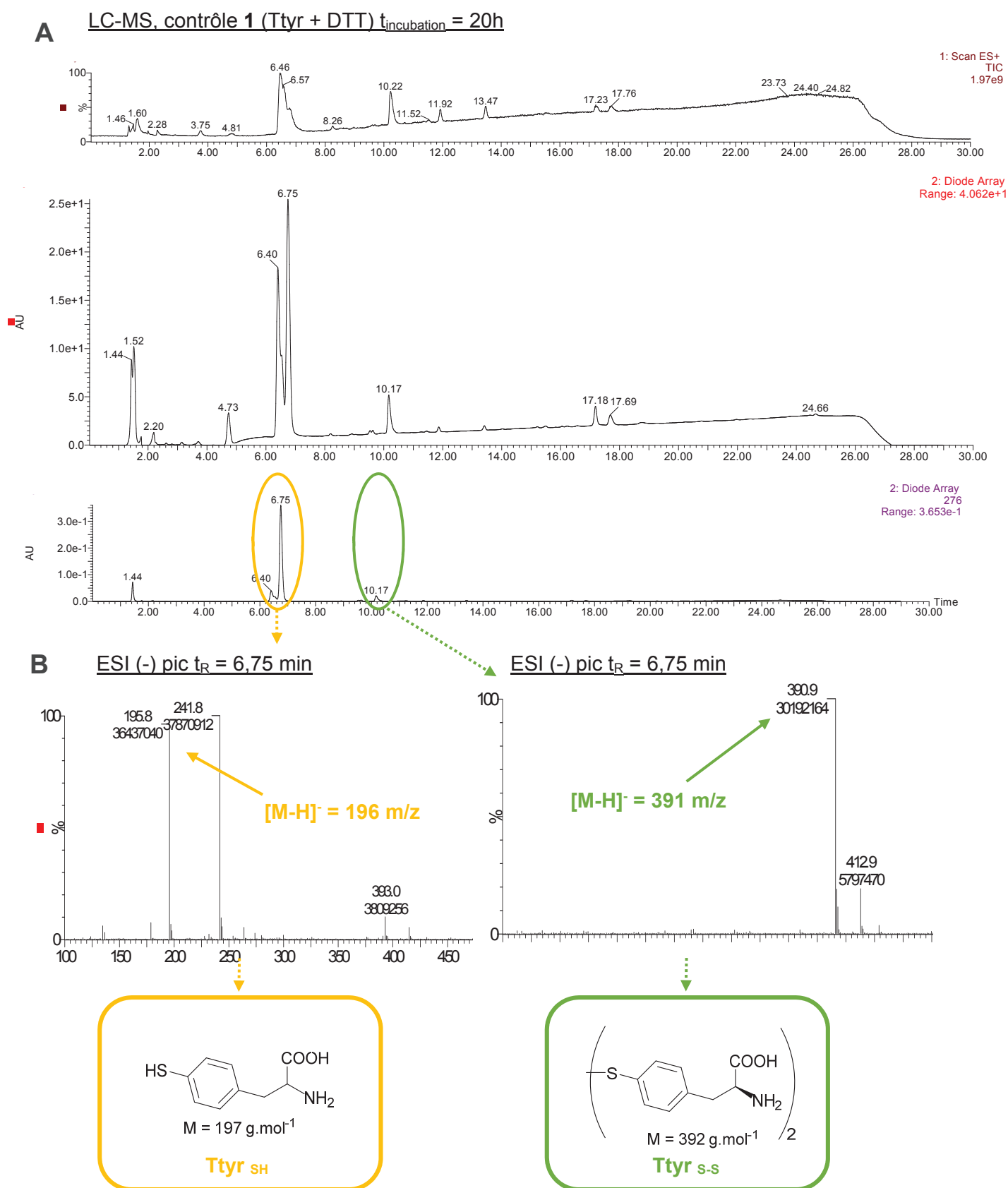
La réaction est conduite à 37 °C pendant 20h et différents prélèvements sont effectués à t = 1h, 2h, 3h, 4h et 20h, pour stopper la réaction enzymatique une solution d'arrêt est ajoutée (mélange ACN/ acide formique, 1:10) à chaque prélèvement (ratio 1:1). Après filtration les échantillons sont injectés et analysés sur la chaîne LC-MS. Deux autres échantillons « contrôle » ont également été préparés dans les mêmes conditions (Tableau 23). Après 20h d'incubation à 37 °C un prélèvement de chaque contrôle est effectué et analysé en LC-MS après ajout de la solution d'arrêt et filtration. L'objectif de l'injection en LC-MS de ces échantillons contrôle est d'obtenir les temps de rétention de chaque constituant de la réaction enzymatique (pNPGlc, Ttyr et DTT). Ils pourront alors être facilement identifiés sur les chromatogrammes enregistrés pour les essais de thioglycosylation.



Contrôle 1	Contrôle 2
Ttyr / DTT (1 :10)	pNPGlc (1 mM)
Tampon phosphate 20 mM pH 8	Tampon phosphate 20 mM pH 8

**Tableau 23** : Composition des solutions contrôle 1 et 2.

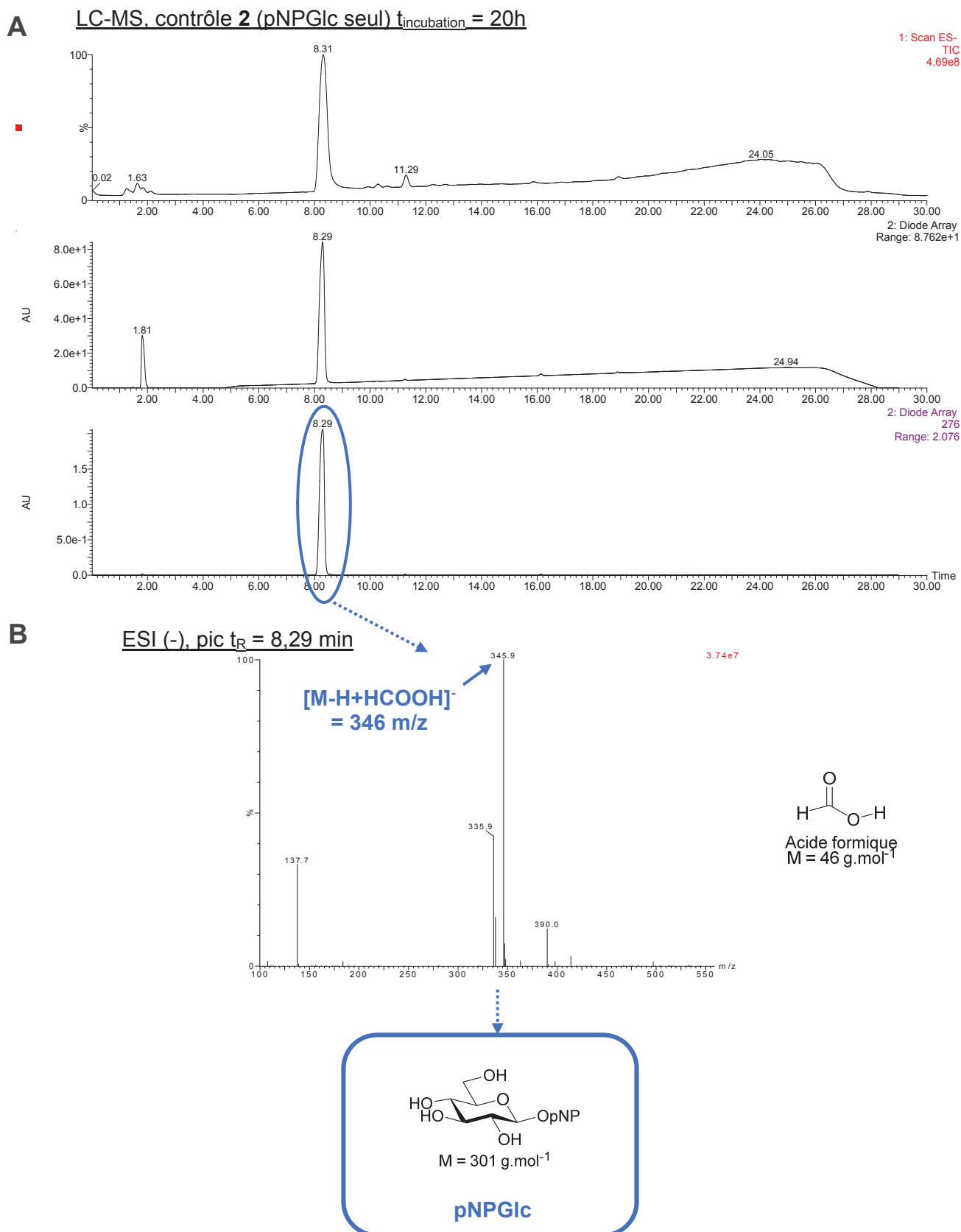
Dans le cas du contrôle 1 Ttyr / DTT, après 20h d'incubation à 37 °C la combinaison des chromatogrammes de masse et UV enregistrés permettent de mettre en évidence 7 pics. L'absorbance maximale de Ttyr est atteinte à  $\lambda = 273 - 276$  nm, pour identifier les pics correspondant à la Thiotyrosine, un second chromatogramme UV a été extrait à  $\lambda = 276$  nm et a permis de discriminer trois pics de temps de rétention  $t_R = 1,44$  ; 6,75 et 10,17 dont le premier pic à  $t_R = 1,44$  correspond au pic d'injection (**A**, Figure 74).



**Figure 74 :** Analyse LC-MS du contrôle 1, identification de la Thiotyrosine. (A) Chromatogrammes de Masse (TIC) et UV (Diode array) du contrôle 1 ; (B) Extraction des spectres de masse des pics à  $t_R = 6,75$  et  $10,17\text{ min}$ .

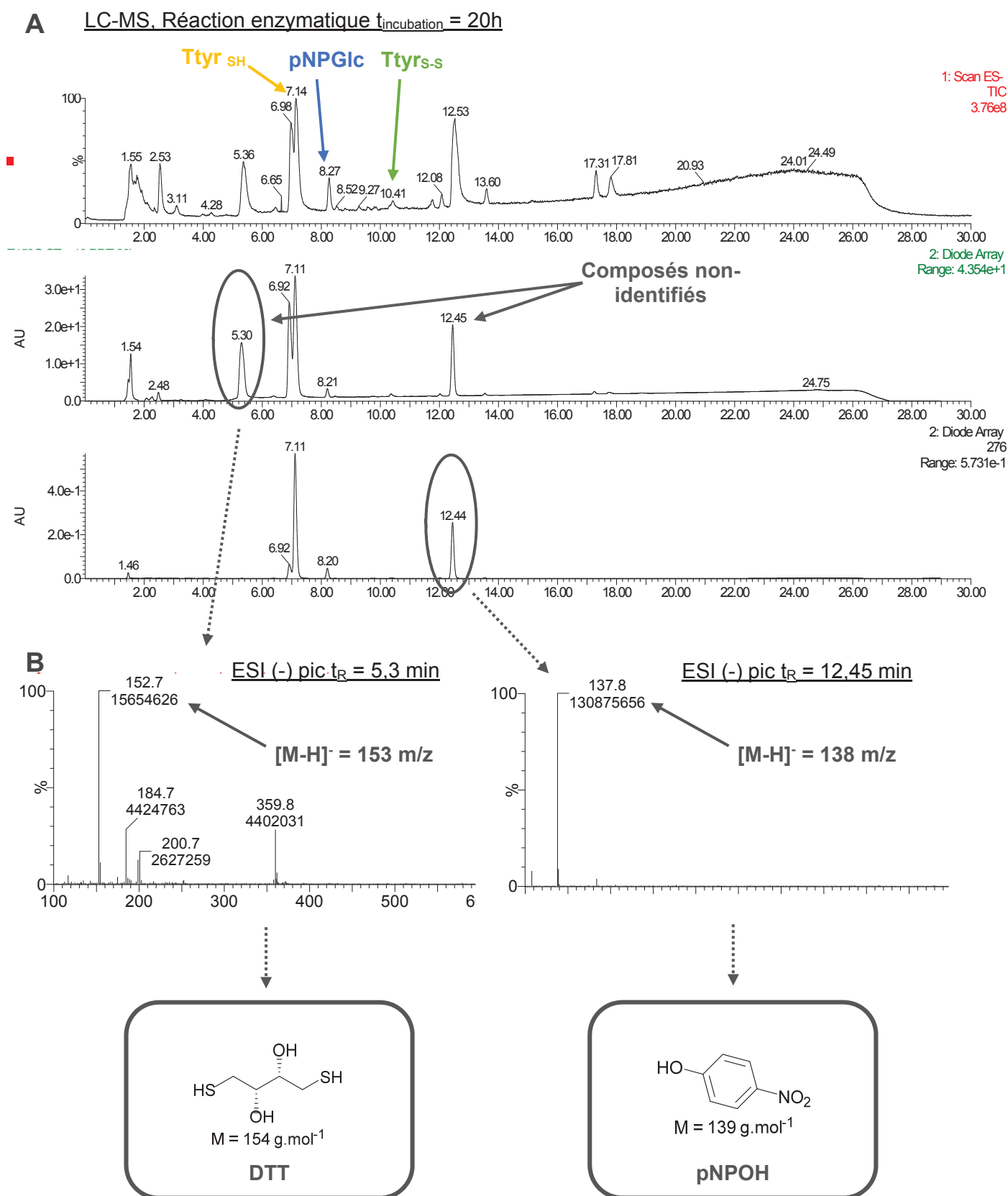
L'extraction des spectres de masse des signaux sortant à  $t_R = 6,75$  et  $10,2$  min permettent d'identifier des pics de masse  $[M-H]^- = 196$  m/z et  $[M-H]^- = 391$  m/z respectivement (**B**, Figure 74). Le poids moléculaire du monomère de Ttyr est de  $197$  g.mol<sup>-1</sup> tandis que le dimère Ttyr<sub>s</sub> possède une masse de  $392$  g.mol<sup>-1</sup>. Ainsi dans le contrôle **1**, le signal caractérisé par un temps de rétention  $t_R = 6,75$  min a pu être attribué à la Thiotyrosine, tandis que le second signal de plus faible intensité ( $t_R = 10,2$  min) correspond au dimère de la Thiotyrosine. L'analyse de cet échantillon par LC-MS permet également de mettre en évidence la stabilité de Ttyr en solution après 20h d'incubation à  $37$  °C et le pouvoir réducteur du DTT qui limite la formation du dimère de l'AAn comme décrit précédemment.

La même stratégie a été suivie pour analyser l'échantillon contrôle **2** contenant  $1$  mM de pNPGlc. Après 20h d'incubation à  $37$  °C, les chromatogrammes de Masse (TIC) UV ( $\lambda = 276$  nm) obtenus ont révélé un seul signal majoritaire de temps de rétention  $t_R = 8,3$  min (**A**, Figure 75). L'extraction du spectre de masse correspondant présente un pic majoritaire de masse  $[M-H]^- = 346$  m/z supérieure à celle attendue ( $[M-H]^- = 300$  m/z, pNPGlc) (**B**, Figure 75). Cette augmentation de masse est due à l'acide formique ( $M = 46$  g.mol<sup>-1</sup>), présent dans l'éluant de séparation et la solution d'arrêt. En chromatographie phase liquide cette molécule est reconnue pour générer facilement des adduits avec les molécules passant sur la colonne de séparation. Ainsi selon notre méthode de séparation, le pNPGlc est identifié avec un temps de rétention de  $8,3$  min et le spectre de masse correspondant est caractérisé par le signal  $[M-H+HCO_2H]^- = 346$  m/z.



**Figure 75 :** Analyse LC-MS du contrôle 2, identification du pNPGlc. (A) Chromatogrammes de Masse (TIC) et UV (Diode array) du contrôle 2 ; (B) Extraction du spectre de masse du pic à  $t_R = 8,29\text{ min}$ .

Les temps de rétention du pNPGlc et de Ttyr connus, l'analyse des chromatogrammes de l'essai de thioglycosylation en présence de DtGly E159Q a pu être réalisée. Pour rappel, la réaction a été initiée après 1h d'incubation à 37 °C par l'ajout de DtGly E159Q (45 nM). Des prélèvements ont alors été réalisés après 1, 2, 3, 4 et 20h d'incubation. Visuellement une faible coloration jaune du milieu réactionnel est observée après 20h d'incubation. Compte-tenu des résultats précédemment décrit sur l'activité thioligase de DtGly (cf. Chapitre 2, p.123) une faible coloration jaune est le résultat de l'activité d'hydrolyse résiduelle de l'enzyme. Les chromatogrammes enregistrés pour les différents échantillons se sont tous révélés similaires. Afin de proposer une comparaison homogène avec les contrôles **1** et **2**, seul le résultat LC-MS de l'échantillon  $t_{\text{incubation}} = 20\text{h}$  est présenté ici. Ainsi sur les chromatogrammes de masse et UV du  $t_{\text{incubation}} = 20\text{h}$  les signaux de  $t_R = 6,7$  ;  $8,3$  et  $10,4$  min ont pu rapidement être identifiés à la Ttyr, au pNPGlc et au dimère TtyrS-S respectivement (**A**, Figure 76). En revanche, deux nouveaux signaux sont observés à  $t_R = 5,3$  min et  $12,53$  min. Sur le chromatogramme UV ( $\lambda = 276$  nm) seul le signal à  $t_R = 12,5$  min est conservé, il s'agit donc d'un composé aromatique, à l'inverse du composé caractérisé par  $t_R = 5,3$  min dont on peut observer l'extinction du signal. L'extraction des spectres de masse de ces 2 pics a permis d'attribuer le signal à  $t_R = 5,3$  min à l'agent réducteur DTT ( $M = 154 \text{ g.mol}^{-1}$ ) avec  $[M-H]^- = 153 \text{ m/z}$ , tandis que le produit sortant de la colonne après  $12,5$  min a été identifié comme le pNPOH ( $M = 139 \text{ g.mol}^{-1}$ ) avec  $[M-H]^- = 138 \text{ m/z}$  (**B**, Figure 76). Parmi les autres pics restants l'extraction des spectres de masse correspondants n'a pas permis l'identification du produit de thioglycosylation ( $[M+H]^+ = 360 \text{ m/z}$ ). La thioglycosylation enzymatique n'a pas eu lieu comme attendu.

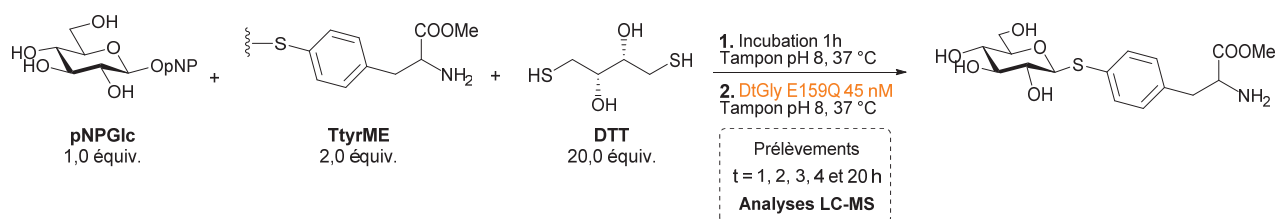


**Figure 76 :** Analyse LC-MS de la réaction enzymatique après 20h d'incubation. (A) Chromatogrammes de Masse (TIC) et UV (Diode array) ; (B) Extraction des spectres de masse des pics à  $t_R = 5,3$  et  $12,45 \text{ min}$ .

La méthode de séparation et de détection mise au point par LC-MS permet de séparer et d'identifier efficacement les différents réactifs impliqués dans la réaction de thioglycosylation catalysée par DtGly E159Q. D'autre part, en tenant compte de la polarité du pNPGlc et de Ttyr ainsi que de leur temps de rétention il est évident que le produit de thioglycosylation ne pourra être attendu à un  $t_R > 9$  min. Grâce à l'ensemble des informations collectées, l'essai de réaction enzymatique en présence de la forme protégée de la Thiotyrosine **31** (TtyrME) a alors pu être réalisé.

## 2. Thioglycosylation enzymatique de TtyrME catalysée par DtGly E159Q

En respectant les conditions décrites précédemment, la réaction de thioglycosylation enzymatique de la Thiotyrosine estérifiée a été effectuée dans un rapport donneur / accepteur / réducteur de 1:2:20 (soit 1 mM de pNPGlc / 2 mM de TtyrME / 20 mM de DTT) dans un tampon phosphate 20 mM pH 8. Ces quantités ont été introduites à partir de solutions mère de pNPGlc (10 mM dans H<sub>2</sub>O), TtyrME (40 mM dans MeOH) et DTT (100 mM dans H<sub>2</sub>O). Après 1h d'incubation à 37 °C la réaction est initiée par l'ajout de DtGly E159Q 45 nM (Schéma 67).



**Schéma 67** : Essai de thioglycosylation enzymatique de la Thiotyrosine estérifiée.

Des prélèvements sont effectués après  $t = 1$  h, 2 h, 4 h et 20 h d'incubation à 37 °C sous agitation. Après ajout de la solution d'arrêt (ratio 1:1) et filtration, les échantillons sont analysés en LC-MS. De la même manière trois échantillons « contrôle » ont été préparés dans les mêmes conditions (Tableau 24). Dans chacun des cas un prélèvement a été effectué après 20h d'incubation à 37 °C et l'ensemble a été analysé après filtration.

Contrôle 1	Contrôle 2	Contrôle 3
TtyrME 2 mM, Tampon phosphate 20 mM pH8	pNPGlc 1 mM Tampon phosphate 20 mM pH8	pNPGlc 1 mM TtyrME 2 mM, DTT 20 mM Tampon phosphate 20 mM pH8

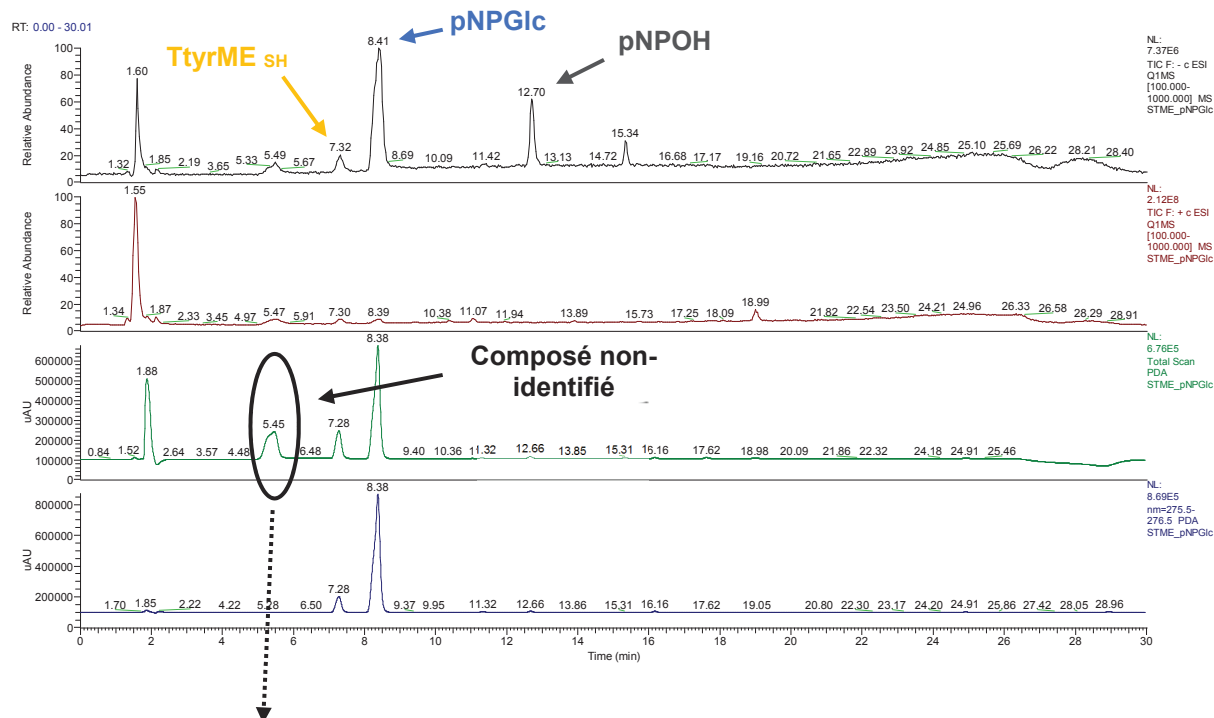
**Tableau 24** : Composition des solutions contrôle de la réaction de thioglycosylation par DtGly E159Q.

En suivant la même stratégie d'analyse que celle décrite précédemment, les chromatogrammes de masse et UV du contrôle **3** ont pu être caractérisés à partir de l'analyse des contrôles **1** et **2** (présentés dans la partie expérimentale p. 240). La première observation

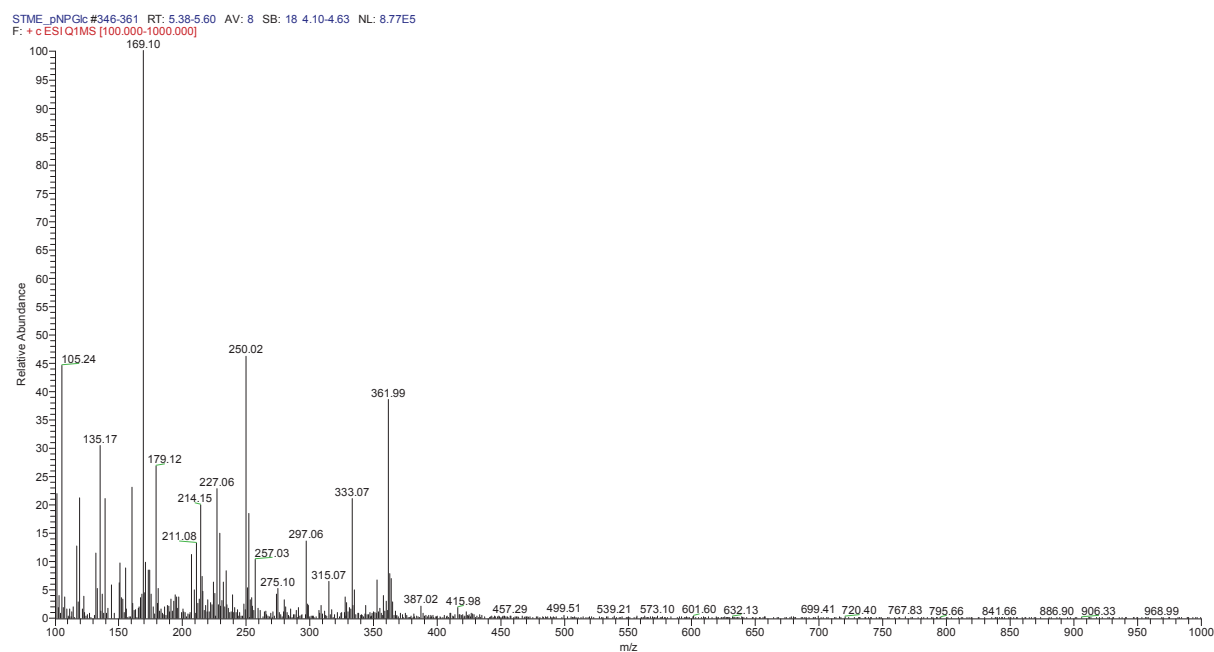
qui peut être faite à la vue des chromatogrammes de masse est que l'ionisation des molécules présentes dans l'échantillon est fortement influencée par le mode utilisé, ESI (+) ou ESI (-). En effet en mode négatif six pics peuvent être identifiés alors que le chromatogramme enregistré en mode positif reflète de l'extinction (ou faible intensité) de la majorité des signaux à l'exception du signal à  $t_R = 1,6$  min alors attribué au pic d'injection. D'autre part, les trois pics caractérisés par un  $t_R = 7,32$  ;  $8,41$  et  $12,7$  min ont pu être attribués à la Thiotyrosine estérifiée (TtyrME), au pNPGlc ainsi qu'au produit d'hydrolyse résiduel pNPOH respectivement (**A**, Figure 77). Malgré les 2 équivalents de TtyrME ajoutés dans la solution, le signal de l'acide aminé est d'une intensité beaucoup plus faible que celui du pNPGlc. Cette différence d'intensité n'a pas été observée précédemment (i.e., essai pNPGlc / Ttyr / DTT) et semble être due à une mauvaise homogénéisation de la solution avant d'effectuer le prélèvement. Les chromatogrammes UV rendent compte de la même observation. Le signal de la Thiotyrosine, qui présente un maximum d'absorbance à  $\lambda = 276$  nm, est de plus faible intensité que celui du pNPGlc (**B**, Figure 77). De plus leur analyse révèle la présence d'un 4<sup>ème</sup> pic non identifié à  $t_R = 5,45$  min qui n'absorbe pas à  $\lambda = 276$  nm (extinction du signal sur le chromatogramme UV  $\lambda = 276$  nm). Le spectre de masse correspondant est constitué d'une multitude de pics dont aucun n'est caractérisé par une masse correspondant aux produits de départ. Enfin, l'absence de pic de masse  $[M+H]^+ = 374$  m/z permet de montrer que la réaction de thioglycosylation spontanée entre le pNPGlc et TtyrME n'est pas possible.



**A** LC-MS, contrôle 3  $t_{\text{incubation}} = 20\text{h}$

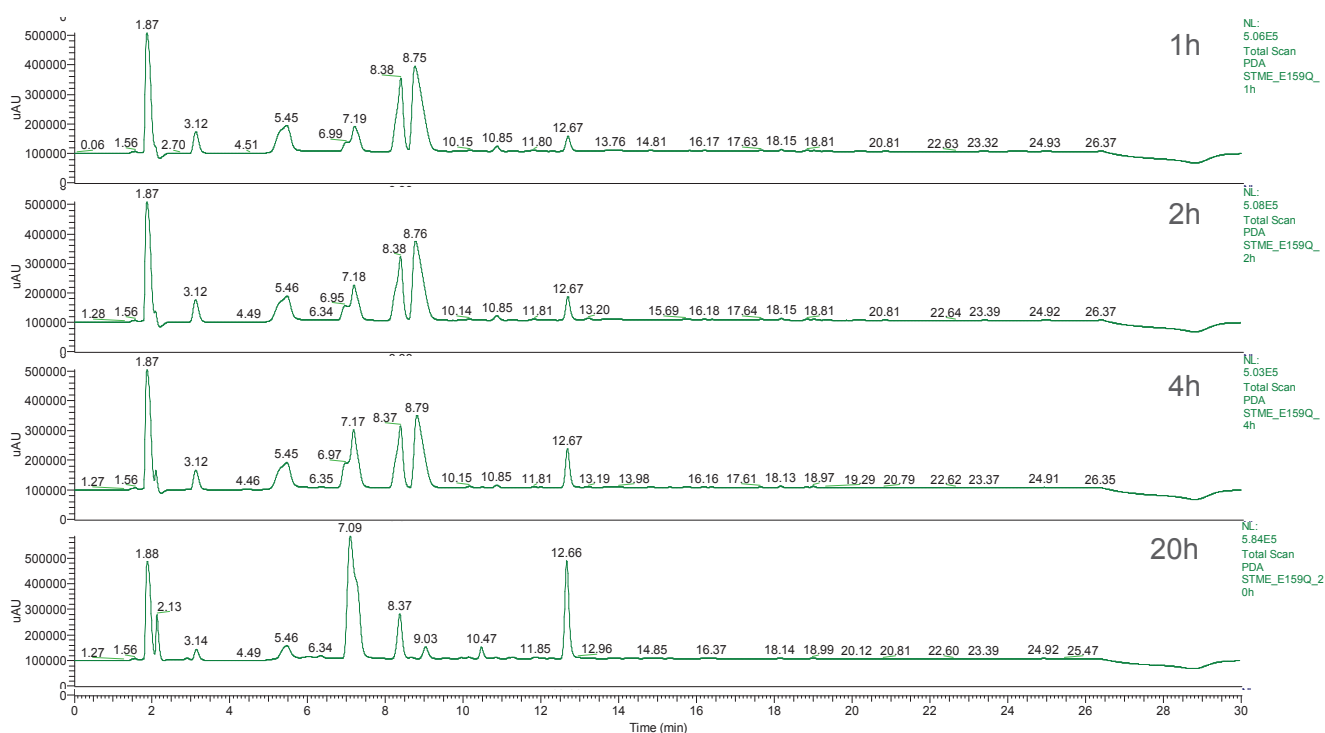


**B** ESI (+) pic  $t_R = 5.45\text{ min}$



**Figure 77** : Analyse LC-MS du contrôle 3 après 20 h d'incubation. **(A)** Chromatogrammes de Masse (TIC) et UV (Diode array) du contrôle 3 ; **(B)** Extraction du spectre de masse du signal à  $t_R = 5,45\text{ min}$ .

A partir de ces résultats préliminaires, le suivi de la réaction de thioglycosylation enzymatique a pu être réalisé. Le milieu réactionnel composé du pNPGlc, de TtyrME et du DTT est donc laissé incuber 1h à 37 °C avant l'ajout de DtGly E159Q (45 nM). Quelques minutes après l'ajout du biocatalyseur, une vive coloration jaune du milieu réactionnel est observée. Compte-tenu de l'étude sur l'activité thioligase de DtGly E159Q décrite dans le chapitre 2 de ce manuscrit (p. 123), la coloration rapide de la réaction est caractéristique de la présence d'un bon accepteur capable d'augmenter le turn-over de l'enzyme et donc de favoriser le couplage glycosidique. Pour confirmer ce fait les résultats de LC-MS des prélèvements effectués à  $t_{\text{incubation}} = 1, 2, 4$  et 20h ont été analysés. Au vu des profils UV de chaque échantillon, un premier constat peut être établi, la réaction enzymatique ne semble pas ou peu évoluer au cours du temps (Figure 78) Ainsi pour faciliter la lecture des résultats et proposer une comparaison homogène avec le contrôle **3**, seule l'analyse de l'échantillon  $t_{\text{incubation}} = 20\text{h}$  est présentée pour l'attribution des signaux observés. Cette dernière peut alors être utilisée comme modèle pour les autres prélèvements de  $t_{\text{incubation}} = 1, 2, 4\text{h}$  (disponibles dans la partie expérimentale, p. 241 et 242).

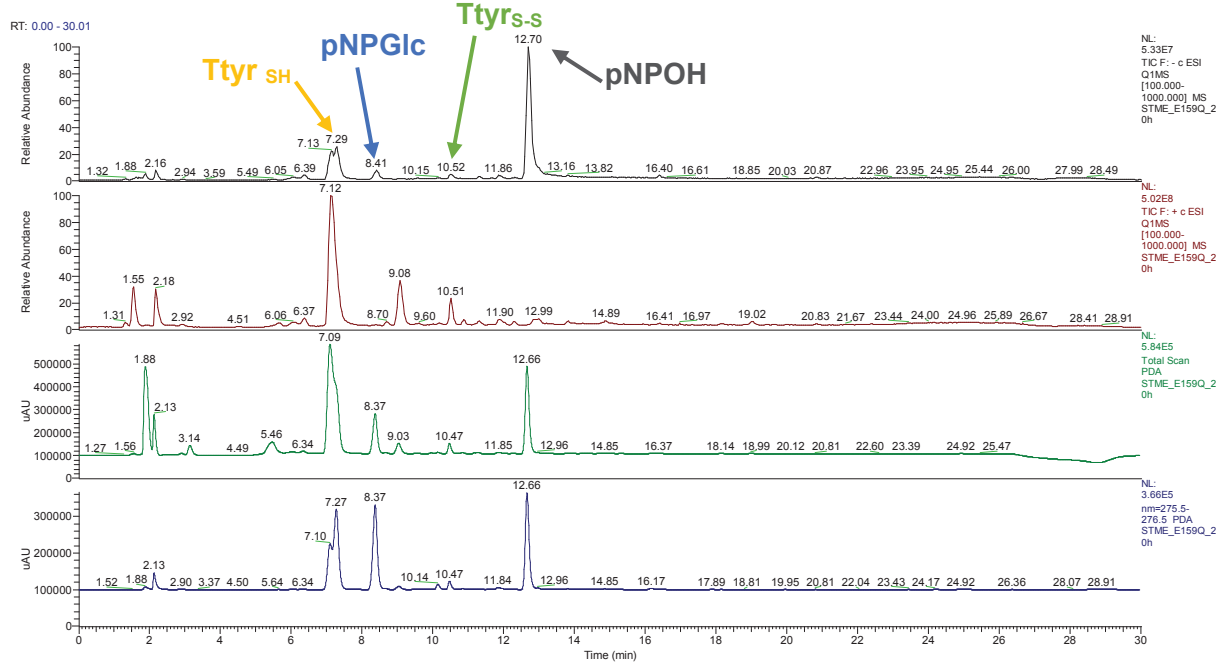


**Figure 78 :** Chromatogrammes UV enregistrés pour les prélèvements effectués à  $t = 1, 2, 4$  et 20h d'incubation à 37 °C.

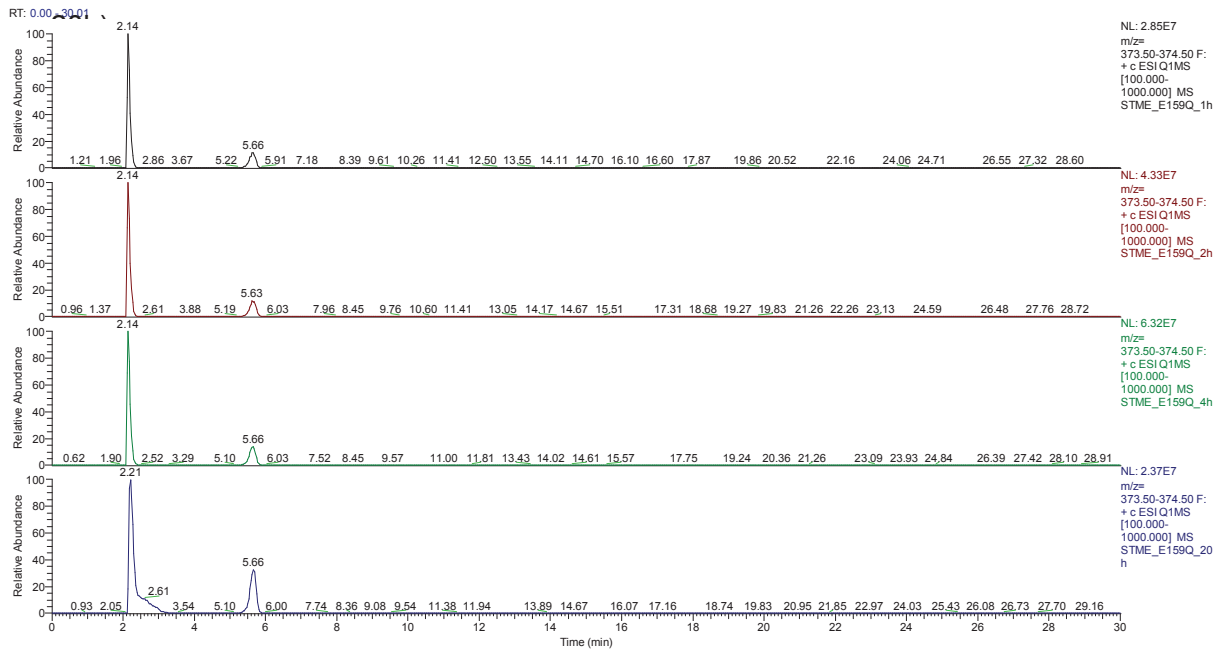
L'observation des chromatogrammes de masse enregistrés en mode ESI (+) et ESI (-) permet de mettre en évidence 7 pics parmi lesquels ont été identifiés : le pic d'injection ( $t_R = 1,55$  min), TtyrME ( $t_R = 7,29$  min), le pNPGlc ( $t_R = 8,41$  min), le dimère de TtyrME ( $t_R = 10,51$  min) et le pNPOH ( $t_R = 12,7$  min) (**A**, Figure 79). Les signaux de  $t_R = 2,1$  et 9,1 min sont également rencontrés sur les chromatogrammes UV ainsi que le signal à  $t_R = 5,45$  min précédemment décrit.

Afin de déterminer si l'un de ces pics pouvait être attribué au produit de thioglycosylation ( $M = 373 \text{ g.mol}^{-1}$ ), une recherche d'ions a été réalisée à partir de l'ensemble des chromatogrammes ESI (+) et ESI (-) enregistrés pour les prélèvements de 1 à 20h d'incubation. L'objectif de cette extraction est donc de mettre en évidence un signal caractérisé par un ion de masse  $[M+H]^+ = 374 \text{ m/z}$  ou  $[M-H]^- = 372 \text{ m/z}$  et si tel est le cas de déterminer le temps minimum de réaction nécessaire pour qu'il se forme. En mode ESI (+) les deux signaux de  $t_R = 2,14$  et  $5,66 \text{ min}$  ont pu être isolés et ce sur le chromatogramme de chaque échantillon (**B**, Figure 79).

**A** LC-MS, Réaction enzymatique  $t_{\text{incubation}} = 20\text{h}$



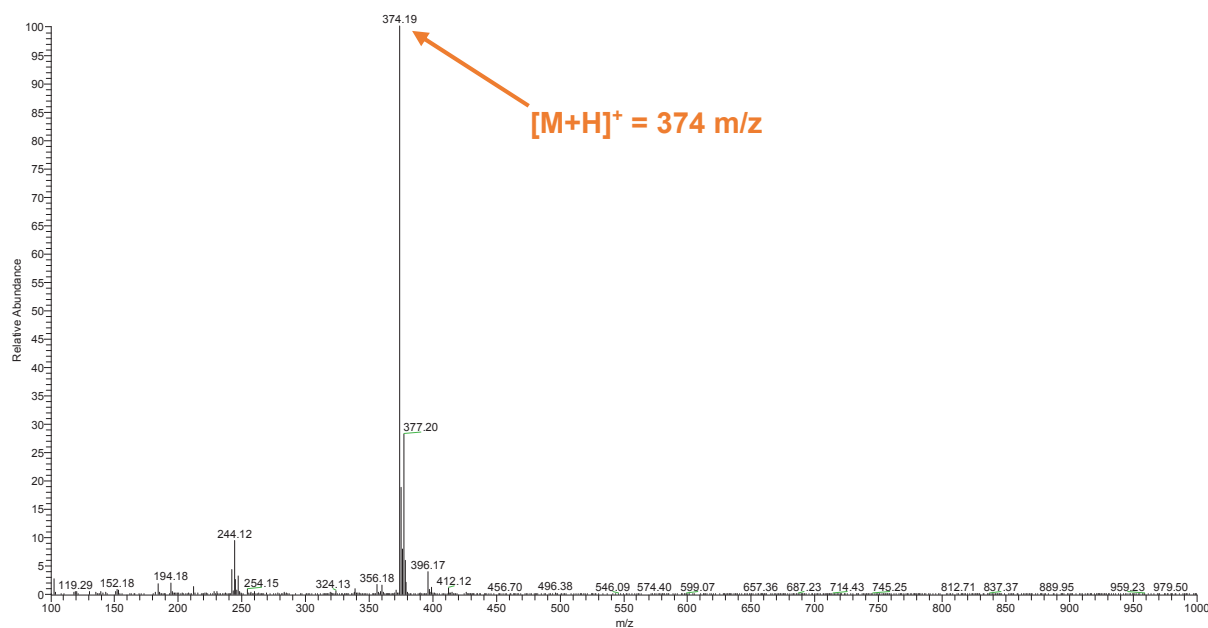
**B** Extraction d'ions de la masse  $[M+H]^+ = 374\text{ m/z}$  ( $t_{\text{incubation}} = 1, 2, 4\text{ et}$



**Figure 79** : Analyse LC-MS de la réaction enzymatique après 20h d'incubation. **(A)** Chromatogrammes de Masse (TIC) et UV (Diode array) ; **(B)** Recherche d'ions de masse  $[M+H]^+ = 374\text{ m/z}$  après 1, 2, 4 et 20h d'incubation.

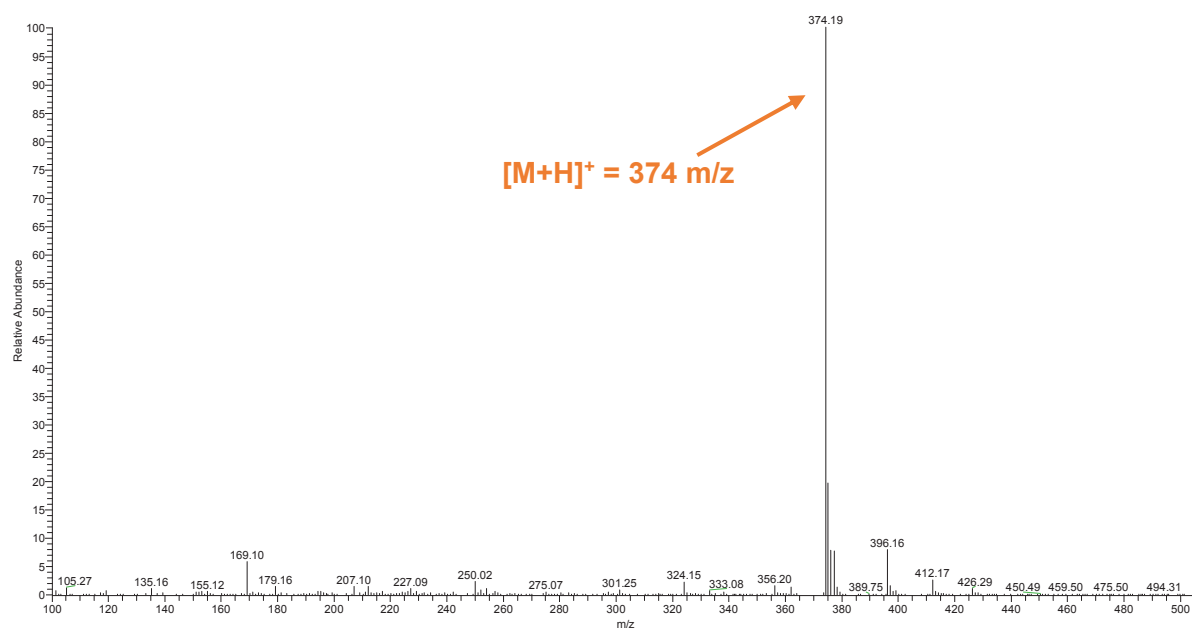
Les spectres de masse extraits pour ces 2 signaux confirment la présence d'un pic majoritaire de masse  $[M+H]^+ = 374 \text{ m/z}$ , caractéristique de la glycosylation de la Thiotyrosine estérifiée, et prouvent par conséquent la réalisation du couplage enzymatique (Figure 80). Néanmoins par cette analyse il n'est pas possible de dire à quel pic correspond l'acide aminé glycosylé. En effet, l'essai de thioglycosylation réalisé à partir d'une faible quantité de produit de départ n'a pas permis d'isoler les 2 fractions correspondants de  $t_R = 2,14$  et  $5,66$  min et donc leur caractérisation par RMN.

**A** ESI (+) pic  $t_R = 2,12$  min



**B**

ESI (+) pic  $t_R = 5,66$  min

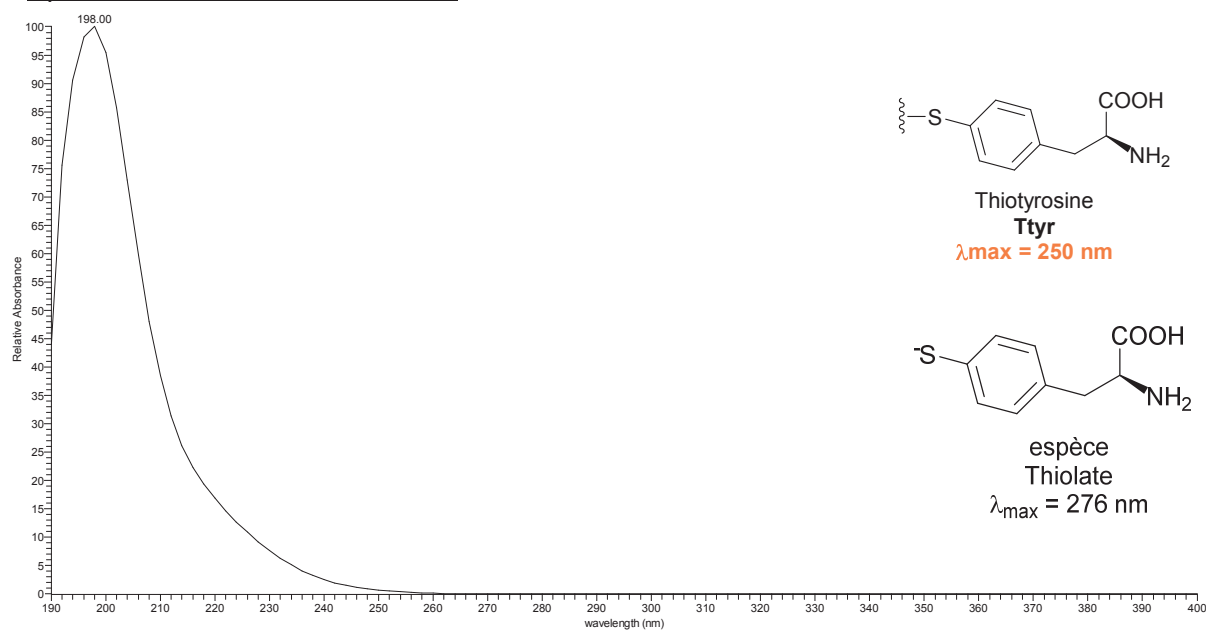


**Figure 80** : Spectre de masse ESI(+) des fractions  $t_R = 2,12$  min (**A**) et  $t_R = 5,66$  min (**B**).

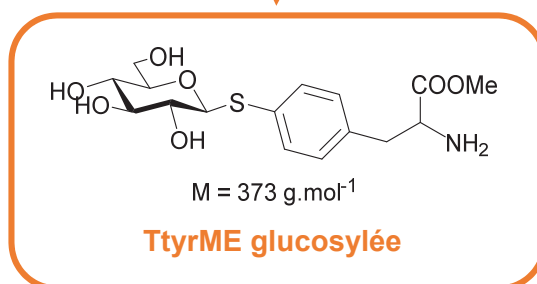
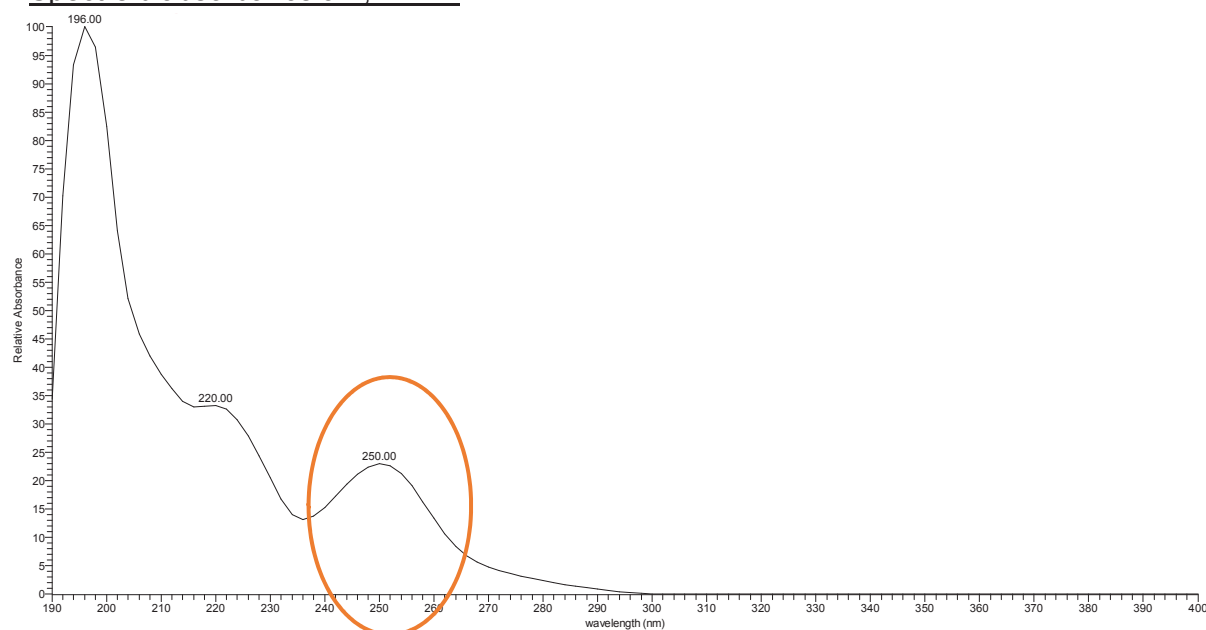
Malgré tout, pour tenter de les différencier un spectre d'absorbance de chaque signal a été enregistré et comparé avec celui de Ttyr précédemment obtenu (p.161). Pour rappel, en solution la Thiotyrosine est caractérisée par un pic d'absorbance à  $\lambda = 250$  nm et à  $\lambda = 276$  nm quand l'espèce thiolate est en solution. Le profil UV-visible de la fraction  $t_R = 2,14$  min présente trois pics d'absorbance à  $\lambda = 196$ , 220 et 250 nm tandis que la fraction  $t_R = 5,66$  min n'est caractérisée que par un seul pic d'absorbance à  $\lambda = 196$  nm (Figure 81). La réaction n'ayant pas été conduite sur une quantité suffisante pour permettre l'isolation et la caractérisation de chaque fraction (fractions  $t_R = 2,14$  et 5,66 min) ce résultat préliminaire semble donc être un faisceau de preuve que la réaction de thioglycosylation enzymatique de la Thiotyrosine a bien eu lieu. Pour confirmer la réalisation de ce couplage, plusieurs essais devront être effectués en augmentant :

- i) la concentration en biocatalyseur ; les chromatogrammes identiques pour les différents échantillons montre que la réaction n'est pas totale et ne semble plus évoluer après 1h d'incubation. En présence d'une plus grande quantité de DtGly E159Q, si la réaction a lieu une nette diminution du pic de la Thiotyrosine devrait être observée en faveur de celui correspondant au produit ;
- ii) la quantité en réactif ; ceci permettrait de purifier les produits de la réaction enzymatique et de les identifier par RMN et HRMS.

**A** Spectre d'absorbance à 5,66 min



**B** Spectre d'absorbance à 2,14 min



**Figure 81** : Spectres d'absorbance du signal à  $t_R = 5,66 \text{ min}$  (**A**) et  $t_R = 2,14 \text{ min}$  (**B**).

## IV. Conclusion

L'objectif de ce dernier chapitre était de montrer que le couplage glycosidique d'un acide aminé était possible à partir d'une thioligase. Pour ce faire, à partir de l'étude de relation structure/activité de notre enzyme modèle DtGly E159Q, nous nous sommes orientés vers la synthèse d'un acide aminé aromatique non-naturel présentant une fonction thiol de faible pKa et dont la fonction acide carboxylique est masquée sous la forme d'un ester. Ainsi la Thiotyrosine (Ttyr) **30** a pu être obtenue au terme d'une synthèse en trois étapes avec un rendement global de 59 % tandis que la forme estérifiée TtyrME **31** a été produite en 4 étapes avec un faible rendement final d'1 %. L'analyse RMN a permis de confirmer la structure des deux composés tandis que l'analyse HRMS a mis en évidence leur organisation sous la forme de dimère en solution. Pour anticiper les essais de thioglycosylation enzymatique une étude a été réalisée à partir de Ttyr pour déterminer les conditions de réaction nécessaires pour assurer la présence de la forme thiolate, l'espèce réactive, en solution. Ainsi en milieu tamponné à pH 8 et contenant 10 équivalents de l'agent réducteur DTT (par rapport à l'accepteur) la forme déprotonée de Ttyr est assurée pendant 18h à 37 °C. Les spectres d'absorbance UV-visible enregistrés pour la Thiotyrosine et sa forme déprotonée ont révélé deux longueurs d'onde caractéristiques à  $\lambda_{\text{max}} = 250 \text{ nm}$  et  $\lambda_{\text{max}} = 273 \text{ nm}$  pour chaque espèce respectivement. Par la suite une méthode de suivi de la réaction de thioglycosylation par LC-MS a été mise au point à partir d'une réaction modèle en présence de pNPGlc comme donneur activé et de Ttyr pour jouer le rôle de l'accepteur de glycoside. La méthode a été validée et permet la séparation et la détection des différents acteurs de la réaction enzymatique. D'autre part nous avons pu mettre en évidence qu'aucun produit de thioglycosylation n'avait été formé par catalyse enzymatique ou par réaction spontanée entre le donneur et l'accepteur de glycoside.

Enfin l'essai de thioglycosylation de TtyrME catalysée par DtGly E159Q à 37 °C a été réalisé et le suivi par LC-MS a permis d'identifier l'apparition d'un nouveau signal avec un temps de rétention de 2,12 min. Le spectre de masse correspondant a présenté un ion de masse attendue  $([M+H]^+ = 374 \text{ m/z})$  et le profil UV-visible enregistré est comparable à celui de la Thiotyrosine avec un pic d'absorbance à  $\lambda = 250 \text{ nm}$ . Ce résultat préliminaire permettrait d'attribuer ce pic à la **TtyrME S-glycosylée** de masse  $M = 372 \text{ g.mol}^{-1}$  néanmoins seule une caractérisation complète de la molécule permettrait d'affirmer que la réaction enzymatique a bien eu lieu. D'autre part, l'analyse des différents prélèvements effectués à  $t = 1, 2, 4$  et 20h d'incubation a révélé que la réaction était non totale et stoppée après 1h d'incubation. Pour pallier à ces deux inconvénients, des améliorations sont à envisager et permettront alors de confirmer ces résultats prometteurs de thioglycosylation enzymatique d'un acide aminé :

- optimiser la synthèse de la Thiotyrosine pour permettre sa production en grande quantité ou envisager une nouvelle voie d'accès qui faciliterait sa purification, notamment en l'intégrant dans une courte séquence peptidique ;
- optimiser le couplage enzymatique en augmentant la concentration en enzyme pour favoriser une réaction totale comme précédemment décrits pour la synthèse des thioglycoconjugués par DtGly E159Q.



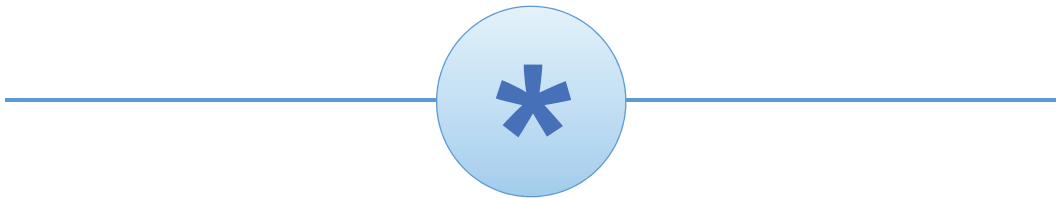


## Bibliographie

- Forbes, C. R., and Zondlo, N. J.; *Org. Lett.* **(2012)**, 14, 464-467
- Gamblin, D. P., Scanlan, E. M., and Davis, B. G.; *Chem. Rev.* **(2009)**, 109, 131-163
- Heidecke, C. D., Ling, Z., Bruce, N. C., Moir, J. W., Parsons, T. B., and Fairbanks, A. J.; *Chembiochem* **(2008)**, 9, 2045-2051
- Hescher, E., Bernier, M., and Parent, P.; *Helv. Chim. Acta* **(1983)**, 66, 1355-1365
- Koeller, K. M., Smith, M. E. B., Huang, R. F., and Wong, C. H.; *J. Am. Chem. Soc.* **(2000a)**, 122, 4142-4242
- Koeller, K. M., Smith, M. E. B., and Wong, C. H.; *J. Am. Chem. Soc.* **(2000b)**, 122, 752-743
- Li, B., Zeng, Y., Hauser, S., Song, H., and Wang, L. X.; *J. Am. Chem. Soc.* **(2005)**, 127, 9692-9693
- Lu, H. S. M., Volk, M., Kholodenko, Y., Gooding, E., Hochstrasser, R. M., and DeGrado, W. F.; *J. Am. Chem. Soc.* **(1997)**, 119, 7173-7180
- Meyer, F. M., Liras, S., Guzman-Perez, A., Perreault, C., Bian, J. W., and James, K.; *Org. Lett.* **(2010)**, 12, 3870-3873
- Rajagopalan, S., Radke, G., Evans, M., and Tomich, J. M.; *Synth. Commun.* **(1996)**, 26, 1431-1440
- Ramakrishnan, B., Boeggeman, E., and Qasba, P. K.; *Bioconjugate Chem.* **(2007)**, 18, 1912-1918
- Rudolf, J. D., and Poulter, C. D.; *ACS Chem. Biol.* **(2013)**, 8, 2707-2714
- Soria-Castro, S. M., and Penénory, A. B.; *Beilstein J. Org. Chem.* **(2013)**, 9, 467-475
- Takegawa, K., Yamaguchi, S., Kondo, A., Iwamoto, H., Nakoshi, M., Kato, I., and Iwahara, S.; *Biochem. Int.* **(1991)**, 24, 849-855
- Thayer, D. A., and Wong, C.-H.; *Top. Curr. Chem.* **(2006)**, 267, 37-63
- Varki, A.; *Glycobiology* **(1993)**, 3, 97-130
- Villalonga, M. L., Diez, P., Sanchez, A., Gamella, M., Pingarron, J. M., and Villalonga, R.; *Chem. Rev.* **(2014)**, 114, 4868-4917
- Wang, H., Oman, T. J., Zhang, R., Garcia De Gonzalo, C. V., Zhang, Q., and van der Donk, W. A.; *J. Am. Chem. Soc.* **(2014)**, 136, 84-87
- Yamamoto, K., Kadowaki, S., Watanabe, J., and Kumagai, H.; *Biochem. Biophys. Res. Commun.* **(1994)**, 203, 244-252



## **Conclusion Générale & Perspectives**





## Conclusion Générale & Perspectives

L'objet de ce travail de thèse a résidé dans le développement d'outils enzymatiques pour le couplage thioglycosidique d'acides aminés dans le but de proposer *in fine* une nouvelle voie d'accès aux protéines S-glycosylées. Dans la communauté scientifique, la glycosylation, modification la plus complexe rencontrée chez les protéines, suscite un vif intérêt tant les liens sucres – protéines sont étroits et influencent de nombreux processus biologiques. Une variété de méthodologies ont alors été développées afin de permettre la production de formes homogènes de ces glycoprotéines et permettre leur étude. Les formes S-glycosidiques ont trouvé un intérêt particulier dans ce domaine de par notamment la réactivité de l'atome de soufre qui facilite le couplage spécifique sucre – acide aminé mais également grâce à la résistance de ce type de liaison vis-à-vis de l'hydrolyse chimique et enzymatique.

La découverte récente de formes naturelles de S-glycoprotéines a permis l'identification et la caractérisation de S-glycosyltransférases responsables de leur biosynthèse. Cependant peu de données sont disponibles à l'heure actuelle sur ce type d'enzymes et les GT sont de manière générale difficile d'accès sous forme recombinantes. Nous nous sommes alors intéressés à la seconde catégorie d'enzymes capables de créer des liaisons thioglycosidiques, les thioglycoligases, qui ont prouvé leur potentiel en synthèse pour la préparation de thioglycoconjugués.

Dans un premier temps, nous nous sommes appliqués à sélectionner des enzymes natives qui constitueraient notre banque de biocatalyseurs. Notre choix s'est alors porté sur les GH de la bactérie thermophile *Dictyoglomus thermophilum*. A partir de son génome trois gènes ont pu être clonés (*dicth\_0289*, *dicth\_0359* et *dicth\_1692*) conduisant par la suite à la surexpression et la caractérisation des protéines thermophiles correspondantes : une  $\alpha$ -L-rhamnosidase (DtRha), une  $\beta$ -D-glycosidase (DtGly) et une  $\beta$ -D-mannosidase (DtMan). Parmi ces dernières, DtGly a retenu notre attention puisqu'elle a présenté une versatilité vis-à-vis du substrat en reconnaissant jusqu'à cinq motifs sucres différents (Glc, Gal, Fuc, Xyl et Man). L'ensemble des biocatalyseurs a été testé sur un panel d'alcools aliphatiques pour déterminer leur potentiel en synthèse. Les résultats obtenus ont permis de mettre en évidence l'activité transglycosidase de DtGly en catalysant la production de glycosides d'alkyle avec des rendements de 30 à 99 %. Dans un souci de valorisation, cette enzyme a été utilisée pour la mise au point d'une voie de synthèse chimio-enzymatique en deux étapes de glycéroglycosides. La détermination de leur activité biologique par microchallenge test, réalisé par l'entreprise GLYcoDiag, a permis de mettre en évidence les propriétés antifongiques du 1-O-( $\beta$ -D-glucopyranosyl)-(S)-glycérol (**14**), propriétés très recherchées à l'heure actuelle pour lutter contre les phénomènes de résistances des souches fongiques aux antibiotiques. Des essais de cristallogenèse ont été réalisés sur les trois glycosidases et ont permis l'obtention de cristaux dans le cas de DtRha et DtMan, dont l'analyse au Synchrotron SOLEIL a permis d'obtenir une image de diffraction de 3 Å pour DtRha et les données collectées pour DtMan nous ont permis d'obtenir un modèle préliminaire avec une résolution de 2,7 Å.

Dans un second temps, à partir de notre banque de GH sauvages, nous nous sommes tournés vers la production des thioglycoligases correspondantes par mutagenèse dirigée. Neuf mutants ont ainsi pu être obtenus, 2 variants dans le cas de DtGly et 7 pour la  $\beta$ Man DtMan, et caractérisés avec une perte générale de l'activité d'hydrolyse suite au remplacement du résidu acide base par un acide aminé catalytiquement inactif. L'activité thioligase du mutant DtGly E159Q a été confirmée sur un panel d'accepteurs de type thioarylique et constitue donc le 9<sup>ème</sup> exemple de thioligase connu à ce jour. L'étude des relations structure/activité de la protéine a démontré une perte d'activité vis-à-vis d'accepteurs portant des groupements fonctionnels polarisables. L'hypothèse d'une répulsion d'ordre électrostatique protéine/accepteur a alors été suggérée puisque la protection de ces groupes par une fonction non polarisable (ester de méthyle) a permis de retrouver l'activité initiale de thioligase de l'enzyme. Malgré tout, un premier essai de thioglycosylation enzymatique a été réalisé sur la Cystéine mais comme nous pouvions le prévoir aucun produit n'a été obtenu. La fonction thiol caractérisée par un pKa élevé n'est pas un bon nucléophile et la fonction acide carboxylique non protégée n'est pas compatible avec l'activité de l'enzyme. La validation du concept de synthèse enzymatique d'acides aminés S-glycosylés a donc été envisagée à partir d'un acide aminé non-usuel. Dans le cas de l'hydrolase DtMan une approche plus fondamentale était recherchée et avait pour but de déterminer si l'activité transglycosidase de l'enzyme native était une condition nécessaire pour obtenir une bonne thioligase. Après l'étude réalisée sur les 7 variants de DtMan nous avons pu répondre à notre problématique en montrant que la simple mutation du résidu acide/base permettait certes de réduire l'activité d'hydrolyse de l'enzyme mais n'avait aucun effet sur son activité de synthèse. Une activité intéressante a néanmoins été rencontrée dans le cas du mutant DtGly E425C dont l'activité d'hydrolyse était dépendante de la présence d'un agent réducteur. Ce comportement a laissé supposer de la formation d'un pont disulfure dans le site actif de l'enzyme. Une étude de cristallogénèse du mutant a été réalisée à partir des conditions de cristallisation préalablement optimisées pour DtMan *wt*. Un cristal en présence de DTT a pu être analysé au Synchrotron SOLEIL mais la résolution du modèle n'était pas complète au moment de l'écriture de ce manuscrit. Néanmoins la résolution partielle de la structure de DtMan type sauvage rend compte de la présence d'un résidu Cystéine à proximité du résidu acide/base et confirme la possibilité d'un pont disulfure pour le mutant E425C.

Enfin, la dernière partie du travail de thèse a été consacrée à la validation du concept de glycosylation enzymatique d'un acide aminé catalysée par le mutant thioligase DtGly E159Q. Pour répondre aux critères de relation structure/activité de l'enzyme, l'acide aminé non-naturel Thietyrosine (Ttyr), présentant un groupement thioarylique de pKa < 7, a été réalisée. Cette synthèse a été effectuée en trois étapes à partir d'un dérivé halogéné commercial de la L-Phénylalanine. L'incorporation de la fonction thiol a été réalisée par couplage catalysé au Palladium (0) et au terme d'une étape de déprotection générale, Ttyr a été obtenue avec un rendement global de 59 %. Pour favoriser la réaction de thioglycosylation, la fonction carboxylique a été protégée sous la forme d'un ester, conduisant à l'analogue TtyrME de la Thietyrosine avec un rendement final de 12 % sur quatre étapes. De par la réactivité de la fonction thiol, les deux molécules se trouvent sous la forme de dimère en solution. Pour s'assurer de la présence de la forme thiolate lors de la réaction enzymatique

une étude de Ttyr en solution a été réalisée. Au terme de cette étude il a été établi que la réaction de couplage enzymatique devait être conduite dans un milieu tamponné à pH 8 en présence de 10 équivalents de DTT par rapport à la Thiotyrosine. Compte-tenu de la faible quantité en accepteur TtyrME disponible, les essais de thioglycosylation n'ont pas pu être conduits dans des conditions classiques de synthèse, une méthode de suivi de réaction a donc dû être développée. Pour mettre au point ce suivi un test a été réalisé en présence de Thiotyrosine pour jouer le rôle de l'accepteur ce qui a permis de séparer et d'identifier efficacement chaque acteur de la réaction enzymatique par LC-MS. Cette méthode a par la suite été appliquée à la réaction de thioglucosylation de TtyrME et a permis de mettre en évidence l'apparition, après 1h d'incubation, d'un nouveau composé dont le spectre de masse présente un ion de masse équivalente à celle de la TtyrME S-glycosylée. L'analyse a également révélé que la réaction n'évolue plus après cette heure d'incubation, l'ajout d'une concentration plus forte en enzyme pourrait alors favoriser la réaction vers une conversion totale du donneur de glycoside et confirmer le couplage enzymatique. Ce premier résultat préliminaire s'avère donc être un faisceau d'indices prometteur pour la validation du concept de S-glycosylation d'un acide aminé catalysée par une thioglycoligase.

Au vu de ces travaux plusieurs perspectives à plus ou moins long terme sont à envisager :

- Utiliser les outils de bio ingénierie pour améliorer les propriétés de transglycosidase de DtRha et de DtMan par évolution dirigée. Une telle stratégie qui aboutit à la génération de banques importantes de mutants nécessite la mise au point d'une méthode de criblage permettant la sélection rapide et efficace des meilleurs variants. Une fois identifiés, ceux-ci pourront être modifiés pour obtenir les thioligases correspondantes et ainsi agrandir la boîte à outils enzymatique pour la glycosylation d'acides aminés.
- Améliorer la synthèse de la Thiotyrosine notamment l'étape de protection de la fonction acide carboxylique. Une alternative serait d'intégrer la Thiotyrosine dans une séquence peptidique synthétisée par SPPS. Pour ce faire il serait intéressant de s'inspirer des méthodologies de synthèse chimique et hémisynthétique de peptides contenant un acide aminé non-naturel.
- Pour réaliser une étude complète et prouver que le concept de thioglycoligase est applicable à la S-glycosylation de protéines, il serait intéressant de s'inspirer des travaux du Prof. Schultz en produisant *in-vivo* une protéine mutée contenant la Thiotyrosine par extension du code génétique.

### **Incorporation d'un acide aminé non-naturel sur une séquence peptidique par synthèse chimique ou hémisynthétique.**

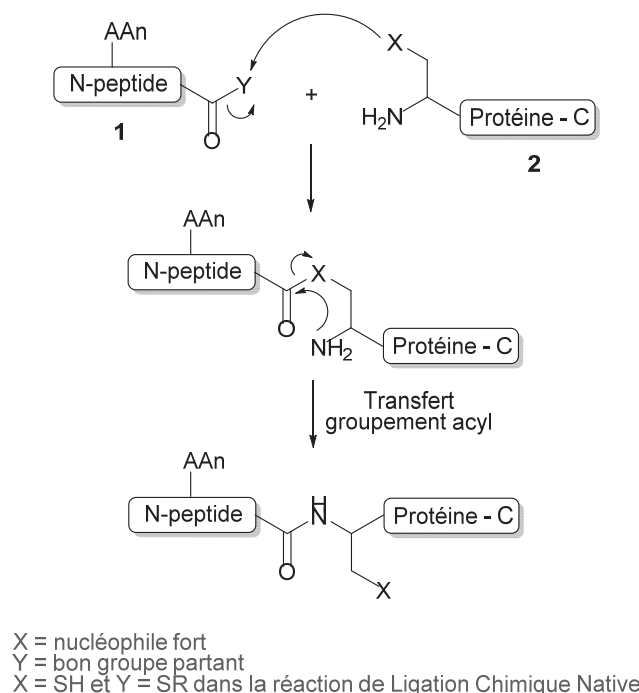
Les techniques conventionnelles de synthèse de peptides et protéines ont été appliquées à l'incorporation d'acides aminés non-naturels. Par la méthode de synthèse



peptidique en phase solide (SPPS) il est aisé d'introduire un nouvel acide aminé soit i) à partir de l'AAn synthétisé en solution et greffé sur le peptide cible supporté en cours d'élongation ou ii) l'AAn est directement produit lors de la synthèse de la chaîne peptidique en phase solide. (Oh, **1999**; Roy, **1997**) Malgré la simplicité de la méthode, celle-ci se limite à la synthèse de peptides ou protéines de petites tailles (< 50 AA). Pour accéder à des protéines plus larges les méthodes de ligation chimique ont été utilisées pour assembler plusieurs fragments peptidiques. En 1998 Low *et coll.* reconstituèrent la rubredoxine par Ligation Chimique Native (NCL) en remplaçant le résidu Tyr10 par différents analogues et étudier leurs effets sur les propriétés redox de la protéine. (Low, **1998**)

Des méthodes d'hémisynthèse ont également été proposées pour synthétiser une protéine d'intérêt dont le principe général repose sur la production du peptide contenant un ou plusieurs AAns par SPPS d'une part et la production de la protéine recombinante cible d'autre part. Le couplage des deux fragments peut alors se faire par voie enzymatique ou par voie chimique *via* :

- Ligation chimique, en introduisant des groupements bio-orthogonaux sur les positions terminales des fragments **1** et **2** (Schéma 68). en résulte alors la protéine hémisynthétique avec un site de ligation non-naturel de type thioester, thioether, disulfure,... ; (Han, **2004**)
- Un couplage suivant la méthode de Ligation Chimique Native (LCN), ou une liaison peptidique native peut être créée par la présence d'un résidu Cys sur la position *N*-term du peptide **1** et d'une fonction  $\alpha$ -thioester sur la position *C*-term de la protéine recombinante **2** (Schéma 68). Cette technique est aussi appelée Ligation de Protéine Exprimée (LPE). Raines *et coll.* ont utilisé cette méthode pour générer un mutant de la RNase A et ont montré l'augmentation de la thermostabilité de l'enzyme par modification des résidus Asn113-Pro114 *via* deux acides aminés  $\beta$ . (Arnold, **2002**)



**Schéma 68 :** Incorporation d'AAn par hémisynthèse de protéines.

Ainsi dans notre cas, il pourrait être envisagé de construire un peptide par SPPS contenant la Thiotyrosine qui pourrait ensuite être glycosylée par DtGly E159Q. L'avantage d'intégrer l'AAn dans une séquence peptidique est que i) la fonction acide carboxylique est engagée dans la liaison peptidique, l'étape de protection qui était limitante dans la synthèse de Tyr n'est plus nécessaire et ii) la préparation de peptides par SPPS aujourd'hui automatisée est une méthode rapide et efficace et permet la production de séquences peptidiques pures sans passer par des étapes de purification fastidieuses. Le peptide ainsi glycosylé pourra par la suite être utilisé en hémisynthèse pour reconstituer une protéine S-glycosylée plus large.

### Biosynthèse de protéines mutées et expansion du code génétique

Dans un premier temps les techniques se sont basées sur l'incorporation d'analogues des acides aminés usuels sur des protéines bactériennes. Cette méthode a été proposée par Munier *et coll.* en 1956 et repose sur l'utilisation de souches d'*E. coli* auxotrophes incapables de biosynthétiser l'acide aminé naturel cible. Après induction, l'expression protéique qui a lieu dans un milieu enrichi par l'analogue de l'acide aminé cible conduit à la génération de la protéine mutée présentant l'acide aminé non-naturel. (Munier, **1956**) Cette méthode s'est notamment révélée utile pour l'incorporation de sondes pour l'analyse RMN ou RX de protéines. L'exemple le plus connu reposant sur la substitution de la Méthionine par la Sélénométhionine pour la résolution de structures de protéines aux RX. (Hendrickson, **1997**) Malgré la simplicité de la méthode celle-ci présente l'inconvénient d'incorporer l'AAn sur tous les sites où l'acide aminé cible est codé. Ces modifications peuvent alors avoir des répercussions importantes sur la structure et la fonction de la protéine ainsi construite. Une alternative a par la suite été proposée par l'équipe du Prof. Schultz basée sur l'utilisation

d'ARNt et Aminoacyl-ARNt synthétase (AARS) orthogonales. Cette méthode décrite dans un premier temps pour la synthèse *in-vitro* de protéines mutées fut par la suite étendue à leur biosynthèse *in-vivo* à travers l'extension du code génétique d'un organisme hôte qui représente à l'heure actuelle une avancée technique considérable.

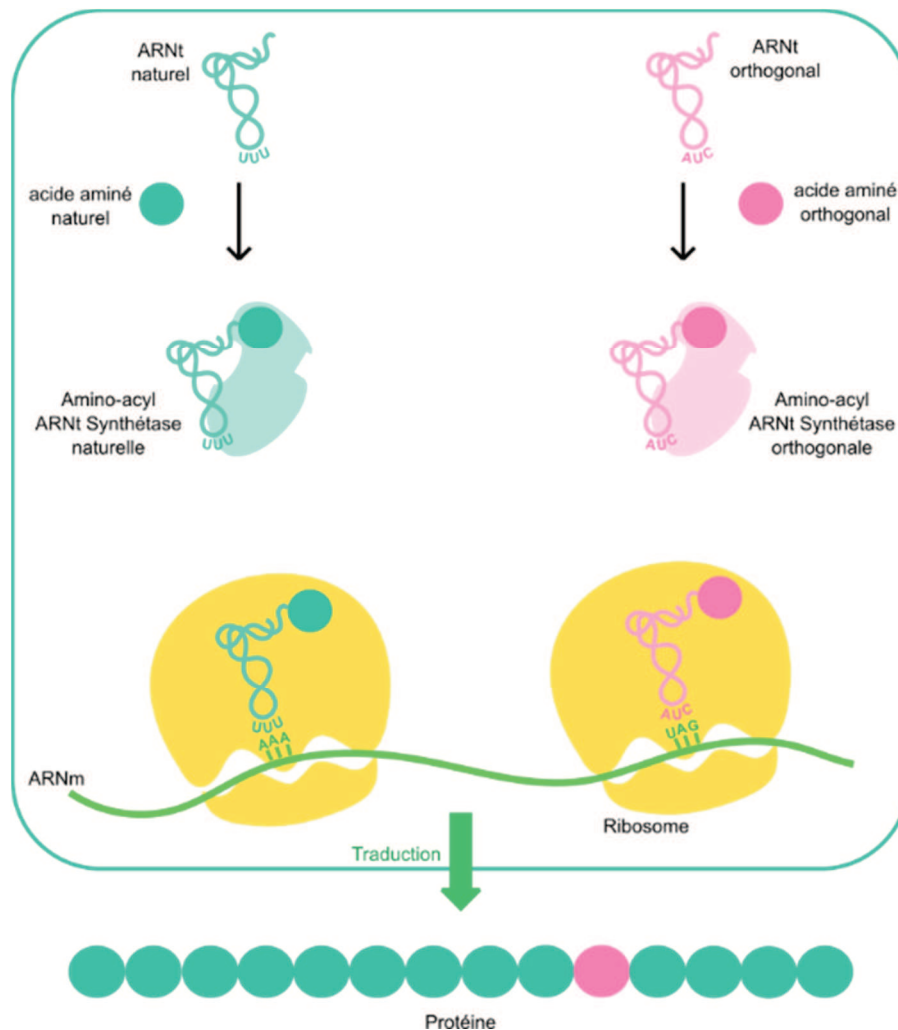
La biosynthèse des protéines peut se résumer selon les 3 étapes suivantes : i) transcription de l'ADN en ARN messager (ARNm) ; ii) traduction de l'ARNm produisant la séquence protéique et iii) modifications post-traductionnelles.

L'incorporation d'un acide aminé sur la séquence protéique lors de la traduction génétique implique de nombreux facteurs, dont les 3 acteurs suivant : i) une ARN transférase (ARNt) présentant l'anti-codon complémentaire du codon lu sur l'ARNm ; ii) l'acide aminé spécifique de codon et iii) l'Aminoacyl-ARNt-synthétase (AARS) correspondante permettant la fixation de l'AA sur l'ARNt (Figure 82).

Chaque codon du code génétique, parmi les 64 disponibles, correspond à un acide aminé spécifique à l'exception des 3 codons stop. Ces derniers, de séquence UAA (codon ocre), UAG (codon ambre) et UGA (codon opale), marquent la fin de la traduction génétique car il n'existe aucun ARNt correspondant.

La méthodologie proposée par Schultz *et coll.* repose sur la mutation de l'anti-codon d'une ARNt alors capable de reconnaître le codon ambre (UAG). Cet ARNt muté appelé **ARNt supprimeur** est alors capable de traduire le codon stop ambre par n'importe quel acide aminé. L'AA est dans un premier temps fixé sur l'ARNt supprimeur par aminoacylation chimique, l'ARNt ainsi modifié peut lire le codon UAG intégré dans la séquence ADN du gène

d'intérêt en résulte alors la protéine correspondante avec l'AAn incorporé sur le site spécifié (Figure 82). (Noren, **1989**)

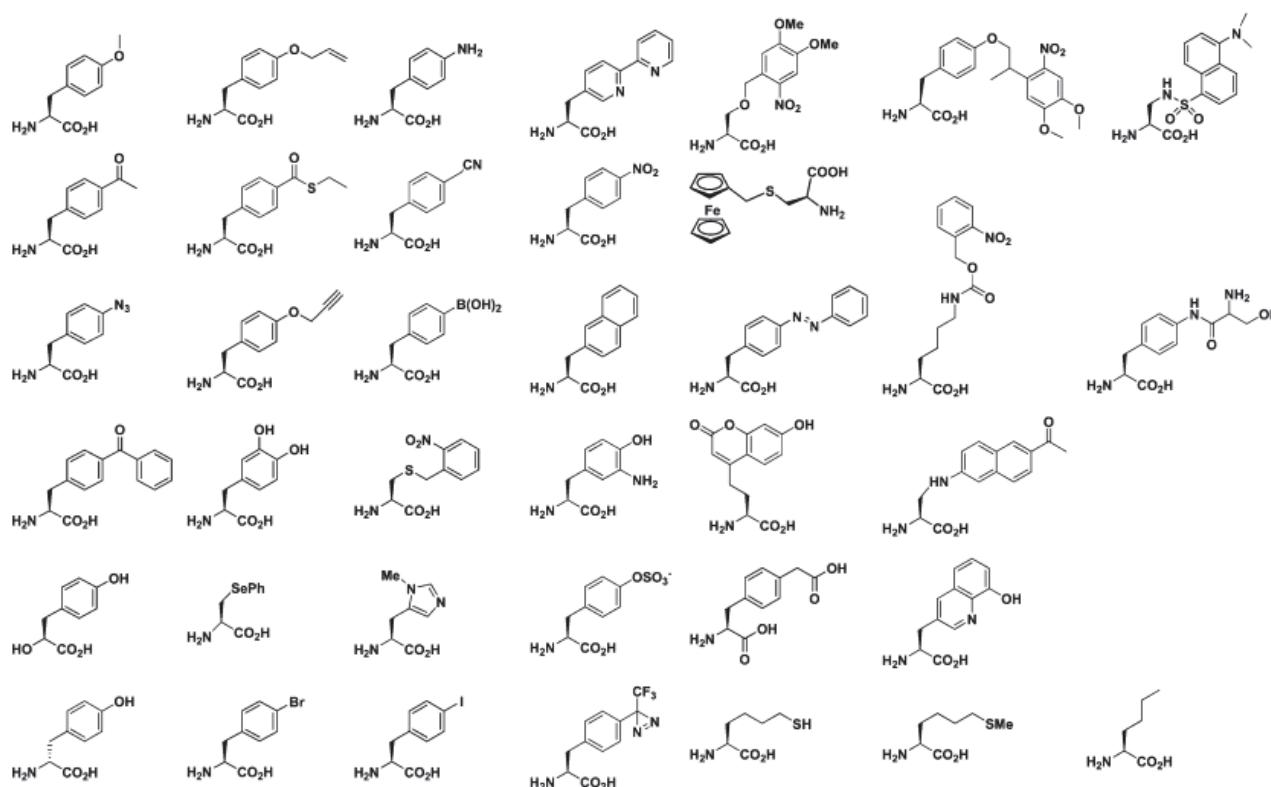


**Figure 82** : Incorporation génétique d'un acide aminé non-naturel.

Pour cette synthèse *in-vitro*, l'ensemble du matériel transcriptionnel et traductionnel est extrait d'un organisme sélectionné (bactérie, levures, cellule de mammifère,...). L'efficacité d'incorporation de l'AAn est assurée par le design d'un ARNt suppresseur orthogonal aux ARNt-synthétases provenant de l'organisme hôte. Dans le cas contraire l'incorporation de l'AAn est en compétition avec l'introduction d'un acide aminé usuel en réponse au codon stop. La première application de cette méthode fut la production de mutants de la  $\beta$ -lactamase par substitution de la Phe66 par les analogues p-nitrophenylalanine, p-fluorophenylalanine ou homophenylalanine. (Noren, **1989**) Depuis les travaux pionniers de Noren *et coll.* la méthodologie de mutagenèse *in-vitro* de protéines a fait l'objet de nombreuses études et permis l'incorporation d'une variété d'AAns permettant la résolution de mécanismes enzymatiques ou de structures protéiques en autre. (Wang, **2005**)

Si la méthode se révèle efficace elle est toutefois limitée par de faibles rendements, l'impossibilité d'étudier les protéines mutées *in-vivo*, et la nature de la chaîne latérale de l'AAn conditionnée par l'étape d'aminocyclisation chimique. Pour contourner ces limitations l'équipe

du Prof. Schultz proposa d'étendre son concept à la mutagenèse *in-vivo* de protéines. La méthode repose alors sur l'altération de la machinerie biosynthétique d'un organisme hôte pour qu'il puisse produire de lui-même le couple AARS/ARNt permettant l'introduction de l'AAn correspondant. Comme décrit précédemment, pour assurer une traduction fidèle le couple AARS/ARNt se doit d'être orthogonal au système existant dans l'organisme hôte. Le couple peut alors être produit par design génétique, en utilisant un plasmide présentant les gènes codant pour le couple AARS/ARNt désiré ou en le prélevant d'un organisme différent de l'organisme hôte (par exemple un système AARS/ARNt extrait d'*E. coli* pourra être utilisé chez les levures). De la même manière l'acide aminé exogène à l'organisme hôte doit alors être conçu pour permettre son transport à travers la membrane cellulaire et ainsi favoriser son incorporation dans la séquence protéique par le système orthogonal AARS/ARNt. A partir de cette méthode élégante de nombreux AAns ont pu être encodés dans le code génétique d'organismes procaryotes et eucaryotes (Figure 83).(Wu, 2009)

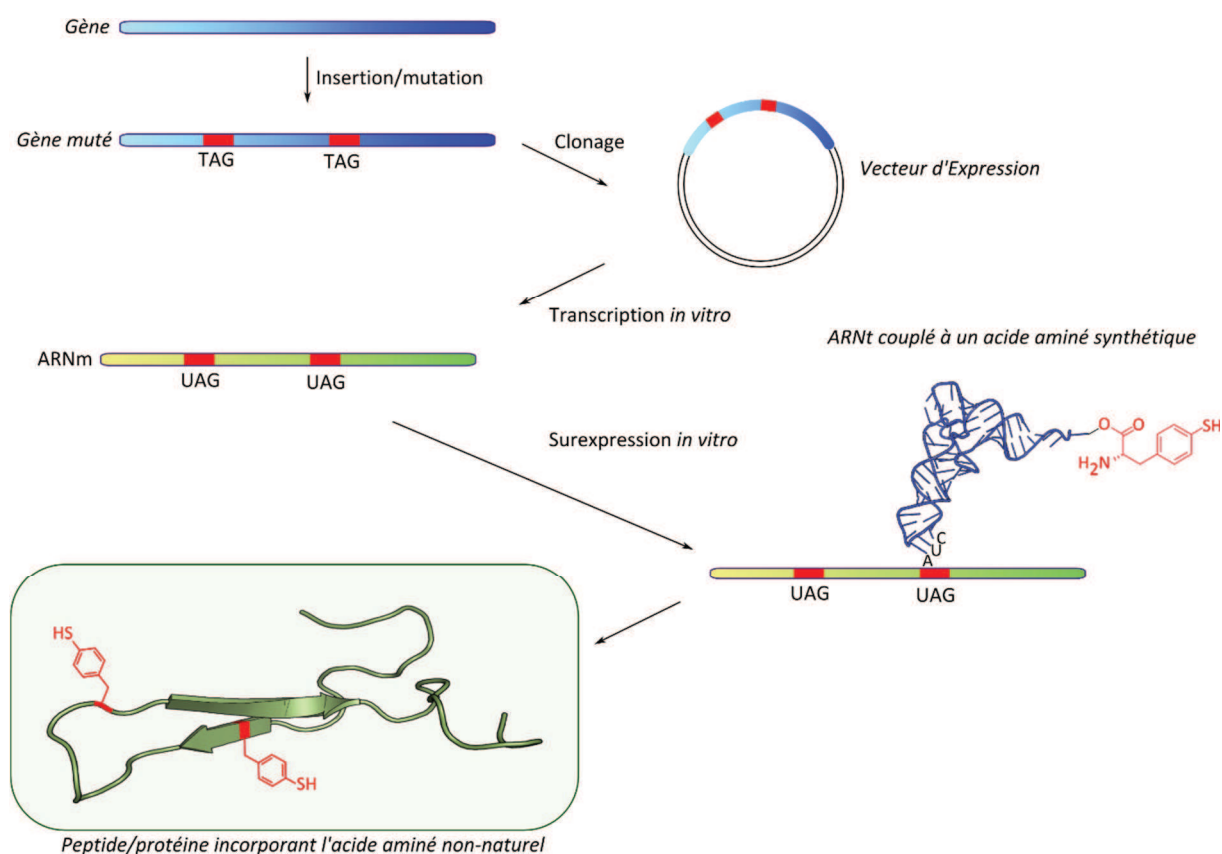


**Figure 83 :** Répertoire d'AAns intégrés dans le code génétique d'organismes procaryotes et eucaryotes.(Wu, 2009)

En 2003 Mehl *et coll.* ont décrit la première souche d'*E. coli* présentant un code génétique codant pour 21 acides aminés. Cette bactérie est alors capable de i) synthétiser l'acide aminé p-Aminophenylalanine (*pAF*) « naturellement » à partir de sources carbonées simples, ii) d'incorporer *pAF* *via* le système *pAF*-AARS/ARNt<sup>*pAF*</sup> sur la séquence protéique en réponse au **codon ambre TAG** avec d'excellents rendements et une haute fidélité.(Mehl, 2003)

De nombreuses méthodes analogues à l'extention du code génétique proposée par Schultz *et coll.* ont également été développées. Notamment Hohsaka et Sisido ont mis au point une méthode originale en basant leurs travaux sur la création de nouveaux codons constitués de 4 bases, UAGN, ou 5 bases, UAGN<sub>1</sub>N<sub>2</sub>, (N = A, G, T ou C) au lieu des 3 bases classiques. L'extension de la taille des codons sens et anti-sens permet ainsi l'incorporation de plusieurs AAns et d'augmenter le panel d'AAns pouvant être encodé par un organisme hôte. (Hohsaka, 2000; Johnson, 2010)

En s'inspirant de la méthodologie développée par le Prof. Schultz, elle pourrait être appliquée à la biosynthèse d'une protéine contenant un ou plusieurs résidus Thiotyrosine. Ce travail impliquerait alors le design du plasmide codant pour le couple AARS/ARNt associé à la Thiotyrosine, la transformation chez *E. coli* ou un autre système d'expression et enfin la surexpression de la protéine cible contenant un ou plusieurs résidu non-naturel (Figure 84). Le concept de S-glycosylation d'un acide aminé par une thioglycoligase pourra alors être étendu au modèle de la protéine et ainsi proposer une voie d'accès innovante aux protéines S-glycosylées.



**Figure 84 :** Production in-vivo d'une protéine contenant un ou plusieurs résidus Thiotyrosine, site(s) de glycosylation enzymatique.



*Les travaux réalisés durant cette thèse ont fait l'objet de plusieurs publications et communications scientifiques :*

**Unraveling the substrate recognition mechanism and specificity of the unusual glycosyl hydrolase family 29 BT2192 from *Bacteroides thetaiotaomicron*.**

Guillot, L., Lafite, P. and Daniellou, R.; *Biochemistry* (2014), 53, 1447-1455

**Enzymatic thioglycosylation: current knowledge and challenges.**

Guillot, L., Lafite, P. and Daniellou, R.; *Carbohydr. Chem.* (2014), 40, 178-194

**Improved synthesis of ethyl 1-Thio- $\alpha$ -D-galactofuranoside**

Jayasuiya, A. B., Peng, W., Guillotin, L., Lowary, T.; *Carbohydr. Chem. : Proven Synth. Methods*, (2015), 3, 143-146

*Ce travail a également été valorisé lors de plusieurs congrès nationaux et internationaux sous la forme de communications orales et de posters:*

Communications orales :

**Guillot, L.**, Lafite, P. and Daniellou, R. Génération de nouvelles thioglycoligases pour la synthèse de thioglycoconjugués. 25<sup>ème</sup> Journées du Groupe Français des Glycosciences (GFG), 12-15 Mai 2014, Paris (75).

**Guillot, L.**, Lafite, P. and Daniellou, R. *Développement d'outils enzymatiques pour la synthèse de protéines S-glycosylées*. Journées de la Société Chimique de France (SCF), section régionale Centre Ouest, 18-20 Février 2015, Orléans (45). Prix de la meilleure communication

Communications par poster :

**Guillot, L.**, Lafite, P. and Daniellou, R. *BT2192, une  $\alpha$ -L-fucosidase aux propriétés remarquables*. (SFBBM), 4-6 Septembre 2013, Paris (75).

**Guillot, L.**, Lafite, P. and Daniellou, R. *BT2192, une  $\alpha$ -L-fucosidase aux propriétés remarquables*. 26<sup>ème</sup> colloque Biotechnocentre, 10-11 Octobre 2013, Seillac (41). Prix du meilleur poster.

**Guillot, L.**, Lafite, P. and Daniellou, R. *Thioglycoligases : innovative biocatalytic tools for S-glycosylated proteins synthesis*. 11th Carbohydrate Bioengineering Meeting (CBM), 10-13 Mai 2015, Espoo (Finlande). Prix du meilleur poster.



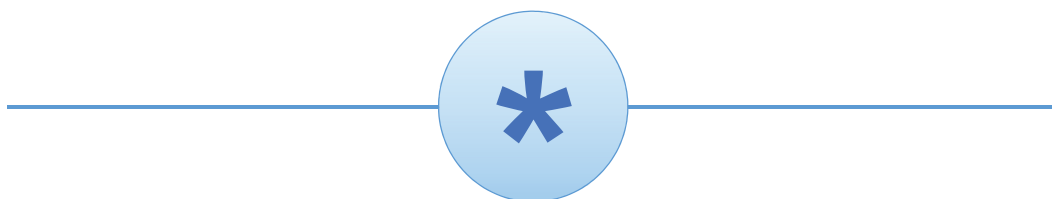


## Bibliographie

- Arnold, U., Hinderaker, M. P., Nilsson, B. L., Huck, B. R., Gellman, S. H., and Raines, R. T.; *J. Am. Chem. Soc.* **(2002)**, 124, 8522-8523
- Han, S., and Viola, R. E.; *Protein Pept. Lett.* **(2004)**, 11, 107-114
- Hendrickson, W. A., and Ogata, C. M.; *Methods Enzymol.* **(1997)**, 276, 494-523
- Hohsaka, T., and Sisido, M.; *Nucleic Acids Symp. Ser.* **(2000)**, 44, 99-100
- Johnson, J. A., Lu, Y. Y., Van Deventer, J. A., and Tirrell, D. A.; *Curr. Opin. Chem. biol.* **(2010)**, 14, 774-780
- Low, D. W., and Hill, M. G.; *J. Am. Chem. Soc.* **(1998)**, 120, 11536-11537
- Mehl, R. A., Anderson, J. C., Santoro, S. W., Wang, L., Martin, A. B., King, D. S., Horn, D. M., and Schultz, P. G.; *J. Am. Chem. Soc.* **(2003)**, 125, 935-939
- Munier, R., and Cohen, G. N.; *Biochim. Biophys. Acta* **(1956)**, 21, 592-593
- Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G.; *Science* **(1989)**, 244, 182-188
- Oh, J. E., and Lee, K. H.; *Bioorg. Med. Chem.* **(1999)**, 7, 2985-2990
- Roy, R. S., and Imperiali, B.; *Protein Eng.* **(1997)**, 10, 691-698
- Wang, L., and Schultz, P. G.; *Angew. Chem., Int. Ed.* **(2005)**, 44, 34-66
- Wu, X., and Schultz, P. G.; *J. Am. Chem. Soc.* **(2009)**, 131, 12497-12515



## Experimental Part





## Molecular Biology, Biochemistry & Enzymology

### Experimental procedures

#### DNA Cloning

The open reading frame encoding DtMan *wt*, DtRha *wt* and DtGly *wt*, were amplified from *D. thermophilum* DSM 3960 genomic DNA (DMSZ Institute, Germany) by PCR using Phusion High-Fidelity DNA polymerase (ThermoFisher) and designed primers containing selected restriction enzymes (FastDigest, ThermoFisher) as described below. PCR products were directly ligating into the pET-28a(+) vector using T4 DNA Ligase (Rapid DNA ligation kit, Fermentas) and final constructs were transformed in *E. coli* DH5 $\alpha$  affording plasmid pET-28a(+) containing desired *D. thermophilum* glycosidase gene.

	Primers	Restriction enzymes
DtMan <i>wt</i>	Forward (F) 5'-AATTGCTAGCATGGAAATATTGAAGCTTGATG-3'	<i>Nhe I</i> / <i>Not I</i>
	Reverse (R) 5'-TTTGCGGCCGCTTAAGACTTTATAACCT-3'	
DtRha <i>wt</i>	Forward (F) 5'-TTGCTAGCATGAAGTCAAGTAATATTTACTCACCTT-3'	<i>Nhe I</i> / <i>Sal I</i>
	Reverse (R) 5'-TTGTCGACTTATATCTTTTCATATAAAAGTTGTATGATCC-3'	
DtGlc <i>wt</i>	Forward (F) 5'-AATTGCTAGCATGCTTAAATACAGGTTTCCTG-3'	<i>Nhe I</i> / <i>Not I</i>
	Reverse (R) 5'-TTTGCGGCCGCTATTTAAGAAAC-3'	

#### Plasmid DNA purification, analysis and sequencing

A single positive colony of transformed *E. coli* DH5 $\alpha$  was selected and used to prepare an inoculum. The bacteria were grown in LB medium (10 ml LB-growth containing Kanamycin antibiotic (34 mg/ml, 1/1000) at 37 °C, 250 rpm, and overnight. Then cells were harvested and plasmid DNA was extracted and purified using E.N.Z.A Minikit (OmegaBiotect).

Plasmid DNA constructs were analysed through agarose electrophoresis. Gels were prepared with 0.8 -1 % agarose in TBE 1x buffer (TBE 5x: 53 g TrisBase, 27.5 g Borate, 1.21 g EDTA) containing SYBRSafe for UV visualization ( $\lambda$  = 254 nm). Migration was performed at 75 – 100 mV electrical field.

DNA sequencing of purified plasmids was performed by EurofinsGenomics laboratories (Germany) and results were analysed with CLC Bio software (<http://www.clcbio.com/>). All sequence alignments are available in the Appendix part p. 249/254/258.

### **Expression and purification of proteins**

Pure plasmid DNA constructs were transformed in *E. coli* Rosetta-(DE3) using heat shock method. 5 µl of pure plasmid DNA was added to 50 µl chemically competent Rosetta strain defrost aliquot at 4 °C. Solution was incubated at 4 °C for 20 minutes before proceeding 90 sec heat shock treatment at 42 °C and mixture was immediately transferred to ice. Cells were then grown in LB medium containing Kanamycin (34 µg/ml) and Chloramphenicol (30 µg/ml) antibiotics, 1h at 37 °C and dropped off on LB-agar plates containing Kan/Can. A single colony was selected and cultivated in LB medium containing usual antibiotics at 37 °C until OD<sub>600</sub> reach 0.6. The culture was then induced by adding IPTG at 0,2 mM final concentration and incubated for 20 hours at 37°C, 250 rpm.

Cells were harvested by centrifugation, resuspended in 100 mL (1/10 culture volume) cocktail solution (TrisHCl 50 mM pH8, NaCl 200 mM, lysozyme and protease inhibitor), and lysed by freeze/thaw cycles and sonication. After centrifugation step (32 000 x g for 20 min) clarification of supernatant was carried out by heat treatment (15 min at 70°C) in order to remove non-thermostable proteins. Final centrifugation (32 000 x g for 20 min) was done and supernatant protein concentrations were determined by the Bradford method.(Bradford, **1976**; Compton, **1985**)

### **Purification and analysis of proteins**

The His-tagged *D. thermophilum* proteins were purified from the cell free extracts on AKTA FPLC Purifier (GE Healthcare) using IMAC (Immobilized Metal Ion Affinity Chromatography) columns (GE Healthcare). 1 ml columns were equilibrated with Tris-HCl (50mM) –NaCl (200mM) buffer (pH8) and the bound proteins were eluted using an imidazole gradient (10-500mM)). For crystallization assays, proteins were further purified through SEC (Size Exclusion Chromatography) purification using tris-HCl (50mM)-NaCl (200mM) buffer (pH8) as eluent. The purity of proteins was confirmed by SDS-PAGE electrophoresis.

### **Bradford Protein assays**

Standard calibrators were prepared in concentration ranging from 0 to 0.5 mg/ml using a BSA stock solution (1 mg/ml). 10 µl of calibrators and protein samples (0 – 0.5 mg/ml) were added in a 96 well-plate in triplicate. To each well 200 µl of Bradford reagent were added (Bio-Rad). After 5 min incubation at room temperature, absorbance was measured at 595 nm. Protein samples concentrations were calculated according experimental calibration curve.(Bradford, **1976**; Compton, **1985**)

### **Enzyme activity assays**

All *D. thermophilum* wt enzymes' activity were assayed in triplicate in a 96 well-plate at 37°C in 200 µL reaction mixtures containing substrate (1 mM), 20 µL of diluted enzyme and Tris buffer (200 mM, pH 7.0). Residual spontaneous hydrolysis of the substrate was determined on sample containing dH<sub>2</sub>O instead of enzyme. For *para*-nitrophenol (pNP) containing substrates,

after 30 min reaction, 100  $\mu$ L of sodium carbonate 1 M were added, and produced pNP was quantified by absorbance measurement at 405 nm.

### ***Kinetic parameters determination***

The kinetic parameters  $K_M$ ,  $k_{cat}$  and  $k_{cat}/K_M$  were quantified by incubation of 200  $\mu$ L reaction mixture containing substrate (0,01 – 10 mM), 20  $\mu$ L of diluted enzyme, and buffer (according to pH Opt of each enzyme) a 96 well-plate at 37 °C. Residual spontaneous hydrolysis of the substrate was determined on sample containing dH<sub>2</sub>O instead of enzyme. The release of pNP was measured at 405 nm during 30 min. All kinetics parameters were calculated by fitting of saturation curves with standard Michaelis-Menten, using Prism 6 software (GraphPad).

For experiments run at 70 °C (DtMan *wt* and mutants), reaction mixtures were prepared in 500  $\mu$ L in microtubes following the same protocol as above. 100  $\mu$ L aliquots were withdrawn in a 96 well-plate after 10, 20 and 30 min incubation and pNP release was measured at 405 nm. All kinetics parameters were calculated by fitting of saturation curves with standard Michaelis-Menten, using Prism 6 software (GraphPad).

### ***Effect of pH and temperature on enzyme activity***

The optimum pH for *D. thermophilum* enzymes was determined by measuring appropriate pNPGlycoside hydrolysis under several pH values ranging from pH 4 to pH 10. The buffers used were :

- Acid citric/sodium phosphate buffer(20 mM) from pH 4.0 to 6.0;
- Imidazole/HCl buffer (20 mM) for pH 7.0;
- Tris/HCl buffer (20 mM) for pH 8.0;
- TrisBase buffer (20 mM) for pH9;
- Sodium carbonate buffer (20 mM) for pH10.

100  $\mu$ L samples containing 1mM pNPGlycoside, 10  $\mu$ L buffer and 5 $\mu$ L of diluted DtMan (*wt* and mutants), DtRha or DtGly (*wt* and mutants), were incubated 5 min at 37°C. Then reactions were quenched by adding 50  $\mu$ L Na<sub>2</sub>CO<sub>3</sub> (1M) and absorbance was measured at 405nm.

The dependence of the enzyme activity on temperature was determined by measuring the hydrolysis of pNPGlycosides for several temperatures ranging from 40°C to 100°C. Samples were prepared following the same protocol than pH-dependence study but only at the determined enzyme optimum pH. After 5 min incubation, reaction was quenched by adding 50  $\mu$ L Na<sub>2</sub>CO<sub>3</sub> (1M) and relative activity was calculated according to absorbance measured for each temperature at 405 nm.



***Influence of co-solvent***

In order to anticipate further uses of DtMan *wt*, DtRha *wt* or DtGly *wt* in organic media, the influence of various co-solvents on those enzymes' activity was studied. MeOH, CH<sub>3</sub>CN, DMSO and DMF in proportion (v/v) ranging from 5 to 50% were used. Reaction mixture were prepared in a 96 well-plate containing substrate (1 mM), 20 µl of diluted enzyme, buffer (according to pH Opt of each enzyme) and organic solvent 5 – 50 % v/v. After 1h incubation at 37 °C, pNP released was quantified at 405 nm and resulting enzyme activity was compared with control sample without organic solvent.

***Sequence analysis***

DtMan *wt*, DtRha *wt* and DtGly *wt* peptidic sequences were retrieved from Uniprot database ([www.uniprot.org](http://www.uniprot.org); accession numbers are B5YAN4, B5YC64 and B5YCI2 respectively) and were blasted using the implemented Blast software in Uniprot against Uniprot and Swiss Prot Databases. Sequences alignments between DtGlycosidase and their respective GH family were performed using Clustal Omega at the EBI server (EBI, <http://www.ebi.ac.uk/Tools/msa/clustalo/>). (Goujon, 2010; Sievers, 2011)

**Directed mutagenesis to DtGly and DtMan thioglycoligase mutants**

The position of catalytic residues of DtGly *wt* and DtMan *wt* was determined by sequence alignment using Clustal Omega at the EBI server and CLC Bio software. Primers containing desired mutation on acid/base residue position (DtGly Glutamate E159; DtMan Glutamate E425) were designed on with Agilent tools website ([www.genomics.agilent.com](http://www.genomics.agilent.com)) as described below (mutated codons are highlighted in bold) :

Mutant	Primers	
DtGly E159A	5'- GAATTACTGGATGACTATAAAT <b>GCG</b> CCAATGCTTATGCCTTT-3'	Forward (F)
	5'- AAAGGCATAAGCATTGGG <b>CG</b> CATTTATAGTCATCCAGTAATTC-3'	Reverse (R)
DtGly E159Q	5'-CATAAGCATTGGGCTGATTTATAGTCATCC <b>CAG</b> TAATTCACAAGAT-3'	Forward (F)
	5'-ATCTTGTGAATT <b>ACT</b> GGATGACTATAAATCAGCCAATGCTTATG-3'	Reverse (R)

Mutant	Primers	
DtMan E425A	5'-CTATAGTTCTTTGGTGTGGAAATAAT <b>GCG</b> AACCATTGGGGATATTA-3'	Forward (F)
	5'-TAATATCCCCAATGGTT <b>CGC</b> ATTATTTCCACACCAAAGAACTATAG-3'	Reverse (R)
DtMan E425C	5'-AATCATCCTTCTATAGTTCTTTGGTGTGGAAATAAT <b>TGCA</b> AACCATTGGGGATATTACG-3'	Forward (F)
	5'-CGTAATATCCCCAATGGTT <b>GCA</b> ATTATTTCCACACCAAAGAACTATAGAAGGATGATT-3'	Reverse (R)
DtMan E425Q	5'-CTATAGTTCTTTGGTGTGGAAATAAT <b>CAGA</b> AACCATTGGGGGA-3'	Forward (F)
	5'-TCCCCAATGGTT <b>CTG</b> ATTATTTCCACACCAAAGAACTATAG-3'	Reverse (R)
DtMan E425G	5'-CTTCTATAGTTCTTTGGTGTGGAAATAAT <b>GGA</b> AACCATTGGGGATATTACGC-3'	Forward (F)
	5'-GCGTAATATCCCCAATGGTT <b>TCC</b> ATTATTTCCACACCAAAGAACTATAGAAG-3'	Reverse (R)
DtMan E425N	5'-ATCCTTCTATAGTTCTTTGGTGTGGAAATAAT <b>AACA</b> AACCATTGGGGATATTA-3'	Forward (F)
	5'-TAATATCCCCAATG <b>GTT</b> GTTATTATTTCCACACCAAAGAACTATAGAAGGAT-3'	Reverse (R)
DtMan E425S	5'-AATCATCCTTCTATAGTTCTTTGGTGTGGAAATAAT <b>TCAA</b> AACCATTGGGGATATTACG-3'	Forward (F)
	5'-CGTAATATCCCCAATGGTT <b>TGA</b> ATTATTTCCACACCAAAGAACTATAGAAGGATGATT-3'	Reverse (R)
DtMan E425Y	5'-TCATCCTTCTATAGTTCTTTGGTGTGGAAATAAT <b>TATA</b> AACCATTGGGGATATTAC-3'	Forward (F)
	5'-GTAATATCCCCAATGGTT <b>ATA</b> ATTATTTCCACACCAAAGAACTATAGAAGGATGA-3'	Reverse (R)

Site-directed mutagenesis were carried out using QuikChange II XL Site-Directed Mutagenesis kit (Agilent) and pure plasmids pET28-a(+)-DtGlc *wt* or pET28-a(+)-DtMan *wt* as DNA template. Mutagenesis products were directly transformed into XL10-Gold Ultracompetent cells (Agilent) by heat shock method. After Agarose electrophoresis analysis, each mutant sequence DNA was sequenced (Eurofins genomics) and confirmed to be identical to the known *wild-type* enzyme sequence DNA except the target codon of catalytic Glutamate replace by the desired amino acid codon. All sequence alignments are available in the Appendix part p. 262 and p. 267.

Mutant enzymes were expressed and purified following the general method of expression and purification of protein described earlier in this experimental part.

### **Thioglycoligase activity screening**

DtGly E159X mutants and DtMan E425X mutants thioglycoligase activity was assayed in microtubes in water bath at 70 °C. 500 µl reaction mixture containing 1 mM substrate (pNPGlc or pNPMAN), 2 to 10 equiv. of various thiol-containing acceptors (from 50 mM mother liquor in MeOH) and in buffer (according to pH Opt of *wild-type* enzymes). Residual hydrolysis activity of enzymes was determined on samples without acceptor. 100 µl aliquots were withdrawn in a 96 well-plate after 5, 10, 20 and 30 min incubation and pNP release was measured at 405 nm. Plotting of Abs vs incubation time allowed to quantified enzyme's  $V_{\max}$  and thus the corresponding  $k_{\text{cat}}$  value.

### **Crystallization, data collection, and model refinement**

Crystallization studies were carried out with the 100 screening conditions Structure Screen 1 kit (Molecular Dimensions) or 96 screening conditions JBScreen Classic HTS I kit (Jena Bioscience). Protein crystal were grown at 20 °C by sitting-drop vapor diffusion in 2 µl drops containing equal amount of pure protein (5 to 20 mg/ml) and screening buffer. After selection of crystal hits, crystallization conditions were optimized by modifying the precipitant and/or the additive concentration. After equilibration of the drops against the crystallization solution, crystals were cryoprotected in mineral oil (MiteGen LV Oil), before freezing in liquid nitrogen. Data were collected on beamline Proxima-1 or Proxima-2 at Soleil Synchrotron Facility, scaled and reduced with the X-Ray Detector Software XDS.(Kabsch, **2010**) For DtMan wt model refinement, Man2A from *Bacteroides thetaiotaomicron* crystal structure (pdb 2JE8)(Tailford, **2007**) was used as a template for molecular replacement in PHASER program.(McCoy, **2007**) The model is currently completed by iterative cycles of manual model building and real space refinement using the program Coot(Emsley, **2010**) and crystallographic refinement using Phenix.refine.(Adams, **2010**) Pymol software was used for generation of all protein structure images (<http://www.pymol.org>).



### **Antimicrobial assay**

The potential antimicrobial activity of the products was assessed by the micro-challenge test method originally described in abdellaoui et al., 2010.(El Abdellaoui, **2010**)

The microorganisms used include all the strains recommended in the current reglementary method for cosmetic preservative efficacy testing (EN ISO11930): *Pseudomonas aeruginosa* (Pa) ATCC 9027, *Staphylococcus aureus* (Sa) ATCC 6538, *Escherichia coli* (Ec) ATCC 8739, *Candida albicans* (Ca) ATCC 10231 and *Aspergillus brasiliensis* (Ab) ATCC 16404. These microorganisms were grown 24-48h at 35°C on tryptone soja broth (Biokar, ref BK046HA) for the bacteria, Yeast Mold broth (Difco, ref 271120) for *Candida albicans* or Sabouraud broth (Biokar, ref BK026HA) for *Aspergillus brasiliensis*, before preparation of a pure saline inoculum

containing between  $10^6$ - $10^7$  (fungus and yeast) or between  $10^7$ - $10^8$  (bacteria) cfu mL<sup>-1</sup>. Prior to the contamination, each sample was first solubilized in water at 5 mg.mL<sup>-1</sup>, then in 0.85 % saline to obtain different concentrations from 12.5 µg.mL<sup>-1</sup> to 500 µg.mL<sup>-1</sup>. Phenonip 0.5 % and 0.85 % saline solutions were respectively used as positive and negative controls. Samples and controls were deposited in duplicate, in five deep-wells plates (one for each tested microorganism) and contaminated with 10 µl of inoculum of each microorganism. The plates were covered with adhesive breathable film and incubated at 22°C. Microbial counts were done at day 1 (24h), days 7 and 14 for each sample. Counting was performed with a TTC micromethod for the yeast and bacteria and with the conventionnal agar plating method on Sabouraud gelose (Biokar, BM17308) for *Aspergillus brasiliensis*. Triphenyltetrazolium chloride (TTC) micromethod was performed as follow : in sterile 96 wells (250 µl) microplates, 20 µl of each sample were ten fold serialy diluted in 180 µl of Letheen broth (LB) (Difco, ref 268110) containing 1.5% Tween 80 (Sigma, ref P1754) and TTC (Sigma, ref T8877) For the strain *Candida albicans*, Sabouraud broth was used instead of LB. The microplates were incubated 48 h at 32,5°C and the microorganisms growth was monitored as color change from colorless to pink/red. The reciprocal highest dilution indicating growth allows the determination of the log number of each microorganism at each time.

---

## Organic & Enzymatic synthesis

### Experimental procedures

---

#### **Chromatography (TLC/FC)**

Thin Layer Chromatography (TLC) analysis were performed on E. Merck 60 F<sub>254</sub> Silica Gel non activated plates. Eluents used were a mixture of dichloromethane-methanol, petroleum ether-ethyl acetate and acetonitrile-water. Ratios are described in experimental part. Compounds were visualized by UV ( $\lambda = 254\text{nm}$ ) and/or spraying with orcinol stain (5 % H<sub>2</sub>SO<sub>4</sub> in ethanol solution containing 5 % orcinol (w/v)) or vanillin stain (1 % H<sub>2</sub>SO<sub>4</sub> in ethanol solution containing 5 % vanillin (w/v)) and followed by charring at 150 °C. Column Flash Chromatography were conducted using Silica gel 60 (230-400 mesh, 40-63  $\mu\text{m}$ ).

#### **Optical rotation**

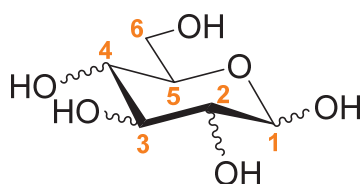
Specific rotation was measured on a Model 341 Perkin-Elmer polarimeter using a sodium lamp ( $\lambda = 589\text{ nm}$ ) at 20 °C in a 1 dm cell.

#### **Nuclear Magnetic Resonance (NMR)**

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on an Avance II Bruker spectrometer at 250 MHz (<sup>13</sup>C, 62.9 MHz) or an Avance III HD NanoBay Bruker at 400 MHz (<sup>13</sup>C: 100 MHz). Chemical shifts ( $\delta$ ) are given in parts per million from tetramethylsilane (TMS) as internal standard. The following abbreviations are used for the proton spectra multiplicities : s : singlet, d : doublet, t : triplet, q : quartet, m : multiplet, b : broad, dd : doublet of doublet.

##### General carbon numbering:

Carbon numbering of carbohydrates were done according to the general method and remembered below.



**Carbohydrates scaffold**

### **HRMS**

Electronic ionspray methodology was used to record mass spectra. HRMS spectra were recorded on a Maxis Bruker 4G.

### **HPLC**

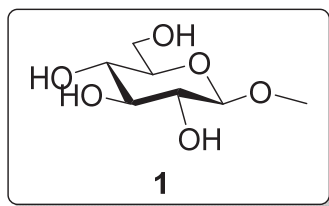
Chromatographic analysis were carried out on the Waters 2690 separation module (Waters) equipped with the Waters 2996 Photodiode Array (PDA) Dectector. HPLC separation was performed with an Alltima C18 column (150x 4.6 mm; 5 $\mu$ ). Separation was done at a flow rate of 1 mL/min. The mobile phase employed for the separation was composed of water (solvent A) and acetonitrile (solvent B) both acidified with 0.1% formic acid and it was used according to the following elution gradient 0-2 min: 5% B; 2-25 min: 5-35% B; 25-29 min: 35-70% B; 29-37 min: 70-5% B. The injection volume was 20  $\mu$ L. The diode-array detector recorded the UV spectra from 190 to 300 nm. The acquisition and data processing were done with the software *Millenium*.

## Alkyl O-glycoside synthesis

### General Procedure 1 :

To a solution of pNP- $\beta$ -D-glycopyranoside (1.0 equiv.) in citric acid/ sodium phosphate buffer (20 mM, pH 6, 2 mL) were added glycoside acceptors (50 equiv.). The reaction mixture was stirred and warmed up to 37°C, the wild type enzyme DtGly (0.35 mg.mL<sup>-1</sup>, 20 nmol, 3 mL) was added and the reaction was monitored by TLC (ACN/ H<sub>2</sub>O, 6:1) until full disappearance of glycoside donor (typically 8h). Then the mixture was concentrated under reduced pressure and purified.



**Methyl 1-O-β-D-glucopyranoside (1)** (Katsumi Kurashima, 2003)

Exact Mass: 194,08

MW: 194,18 g.mol<sup>-1</sup>

The title compound **1** was prepared from pNP-β-D-glucopyranoside (20.7 mg, 69 μmol) according the general procedure 1. The expected product was obtained after a flash chromatography (DCM/MeOH, 9:1) as uncoloured oil (8.3 mg; 54 %),  $R_f = 0.43$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -24$  (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.4 (d,  $J_{1/2} = 8$  Hz, 1H, H-1); 3.95 (dd,  $J_{5/6'} = 1.2$  Hz,  $J_{6'/6''} = 12.4$  Hz, 1H, H-6'); 3.75 (dd,  $J_{5/6''} = 6$  Hz,  $J_{6'/6''} = 11.6$  Hz, 1H, H-6''); 3.6 (s, 3H, OCH<sub>3</sub>); 3.5 (m, 2H, H-3, H-5); 3.4 (bt,  $J_{3/4} = J_{4/5} = 9.6$  Hz, 1H, H-4); 3.29 (bt,  $J_{1/2} = 8,9$  Hz, 1H, H-2), assignments were confirmed by gCOSY.

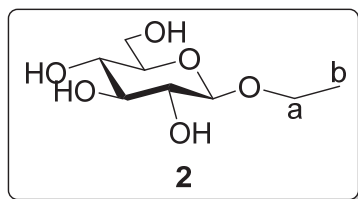
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 103,2 (C-1), 75,9 (C-5), 75,8 (C-3), 73,1 (C-2), 69,7 (C-4), 60,8 (C-6), 57,2 (OCH<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>     $\text{C}_7\text{H}_{14}\text{NaO}_6$      $m/z$  theoretical : 217.0682

$m/z$  measured : 217.0678

[M + H]<sup>+</sup>     $\text{C}_7\text{H}_{15}\text{O}_6$      $m/z$  theoretical : 195.0863

$m/z$  measured : 195.0859

**Ethyl 1-O-β-D-glucopyranoside (2)** (Katsumi Kurashima, 2003)

Exact Mass: 208,09

MW: 208,21 g.mol<sup>-1</sup>

Title compound **2** was obtained from pNP-β-D-glucopyranoside (20.0 mg, 66 μmol) following the general procedure 1. A flash chromatography (DCM/MeOH, 9:1) afforded the expected compound as colourless oil (7.9 mg; 57 %),  $R_f = 0.51$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -34,3$  ( $c = 1$ , H<sub>2</sub>O).

Analytica data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta = 4.5$  (d,  $J_{1/2} = 8$  Hz, 1H, H-1); 4.0 (m, 1H, OC<sub>a</sub>H'); 3.94 (dd,  $J_{5/6} = 4$  Hz,  $J_{6/6''} = 12.4$  Hz, 1H, H-6'); 3.75 (m, 2H, OC<sub>a</sub>H'', H-6''); 3.5 (m, 2H, H-3, H-5); 3.4 (bt,  $J_{3/4} = J_{4/5} = 9.2$  Hz, 1H, H-4); 3.29 (bt,  $J_{2/3} = 9.2$  Hz, 1H, H-2); 1.26 (bt,  $J_{a/b} = 8$  Hz, 3H, OCH<sub>2</sub>C<sub>b</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

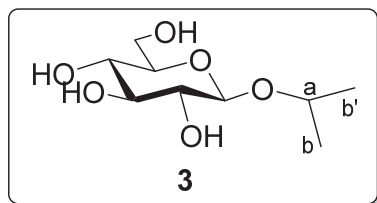
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta = 101.9$  (C-1); 75.9 (C-5), 75.8 (C-3); 73.1 (C-2); 69.7 (C-4); 66,2 (OC<sub>a</sub>H<sub>2</sub>CH<sub>3</sub>); 60. 8 (C-6); 14.3 (OCH<sub>2</sub>C<sub>b</sub>H<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI+) : [M + Na]<sup>+</sup>      C<sub>8</sub>H<sub>16</sub>NaO<sub>6</sub>       $m/z$  theoretical : 231.0839

$m/z$  measured : 231.0837

[M + H]<sup>+</sup>      C<sub>8</sub>H<sub>17</sub>O<sub>6</sub>       $m/z$  theoretical : 209.1019

$m/z$  measured : 209.1021

**Isopropyl 1-O-β-D-glucopyranoside (3) (Li, 2003)**

$C_9H_{18}O_6$   
 Exact Mass: 222,11  
 MW: 222,24 g.mol<sup>-1</sup>

Title compound **3** was synthesized from pNP-β-D-glucopyranoside (24.0 mg, 80 μmol) following the general procedure 1. The desired product was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (5.9 mg; 33 %).  $R_f$  = 0.64 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = -26 (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.56 (d,  $J_{1/2}$  = 8 Hz, 1H, H-1); 4.14 (m, 1H, OC<sub>a</sub>H); 3.93 (bdd,  $J_{6'/6''}$  = 12.3 Hz, 1H, H-6'); 3.73 (dd,  $J_{5/6''}$  = 2 Hz,  $J_{6'/6''}$  = 12.3 Hz, 1H, H-6''); 3.5 (m, 2H, H-3, H-5); 3.4 (bt,  $J_{3/4}$  =  $J_{4/5}$  = 9.6 Hz, 1H, H-4); 3.24 (bt,  $J_{1/2}$  = 8 Hz, 1H, H-2); 1.27 (d,  $J_{a/b}$  = 6 Hz, 3H, OCHC<sub>b</sub>H<sub>3</sub>); 1.24 (d,  $J_{a/b'}$  = 6 Hz, 3H, OCHC<sub>b'</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

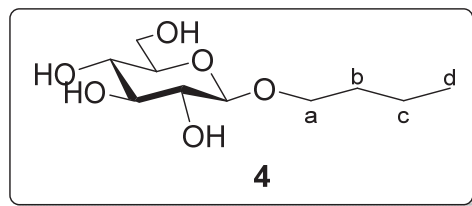
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 100 (C-1); 76 (C-3, C-5); 73.2 (C-2); 73 (OC<sub>a</sub>HCH<sub>3</sub>); 69.7 (C-4); 60.8 (C-6); 22 (OCH<sub>2</sub>C<sub>b</sub>H<sub>3</sub>); 21 (OCH<sub>2</sub>C<sub>b'</sub>H<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>      $C_9H_{18}NaO_6$       $m/z$  theoretical : 245.0996

$m/z$  measured : 245.0994

[M + H]<sup>+</sup>      $C_9H_{19}O_6$       $m/z$  theoretical : 223.1176

$m/z$  measured : 223.1177

**Butyl 1-O-β-D-glucopyranoside (4)** (Katsumi Kurashima, 2003)

$C_{10}H_{20}O_6$   
 Exact Mass: 236,13  
 MW: 236,26 g.mol<sup>-1</sup>

The title compound **4** was prepared from pNP-β-D-glucopyranoside (20.0 mg, 66 μmol) according the general procedure 1. The expected product was obtained after a flash chromatography (DCM/MeOH, 9:1) as a colourless oil (10 mg; 64 %).  $R_f$  = 0.63 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = -35,6(c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.5 (d,  $J_{1/2}$  = 8 Hz, H-1); 3.95 (m, 2H, OC<sub>a</sub>H'CH<sub>2</sub>, H-6'); 3.73 (m, 2H, OC<sub>a</sub>H''CH<sub>2</sub>, H-6''); 3.5 (m, 2H, H-3, H-5); 3.4 (bt,  $J_{3/4}$  = 9.2 Hz, 1H, H-4); 3.28 (bt,  $J_{1/2}$  = 8 Hz,  $J_{2/3}$  = 9.2 Hz, 1H, H-2); 1.63 (m, 2H, CH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>CH<sub>2</sub>); 1.4 (m, 2H, CH<sub>2</sub>C<sub>c</sub>H<sub>2</sub>CH<sub>3</sub>); 0.93 (bt,  $J_{c/d}$  = 7.6 Hz, 3H, CH<sub>2</sub>C<sub>d</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

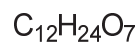
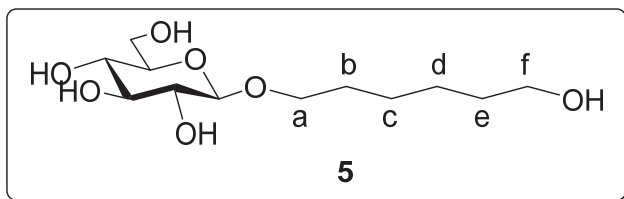
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 102.2 (C-1); 75.9-75.8 (C-3, C-5); 73.2 (C-2); 70.4 (OC<sub>a</sub>H<sub>2</sub>CH<sub>2</sub>); 69.7 (C-4); 60.8 (C-6); 30.9 (C<sub>b</sub>H<sub>2</sub>CH<sub>2</sub>); 18.45 (C<sub>c</sub>H<sub>2</sub>CH<sub>3</sub>); 13.0 (CH<sub>2</sub>C<sub>d</sub>H<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>     C<sub>10</sub>H<sub>20</sub>NaO<sub>6</sub>      $m/z$  theoretical : 259.1152

$m/z$  measured : 259.1151

[M + H]<sup>+</sup>     C<sub>10</sub>H<sub>21</sub>O<sub>6</sub>      $m/z$  theoretical : 237.1332

$m/z$  measured : 237.1332

**Hexan-6-ol-1-O-β-D-Glucopyranoside (5)** (K. Kurashima, 2004)

Exact Mass: 280,15

MW: 280,32 g.mol<sup>-1</sup>

The title compound **5** was synthesized from pNP-β-D-glucopyranoside (20.0 mg, 66 μmol) according the general procedure 1. The expected compound was obtained after a flash chromatography (DCM/MeOH, 9:1) as colorless oil (15,6 mg; 86 %).  $R_f = 0.49$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -32,1$  (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.48 (d,  $J_{1,2} = 8\text{Hz}$ , 1H, H-1); 3.94 (m, 2H, H-6', OC<sub>a</sub>H'CH<sub>2</sub>); 3.74 (m, 2H, H-6'', OC<sub>a</sub>H''CH<sub>2</sub>); 3.63 (t,  $J_{e,f} = 6.8\text{Hz}$ , 3H, C<sub>f</sub>H<sub>2</sub>OH); 3.49 (m, 2H, H-3, H-5); 3.42 (bt,  $J_{3,4} = J_{4,5} = 9.6\text{Hz}$ , 1H, H-4); 3.28 (d  $J_{1,2} = 8\text{Hz}$ , 1H, H-2); 1.65 (m, 2H, CH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>); 1.59 (m, 2H, CCH<sub>2</sub>C<sub>e</sub>CH<sub>2</sub>); 1.40 (m, 4H, C<sub>c</sub>H<sub>2</sub>C<sub>d</sub>CH<sub>2</sub>), assignments were confirmed by gCOSY.

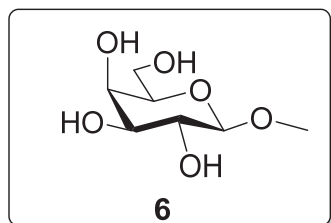
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 102.3 (C-1); 76 (C-5); 75.9 (C-3); 73.2 (C-2); 70.6 (-OC<sub>a</sub>H<sub>2</sub>); 69.8 (C-4); 61.8 (C<sub>f</sub>H<sub>2</sub>OH); 60.9 (C-6); 31.3 (C<sub>e</sub>H<sub>2</sub>CH<sub>2</sub>OH); 28.8 (CH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>); 24.8 (C<sub>c</sub>H<sub>2</sub>C<sub>d</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>      C<sub>12</sub>H<sub>24</sub>NaO<sub>7</sub>       $m/z$  theoretical : 303.1414

$m/z$  measured : 303.1414

[M + H]<sup>+</sup>      C<sub>12</sub>H<sub>25</sub>O<sub>7</sub>       $m/z$  theoretical : 281.1595

$m/z$  measured : 281.1592

**Methyl 1-O-β-D-galactopyranoside (6)** (Hancock, 2005)

$C_7H_{14}O_6$   
 Exact Mass: 194,08  
 MW: 194,18 g.mol<sup>-1</sup>

The title compound **6** was prepared from pNP-β-D-galactopyranoside (20 mg, 66 μmol) according the general procedure 1. The desired product was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (7.4 mg; 58 %).  $R_f = 0.38$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -2$  (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.35 (d,  $J_{1,2} = 8$  Hz, 1H, H-1); 3.95 (d,  $J_{3,4} = 3.2$  Hz, 1H, H-4); 3.80 (m, 2H, H-6); 3.73 (m, 1H, H-5); 3.7 (dd,  $J_{3,4} = 3.6$  Hz,  $J_{2,3} = 10$  Hz, 1H, H-3); 3.6 (s, 3H, OCH<sub>3</sub>); 3.53 (dd,  $J_{1,2} = 8$  Hz,  $J_{2,3} = 10$  Hz, 1H, H-2), assignments were confirmed by gCOSY.

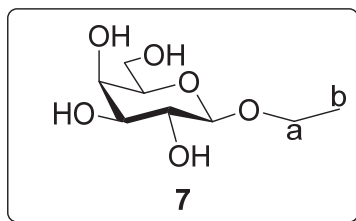
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 103.8 (C-1); 75.2 (C-5), 72.8 (C-3); 70.7 (C-2); 68.7 (C-4); 61 (C-6); 57.2 (OCH<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>      $C_7H_{14}NaO_6$       $m/z$  theoretical : 217.0683

$m/z$  measured : 217.0680

[M + H]<sup>+</sup>      $C_7H_{15}O_6$       $m/z$  theoretical : 195.0863

$m/z$  measured : 195.0863

**Ethyl 1-O- $\beta$ -D-galactopyranoside (7)** (Stevenson, 1993)

$C_8H_{16}O_6$   
 Exact Mass: 208,09  
 MW: 208,21 g.mol<sup>-1</sup>

The title compound **7** was synthesized starting from pNP- $\beta$ -D-galactopyranoside (20 mg, 66  $\mu$ mol) according the general procedure 1. The desired product was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (9.7 mg; 57 %).  $R_f$  = 0.46 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = -6,3 (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.44 (d,  $J_{1/2}$  = 8 Hz, 1H, H-1); 4.0 (m, 1H, OC<sub>a</sub>H<sub>2</sub>'CH<sub>3</sub>); 3.83 (bdd,  $J_{3/4}$  = 3.2 Hz, 1H, H-4); 3.76 (m, 4H, OC<sub>a</sub>H<sub>2</sub>"CH<sub>3</sub>, H-6, H-5); 3.66 (dd,  $J_{3/4}$  = 3.2 Hz,  $J_{2/3}$  = 8 Hz, 1H, H-3); 3.53 (bt,  $J_{1/2}$  =  $J_{2/3}$  = 8 Hz, 1H, H-2); 1.26 (bt,  $J_{a/b}$  = 8 Hz, 3H, OCH<sub>2</sub>C<sub>b</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

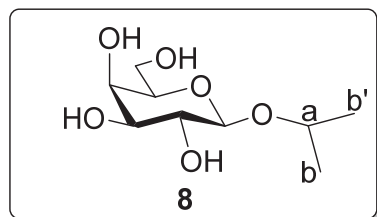
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 102.4 (C-1); 75.1 (C-5), 72.8 (C-3); 70.6 (C-2); 68.6 (C-4); 66.1 (OCH<sub>2</sub>CH<sub>3</sub>); 60.9 (C-6); 14.2 (OCH<sub>2</sub>CH<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI+) : [M + Na]<sup>+</sup>      $C_8H_{16}NaO_6$       $m/z$  theoretical : 231.0839

$m/z$  measured : 231.0836

[M + H]<sup>+</sup>      $C_8H_{17}O_6$       $m/z$  theoretical : 209.1019

$m/z$  measured : 209.1018

**Isopropyl 1-O-β-D-galactopyranoside (8) (Ooi, 1985)**

$C_9H_{18}O_6$   
 Exact Mass: 222,11  
 MW: 222,24 g.mol<sup>-1</sup>

The title compound **8** was prepared from pNP-β-D-galactopyranoside (20 mg, 66 μmol) according the general procedure 1. The desired product was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (4,4mg; 31%).  $R_f = 0.48$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -19$  (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.50 (d,  $J_{1/2} = 7.9$  Hz, 1H, H-1 $\alpha$ ); 4.14 (m, 1H, OC<sub>a</sub>H); 3.95 (d,  $J_{3,4} = 3.1$  Hz, 1H, H-4); 3.86-3.63 (m, 4H, H-3, H-5, H-6); 3.49 (bt,  $J_{2,3} = 8.9$  Hz, 1H, H-2); 1.28 (d,  $J_{a/b} = 6$  Hz, 3H, C<sub>b</sub>H<sub>3</sub>); 1.24 (d,  $J_{a/b'} = 6$  Hz, 3H, C<sub>b'</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 101 (C-1); 75 (C-5); 73 (OC<sub>a</sub>H); 72.9 (C-3); 70.8 (C-2); 69 (C-4); 60.9 (C-6); 22.4 (C<sub>b</sub>H<sub>3</sub>); 21 (C<sub>b'</sub>H<sub>3</sub>), assignments were confirmed by gHSQC.

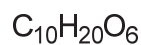
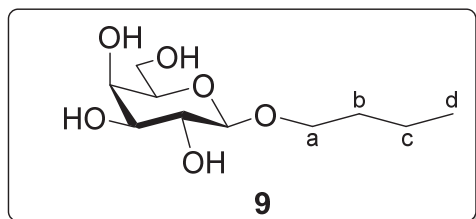
HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>      $C_9H_{18}NaO_6$       $m/z$  theoretical : 245.0996

$m/z$  measured : 245.0994

[M + H]<sup>+</sup>      $C_9H_{19}O_6$       $m/z$  theoretical : 223.1176

$m/z$  measured : 223.1175



**Butyl 1-O- $\beta$ -D-galactopyranoside (9) (Ooi, 1985)**

Exact Mass: 236,13

MW: 236,26 g.mol<sup>-1</sup>

The title compound **9** was prepared from pNP- $\beta$ -D-galactopyranoside (16.8 mg, 61  $\mu$ mol) according the general procedure 1. The desired product was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (11.4 mg; 79%).  $R_f$  = 0.61 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = +2 (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.41 (d,  $J_{1/2}$  = 8 Hz, H-1); 3.96 (m, 2H, OC<sub>a</sub>H'CH<sub>2</sub>, H-4); 3.79 (m, 2H, H-6); 3.7 (m, 3H, OC<sub>a</sub>H''CH<sub>2</sub>, H-3, H-5); 3.53 (bt,  $J_{1/2}$  = 8 Hz,  $J_{2/3}$  = 9.2 Hz, 1H, H-2); 1.63 (m, 2H, OCH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>CH<sub>2</sub>); 1.4 (m, 2H, C<sub>c</sub>H<sub>2</sub>CH<sub>3</sub>); 0.93 (t,  $J_{c/d}$  = 7.6 Hz, 3H, CH<sub>2</sub>C<sub>d</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

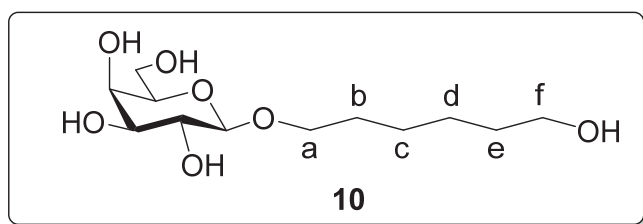
<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O),  $\delta$  = 102.8 (C-1); 75.1 (C-5), 72.9 (C-3); 70.9 (C-2); 70.4 (OC<sub>a</sub>H<sub>2</sub>CH<sub>2</sub>); 68.7 (C-4); 60.9 (C-6); 31.0 (OCH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>); 18.7 (C<sub>c</sub>H<sub>2</sub>CH<sub>3</sub>); 13.2 (CH<sub>2</sub>C<sub>d</sub>H<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>    C<sub>10</sub>H<sub>20</sub>NaO<sub>6</sub>     $m/z$  theoretical : 259.1152

$m/z$  measured : 259.1151

[M + H]<sup>+</sup>    C<sub>10</sub>H<sub>21</sub>O<sub>6</sub>     $m/z$  theoretical : 237.1333

$m/z$  measured : 237.1331

**6-Hydroxyhexyl 1-O-β-D-galactopyranoside (10)** (Stevenson, 1993)

$C_{12}H_{24}O_7$   
 Exact Mass: 280,15  
 MW: 280,32 g.mol<sup>-1</sup>

The title compound **10** was synthesized from pNP-β-D-galactopyranoside (21.0 mg, 66 μmol) according the general procedure 1. The expected compound was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (19.5 mg; 99 %).  $R_f = 0.43$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = +5$  (c = 1, H<sub>2</sub>O).

Analytical data :

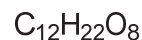
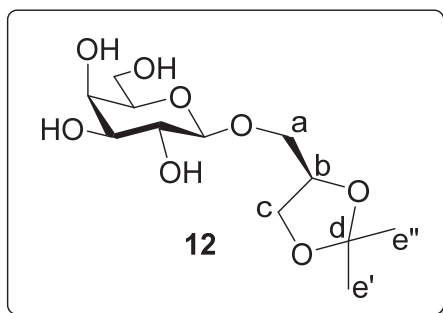
<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.43 (d,  $J_{1/2} = 8$  Hz, 1H, H-1); 3.96 (m, 2H, H-4, OC<sub>a</sub>H'CH<sub>2</sub>); 3.80 (m, 2H, H-6); 3.71 (m, 3H, H-3, H-5, OC<sub>a</sub>H''CH<sub>2</sub>); 3.64 (t,  $J_{e/f} = 6.8$  Hz, 2H, CH<sub>2</sub>C<sub>f</sub>H<sub>2</sub>OH); 3.53 (bt,  $J_{3/4} = 9.6$  Hz, 1H, H-4); 3.28 (bt,  $J_{2/3} = 9.6$  Hz, 1H, H-2); 1.68 (m, 2H, CH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>); 1.59 (m, 2H, CH<sub>2</sub>C<sub>e</sub>H<sub>2</sub>); 1.41 (m, 4H, C<sub>c</sub>H<sub>2</sub>C<sub>d</sub>H<sub>2</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 103 (C-1); 75 (C-5); 73 (C-3); 71 (C-2); 70 (OC<sub>a</sub>H<sub>2</sub>CH<sub>2</sub>); 69 (C-4); 62 (CH<sub>2</sub>OH); 61 (C-6); 31 (C<sub>e</sub>H<sub>2</sub>CH<sub>2</sub>OH); 29 (OCH<sub>2</sub>C<sub>b</sub>H<sub>2</sub>); 25 (C<sub>c</sub>H<sub>2</sub>C<sub>d</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>     $C_{12}H_{24}NaO_7$      $m/z$  theoretical : 303.1414  
 $m/z$  measured : 303.1413

[M + H]<sup>+</sup>     $C_{12}H_{25}O_7$      $m/z$  theoretical : 281.1595  
 $m/z$  measured : 281.1593



**1,2-Isopropylidene 1-O-( $\beta$ -D-galactopyranosyl)-(S)-glycerol (**12**)**

Exact Mass: 294,13

MW: 294,30 g.mol<sup>-1</sup>

The title compound **12** was synthesized from pNP- $\beta$ -D-galactopyranoside (27.4 mg, 91  $\mu$ mol) according the general procedure 1. The expected compound was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (19.7 mg; 73 %).  $R_f$  = 0,48 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = +5 ( $c$  = 1, H<sub>2</sub>O).

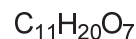
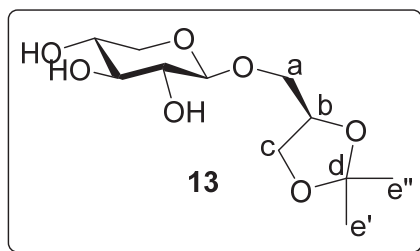
Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.50 (m, 1H, OCH<sub>2</sub>C<sub>b</sub>H); 4.45 (d,  $J_{1/2}$  = 8Hz, 1H, H-1); 4.2 (t,  $J_{b/c}$  = 8 Hz, 1H, CHC<sub>c</sub>H'O); 3.98 (m, 2H, H-4, OC<sub>a</sub>H'CH); 3.87 (bt,  $J_{b/c''}$  = 8 Hz, 1H, CHC<sub>c</sub>H''O); 3.8 (m, 3H, H-6, OC<sub>a</sub>H'CH); 3.7 (m, 2H, H-3, H-5); 3.56 (bt,  $J_{2/3}$  = 6.4 Hz, 1H, H-2); 1.5 (s, 3H, C<sub>e</sub>H<sub>3</sub>); 1.43 (s, 3H, C<sub>e'</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 110 (CH<sub>3</sub>C<sub>d</sub>CH<sub>3</sub>); 103 (C-1); 75 (C-5); 74 (OCH<sub>2</sub>C<sub>b</sub>HCH<sub>2</sub>); 73 (C-3); 71 (C-2); 70 (OC<sub>a</sub>H<sub>2</sub>CH); 69 (C-4); 65 (CHC<sub>c</sub>H<sub>2</sub>O); 61 (C-6); 26 (C<sub>e</sub>H<sub>3</sub>); 24 (C<sub>e'</sub>H<sub>3</sub>), assignments were confirmed by gHSQC and gHMBC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>    C<sub>12</sub>H<sub>22</sub>NaO<sub>8</sub>     $m/z$  theoretical : 317.1207  
 $m/z$  measured : 317.1208

[M + H]<sup>+</sup>    C<sub>12</sub>H<sub>23</sub>O<sub>8</sub>     $m/z$  theoretical : 295.1387  
 $m/z$  measured : 295.1396

**1,2-Isopropylidene 1-O-( $\beta$ -D-xylopyranosyl)-(S)-glycerol (**13**)**

Exact Mass: 264,12

MW: 264,27 g.mol<sup>-1</sup>

The title compound **13** was synthesized from pNP- $\beta$ -D-xylopyranoside (21.3 mg, 114  $\mu$ mol) according the general procedure 1. The expected compound was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (10.6 mg; 51 %).  $R_f$  = 0.68 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = -32 ( $c$  = 1, MeOH).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.48 (m, 1H, OCH<sub>2</sub>C<sub>b</sub>H); 4.45 (d,  $J_{1/2}$  = 8 Hz, 1H, H-1); 4.19 (t,  $J_{b/c}$  = 8 Hz, 1H, CHC<sub>c</sub>H'O); 3.98 (dd,  $J_{4/5}$  = 5.2 Hz,  $J_{5'/5''}$  = 11.6 Hz, 1H, H-5'); 3.92 (bdd,  $J_{a'/b}$  = 6.4 Hz,  $J_{a'/a''}$  = 10.8 Hz, 1H, OC<sub>a</sub>H'CH); 3.89 (bt,  $J_{c''/d}$  = 6.8 Hz, 1H, CHC<sub>c</sub>H''O); 3.79 (bdd,  $J_{a''/b}$  = 3.6 Hz;  $J_{a''/a''}$  = 10.8 Hz, 1H, OC<sub>a</sub>H<sub>2''</sub>CH); 3.65 (m, 1H, H-4); 3.46 (t,  $J_{2/3}$  = 9.2 Hz, 1H, H-3); 3.34 (m, 2H, H-5'', H-2), 1.49 (s, 3H, C<sub>e'</sub>H<sub>3</sub>); 1;42 (s, 3H, C<sub>e''</sub>H<sub>3</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 110 (CH<sub>3</sub>-C<sub>d</sub>-CH<sub>3</sub>); 103 (C-1); 76 (C-3); 74 (OCH<sub>2</sub>-C<sub>b</sub>H-); 73 (C-2); 70 (OC<sub>a</sub>H<sub>2</sub>-CH-); 69 (C-4); 65.4 (-CH-C<sub>c</sub>H<sub>2</sub>O); 65 (C-5); 26 (-C<sub>e'</sub>H<sub>3</sub>); 24 (-C<sub>e''</sub>H<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>    C<sub>11</sub>H<sub>20</sub>NaO<sub>7</sub>     $m/z$  theoretical : 287.1101  
 $m/z$  measured : 287.1102

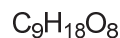
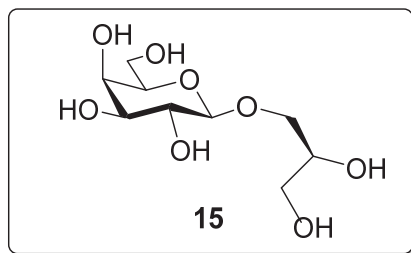
[M + H]<sup>+</sup>    C<sub>11</sub>H<sub>21</sub>O<sub>7</sub>     $m/z$  theoretical : 265.1282  
 $m/z$  measured : 265.1282

## Glyceride-O-glycosides synthesis

### General Procedure 2 :

1,2-Isopropylidene-1-O-( $\beta$ -D-glycopyranosyl)-glycerol was dissolved in a solution of acetic acid (70%) in water (3ml). The reaction was warmed up to 60 °C and left to stir until full deprotection was evident by TLC (~ 3h). Then the mixture was concentrated, co-evaporated with toluene (3 times) and purified



**1-O-(β-D-galactopyranosyl)-(S)-glycerol (15) (Tyl, 2004)**

Exact Mass: 254,10

MW: 254,24g.mol<sup>-1</sup>

The title compound **15** was prepared from 1,2-Isopropylidene-1-O-(β-D-galactopyranosyl)-(S)-glycerol **12** (20.6 mg, 70 μmol) according the general procedure 2. The expected product was obtained after a flash chromatography (DCM/MeOH, 9:1) as colourless oil (8.9 mg, 50 %),  $R_f = 0.18$  (ACN/H<sub>2</sub>O; 6:1),  $[\alpha]^{20}_D = +7$  (c = 1, H<sub>2</sub>O).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.44 (d,  $J_{1/2} = 8$  Hz, 1H, H-1 $\alpha$ ); 3.95 (m, 3H, H-4, OCH<sub>2</sub>C<sub>b</sub>H, CHC<sub>c</sub>H'OH); 3.86-3.77 (m, 3H, H-6, CHC<sub>c</sub>H''OH); 3.75-3.62 (m, 3H, H-3, H-5, OC<sub>a</sub>H<sub>2</sub>); 3.58 (bt,  $J_{1/2} = 8$  Hz, 1H, H-2), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 103 (C-1); 75 (C-5); 72.6 (C-3); 70.8 (C<sub>b</sub>HOH); 70.7 (C<sub>c</sub>H<sub>2</sub>OH); 70.4 (C-2); 68.6 (C-4); 62.3 (C-6); 61 (OC<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

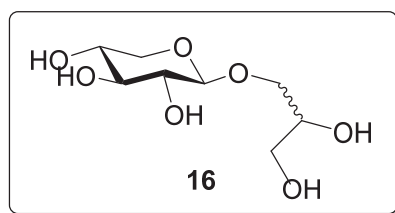
HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>      C<sub>9</sub>H<sub>18</sub>NaO<sub>8</sub>       $m/z$  theoretical : 277.0894

$m/z$  measured : 277.0894

[M + H]<sup>+</sup>      C<sub>9</sub>H<sub>19</sub>O<sub>8</sub>       $m/z$  theoretical : 255.1074

$m/z$  measured : 255.1075



**1-O-( $\beta$ -D-xylopyranosyl)-(S)-glycerol (**16**)**

$C_8H_{16}O_7$   
 Exact Mass: 224,09  
 MW: 224,21 g.mol<sup>-1</sup>

The title compound **16** was prepared from 1,2-Isopropylidene-1-O-( $\beta$ -D-xylopyranosyl)-(S)-glycerol **13** (15.2 mg, 57.6  $\mu$ mol) according the general procedure 2. The expected product was obtained after a flash chromatography (DCM/MeOH, 9:1) as uncoloured oil (8 mg, 62 %),  $R_f$  = 0.39 (ACN/H<sub>2</sub>O; 6:1),  $[\alpha]^{20}_D$  = - 38.2 (c = 1, MeOH).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 4.44 (d,  $J_{1/2}$  = 8 Hz, 1H, H-1 $\alpha$ ); 3.93 (m, 3H, H-5', OCH<sub>2</sub>C<sub>b</sub>H, CHC<sub>c</sub>H'OH); 3.76 (bdd,  $J_{b/c''}$  = 3.6 Hz,  $J_{c'/c''}$  = 10.4 Hz, 1H, CHC<sub>c</sub>H''OH); 3.64 (m, 3H, H-4, OC<sub>a</sub>H<sub>2</sub>); 3.47 (bt,  $J_{2/3}$  = 9.2 Hz, 1H, H-3); 3.33 (m, 2H, H-2, H-5''), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 103.2 (C-1); 75.6 (C-3); 72.9 (C-2); 70.7 (C<sub>c</sub>HOH); 70.3 (C<sub>b</sub>H<sub>2</sub>OH); 69.2 (C-4); 65 (C-5); 62.3 (OC<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

HRMS (ESI+) : [M + Na]<sup>+</sup>    C<sub>9</sub>H<sub>18</sub>NaO<sub>8</sub>     $m/z$  theoretical : 247.0788

$m/z$  measured : 247.0788

[M + H]<sup>+</sup>    C<sub>9</sub>H<sub>19</sub>O<sub>8</sub>     $m/z$  theoretical : 225.0969

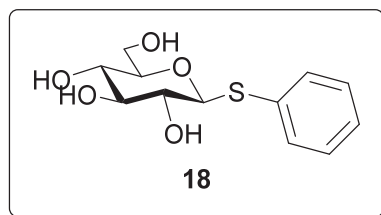
$m/z$  measured : 225.0967

## Enzymatic thioglycosylation

### General procedure 3 :

To a solution of pNP- $\beta$ -D-glucopyranoside (1.0 equiv.) in TrisBase/ Tris HCl buffer (20 mM, pH9, 2ml) were added successively 10 equiv. of glycoside acceptor and the thioglycosylase mutant DtGly E159Q (5.22 mg.ml<sup>-1</sup>, 49 nmol, 500  $\mu$ l). The reaction was allowed to stir at 37 °C overnight. TLC (ACN/ H<sub>2</sub>O, 6:1) revealed full disappearance of starting material, reaction mixture was concentrated under reduced pressure and purified.



**Phenyl 1-thio- $\beta$ -D-glucopyranoside (**18**) (Samoshin, 2014)**

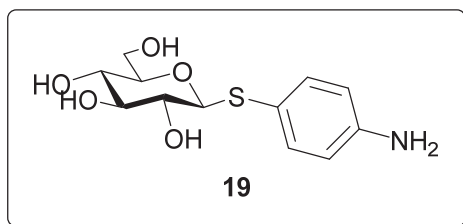
$C_{12}H_{16}O_5S$   
 Exact Mass: 272,07  
 MW : 272,32 g.mol<sup>-1</sup>

Titlle compound **18** was obtained using pNP- $\beta$ -D-glucopyranoside (0.03 mmol, 1.0 equiv.) and commercial thiophenol (0.3 mmol, 10 equiv.) as glycoside acceptor according the general procedure 3. The desired product was obtained after C<sub>18</sub> flash chromatography (H<sub>2</sub>O/MeOH, 100:0 – 50:50) as white solid (9 mg, 98 %),  $R_f$  = 0.3 (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D$  = - 70 (c = 1, H<sub>2</sub>O).

**Analytical data :**

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 7.55 (d,  $J_{Har}$  = 8 Hz, 2H, H<sub>ar</sub>); 7.31 (d,  $J_{Har}$  = 8 Hz, 2H, H<sub>ar</sub>); 7.22 (m, 1H, H<sub>ar</sub>); 4.57 (d,  $J_{1/2}$  = 10 Hz, 1H, H-1); 3.86 (m, 1H, H-6'); 3.67 (m, 1H, H-6''); 3.33 (m, 3H, H-3, H-4, H-5); 3.22 (bt,  $J_{1/2}$  = 10Hz; 1H, H-2), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 134,4 (C<sub>Ar</sub>); 134,3 (CH<sub>ar</sub>); 134 (C<sub>Ar</sub>); 130 (CH<sub>ar</sub>); 89 (C-1); 80 (C-5), 74 (C-3); 71 (C-2); 71 (C-4); 63 (C-6), assignments were confirmed by gHSQC.

**4-Aminophenyl 1-thio-β-D-glucopyranoside (19)** (Brachet, 2013)

$C_{12}H_{17}NO_5S$   
 Exact Mass: 287,08  
 MW : 287,33 g.mol<sup>-1</sup>

Titlle compound **19** was obtained using pNP-β-D-glucopyranoside (0.13 mmol, 1.0 equiv.) and commercial 4-aminothiophenol (1.6 mmol, 10 equiv.) as glycoside acceptor according the general procedure 3. The desired product was obtained after flash chromatography (DCM/MeOH, 9:1) as white solid (45.8 mg, 100 %),  $R_f$  = 0.33 (ACN/H<sub>2</sub>O, 6:1).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 7.44 (d,  $J_{Har}$  = 8 Hz, 2H, H<sub>ar</sub>); 6.85 (d,  $J_{Har}$  = 8 Hz, 2H, H<sub>ar</sub>); 4.57 (d,  $J_{1/2}$  = 12 Hz, 1H, H-1); 3.90 (dd,  $J_{5/6'}$  = 1,2 Hz,  $J_{6'/6''}$  = 16 Hz 1H, H-6'); 3.72 (dd,  $J_{5/6''}$  = 4 Hz,  $J_{6'/6''}$  = 16 Hz, 1H, H-6''); 3.51 (bt,  $J_{2/3}$  =  $J_{3/4}$  = 8 Hz, 1H, H-3); 3.44 (m, 2H, H-4, H-5); 3.28 (bt,  $J_{2/3}$  = 8Hz, 1H, H-2), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 147 (C<sub>Ar</sub>); 135 (CH<sub>ar</sub>); 119 (C<sub>Ar</sub>); 117 (CH<sub>ar</sub>); 88 (C-1); 80 (C-5), 77 (C-3); 72 (C-2); 69 (C-4); 61 (C-6), assignments were confirmed by gHSQC.

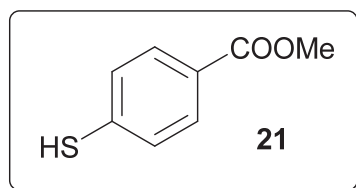
HRMS (ESI+) : [M + Na]<sup>+</sup>     $C_{12}H_{17}NNaO_5S$      $m/z$  theoretical : 310.0720  
 $m/z$  measured : 310.0720

[M + H]<sup>+</sup>     $C_{12}H_{18}NO_7S$      $m/z$  theoretical : 288.0900  
 $m/z$  measured : 288.0899

## Esterification of 4-Mercaptoaryl glycoside acceptors

### General Procedure 4 :

To a solution of commercial 4-Mercaptoaryl acid in methanol was added a catalytic amount of  $\text{H}_2\text{SO}_4$  conc (18 M). The reaction mixture was allowed to stir under reflux for 18 hours. Resulting solution was concentrated under reduced pressure then diluted with EtOAc (10 ml), wash with  $\text{H}_2\text{O}$  (3 x 15 ml) and  $\text{Na}_2\text{CO}_3$  sat (3 x 15 ml). Organic layers were combined, dried over  $\text{MgSO}_4$  and concentrated

**4-Mercaptobenzoic methyl ester (21)** (Novoa, 2014)

Exact Mass: 168,02

MW: 168,21 g.mol<sup>-1</sup>

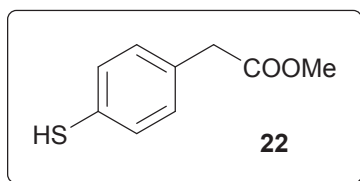
Titlle compound **21** was obtained following the general method 3 starting from commercial 4-Mercaptobenzoic acid (1.0 g, 6.5 mmol). The desired compound was obtained after flash chromatography (PE/EtOAc, 1:1) affording white solid (408,6 mg, 40 %),  $R_f = 0.58$  (EP/EtOAc, 2:1).

Analytical data :

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>),  $\delta$  = 7.9 (d,  $J_{\text{Har}} = 7.2$  Hz, 2H,  $\text{H}_{\text{ar}}$ ); 7.3 (d,  $J_{\text{Har}} = 7.2$  Hz, 2H,  $\text{H}_{\text{ar}}$ ); 3.9 (s, 3H,  $\text{OCH}_3$ ); 3.6 (s, 1H,  $\text{SH}$ ), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>),  $\delta$  = 167 ( $\text{COOCH}_3$ ); 138 ( $\text{C}_{\text{ar}}$ ); 130 ( $\text{CH}_{\text{ar}}$ ); 128 ( $\text{CH}_{\text{ar}}$ ); 127 ( $\text{C}_{\text{ar}}$ ); 52 ( $\text{OCH}_3$ ), assignments were confirmed by gHSQC.

HRMS (ESI+) :  $[\text{M} + \text{H}]^+$      $\text{C}_8\text{H}_9\text{O}_2\text{S}$  $m/z$  theoretical : 169.0318 $m/z$  measured : 169.0317

**4-Mercaptophenylacetic methyl ester (22)**

$C_9H_{10}O_2S$   
 Exact Mass: 182,04  
 MW: 182,24 g.mol<sup>-1</sup>

The desired compound **22** was obtained according the general procedure 4 starting from commercial 4-Mercaptophenylacetic acid (319 mg, 1.90 mmol). Flash chromatography (PE/EtOAc, 1:1) yielded tittle compound **22** as colourless oil (100 mg, 30 %),  $R_f$  = 0.48 (PE/EtOAc, 2:1).

Analytical data :

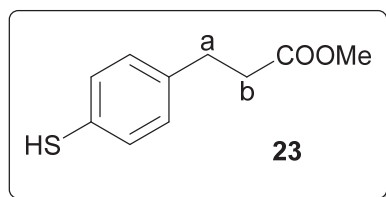
$^1H$  NMR (400 MHz,  $CDCl_3$ ),  $\delta$  = 7.25 (d,  $J_{Har}$  = 6.8 Hz, 2H,  $H_{ar}$ ); 7.16 (d,  $J_{Har}$  = 7.6 Hz, 2H,  $H_{ar}$ ); 3.7 (s, 3H,  $OCH_3$ ); 3.61 (s, 2H,  $CH_2$ );, assignments were confirmed by gCOSY.

$^{13}C$  NMR (125 MHz,  $CDCl_3$ ),  $\delta$  = 174 ( $COOCH_3$ ); 135 ( $C_{ar}$ ); 133 ( $C_{ar}$ ); 131 ( $CH_{ar}$ ); 130 ( $CH_{ar}$ ); 52 ( $OCH_3$ ). 41 (s, 2H,  $CH_2$ ), assignments were confirmed by gHSQC.

HRMS (ESI+) :  $[M + H]^+$      $C_{18}H_{19}O_4S_2$      $m/z$  theoretical : 363.0719  
 $m/z$  measured : 363.0720

HRMS (ESI+) :  $[M + Na]^+$      $C_{18}H_{18}NaO_4S_2$      $m/z$  theoretical : 385.0539  
 $m/z$  measured: 385.0539



**4-Mercaptohydrocinnamic methyl ester (23)** (Sasaki, 2011)

$C_{10}H_{12}O_2S$   
 Exact Mass: 196,06  
 MW: 196,26 g.mol<sup>-1</sup>

Title product **23** was recovered from commercial 4-Mercaptohydrocinnamic acid (300 mg, 1.65 mmol) following the general method 4. Flash chromatography (PE/EtOAc, 1:1) affording title compound **23** as white solid (300 mg, R = 93 %),  $R_f$  = 0.68 (PE/EtOAc, 2:1).

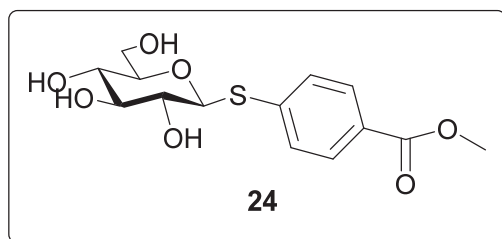
Analytical data :

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>),  $\delta$  = 7,22 (d,  $J_{\text{Har}}$  = 7,2 Hz, 2H, H<sub>ar</sub>); 7,11 (d,  $J_{\text{Har}}$  = 7,6 Hz, 2H, H<sub>ar</sub>); 3,7 (s, 3H, -OCH<sub>3</sub>); 2,89 (bt  $J_{a/b}$  = 7,6 Hz, 2H, C<sub>b</sub>H<sub>2</sub>); 2,64 (bt  $J_{a/b}$  = 7,6 Hz, 2H, C<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>),  $\delta$  = 175 (COOCH<sub>3</sub>); 130,3 (CH<sub>ar</sub>); 130 (CH<sub>ar</sub>); 129,5 (C<sub>ar</sub>); 52 (OCH<sub>3</sub>); 36,5 (C<sub>b</sub>H<sub>2</sub>); 31,3 (C<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + H]<sup>+</sup>    C<sub>20</sub>H<sub>23</sub>O<sub>4</sub>S<sub>2</sub>     $m/z$  theoretical : 391.1032  
 $m/z$  measured : 391.1032

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>    C<sub>20</sub>H<sub>22</sub>NaO<sub>4</sub>S<sub>2</sub>     $m/z$  theoretical : 413.0852  
 $m/z$  measured : 413.0850

**4-Thio-1-β-D-glucopyranoside benzoate methyl ester (24)**

$C_{14}H_{18}O_7S$   
 Exact Mass: 330,08  
 MW: 330,35 g.mol<sup>-1</sup>

Titlle compound **24** was obtained using pNP-β-D-glucopyranoside (0.11 mmol, 1.0 equiv.) and 4-Mercaptobenzoate methyl ester **21** (1 mmol, 10 equiv.) as glycoside acceptor according the general procedure 3. The desired product was obtained after flash chromatography (DCM/MeOH, 9:1) as colourless oil (34 mg, 90 %),  $R_f = 0.62$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -115$  ( $c = 1$ , MeOH).

Analytical data :

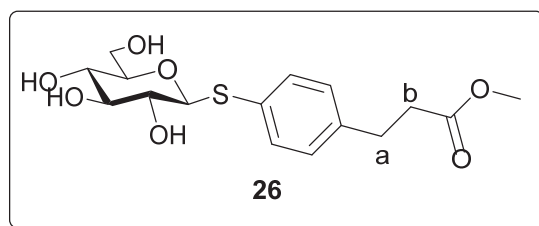
<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta = 7.93$  (d,  $J_{Har} = 7.6$  Hz, 2H, H<sub>ar</sub>); 7.59 (d,  $J_{Har} = 7.6$  Hz, 2H, H<sub>ar</sub>); 4.82 (d,  $J_{1/2} = 10$  Hz, 1H, H-1); 3.93 (bs, 1H, H-6'); 3.90 (s, 3H, OCH<sub>3</sub>); 3.71 (dd,  $J_{5/6''} = 5.2$ ,  $J_{6/6''} = 12$  Hz, 1H, H-6''); 3.38 (m, 4H, H-2, H-3, H-4, H-5), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta = 168$  (COOCH<sub>3</sub>); 143 (C<sub>Ar</sub>); 131 (CH<sub>ar</sub>); 130 (CH<sub>ar</sub>); 129 (C<sub>Ar</sub>); 88 (C-1); 82 (C-3), 79 (C-5); 74 (C-2); 71 (C-4); 63 (C-6); 53 (OCH<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>     $C_{14}H_{18}NaO_7S$      $m/z$  theoretical : 348.1111  
 $m/z$  measured : 348.1110

[M + H]<sup>+</sup>     $C_{14}H_{19}O_7S$      $m/z$  theoretical : 331.0846  
 $m/z$  measured : 331.0845



**4-Thio-1-β-D-glucopyranoside hydrocinnamic methyl ester (26)**

$C_{16}H_{22}O_7S$   
 Exact Mass: 358,11  
 MW: 358,41 g.mol<sup>-1</sup>

Titlle compound **26** was obtained using pNP-β-D-glucopyranoside (0.13 mmol, 1.0 equiv.) and 4-Mercaptohydrocinnamic methyl ester **23** (1 mmol, 8 equiv.) as glycoside acceptor according the general procedure 3. The desired product was obtained after flash chromatography (DCM/MeOH, 9:1) as colourless oil (40.6 mg, 87 %),  $R_f = 0.66$  (ACN/H<sub>2</sub>O, 6:1),  $[\alpha]^{20}_D = -122$  (c = 1, MeOH).

Analytical data :

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 7.52 (d,  $J_{\text{Har}} = 8$  Hz, 2H, H<sub>ar</sub>); 7.19 (d,  $J_{\text{Har}} = 8$  Hz, 2H, H<sub>ar</sub>); 4.58 (d,  $J_{1/2} = 9.6$  Hz, 1H, H-1); 3.88 (bdd,  $J_{6/6''} = 12$  Hz, 1H, H-6''); 3.69 (m, 4H, H-6'', OCH<sub>3</sub>); 3.37 (m, 3H, H-3, H-4, H-5); 3.23 (bt,  $J_{1/2} = 9.6$  Hz, 1H, H-2); 2.92 (t,  $J_{a/b} = 7.6$  Hz, 2H, C<sub>b</sub>H<sub>2</sub>); 2.65 (t,  $J_{a/b} = 7.6$  Hz, 2H, C<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 175 (COOCH<sub>3</sub>); 141 (C<sub>ar</sub>); 133 (CH<sub>ar</sub>); 132 (C<sub>ar</sub>); 130 (CH<sub>ar</sub>); 89 (C-1); 82 (C-5), 80 (C-3); 74 (C-2); 71 (C-4); 63 (C-6); 52 (OCH<sub>3</sub>); 35 (C<sub>b</sub>H<sub>2</sub>); 30 (C<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

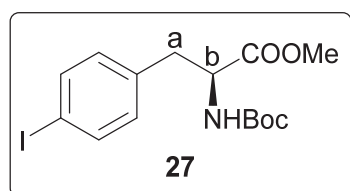
HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>    C<sub>15</sub>H<sub>20</sub>NaO<sub>7</sub>S     $m/z$  theoretical : 381.0978  
 $m/z$  measured : 381.0977

[M + H]<sup>+</sup>    C<sub>15</sub>H<sub>21</sub>O<sub>7</sub>S     $m/z$  theoretical : 359.1159  
 $m/z$  measured : 359.1158

## Thiotyrosine synthesis

### General method :

#### Methyl (S)-2-(tert-butoxycarbonylamino)-3-(4-iodophenyl) propanoate (**27**) (Meyer, 2010)



$C_{15}H_{20}INO_4$   
 Exact Mass: 405,04  
 MW: 405,23 g.mol<sup>-1</sup>

To a solution of *N*-(tert-butoxycarbonyl)-4-iodo-L-phenylalanine (1.28 mmol, 1 equiv.) and NaHCO<sub>3</sub> (2.88 mmol, 2.25 equiv.) in DMF (5ml) was added MeI (6.4 mmol, 5 equiv.). The mixture was allowed to stir at rt for 18 hours. The resulting mixture was poured into water (15mL) and extracted with EtOAc (3 x 10 ml). Organic layers were combined, washed with brine (3 x 10 ml), dried over MgSO<sub>4</sub> and concentrated. Crude extract was further purified through flash chromatography (hexanes/ EtOAc, 9:1) affording the title compound **27** as white solid (440 mg; 85 %), *R*<sub>f</sub> = 0.61 (PE/EtOAc, 2:1), [ $\alpha$ ]<sub>D</sub><sup>20</sup> = +103 (c = 1, EtOAc).

#### Analytical data :

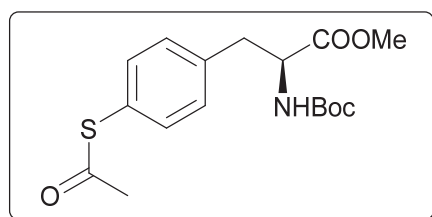
<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 7.61 (d, *J*<sub>Har</sub> = 7.8Hz, 2H, H<sub>ar</sub>); 6.87 (d, *J*<sub>Har</sub> = 7.8 Hz, 2H, H<sub>ar</sub>); 4.97 (d, *J* = 7.1 Hz, 1H, CHNH); 4.56 (m, 1H, CHNH); 3.71 (s, 3H, OCH<sub>3</sub>); 3.02 (m, 2H, CH<sub>2</sub>CH); 1.42 (s, 9H, 3CH<sub>3</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, D<sub>2</sub>O),  $\delta$  = 172.1 (COOCH<sub>3</sub>); 155.1 (COOC(CH<sub>3</sub>)<sub>3</sub>); 137.7 (CH<sub>ar</sub>); 135.9 (C<sub>ar</sub>); 131.5 (CH<sub>ar</sub>); 92.65 (C<sub>ar</sub>); 80.2 (C(CH<sub>3</sub>)<sub>3</sub>); 54.3 (NHCH); 52.5 (OCH<sub>3</sub>); 38.1 (CH<sub>2</sub>CH); 28.4 (C(CH<sub>3</sub>)<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>     $C_{15}H_{20}NaINO_4$     *m/z* theoretical : 428.0329  
*m/z* measured : 428.0326

[M + H]<sup>+</sup>     $C_{15}H_{21}INO_4$     *m/z* theoretical : 406.0509  
*m/z* measured : 406.0506

**Methyl (S)-2-(tert-butoxycarbonylamino)-3-(4-iodophenyl) propanoate (28)** (Soria-Castro, 2013)



$C_{17}H_{23}NO_5S$   
Exact Mass: 353,13  
MW: 353,43 g.mol<sup>-1</sup>

To a solution of Methyl (S)-2-(tert-butoxycarbonylamino)-3-(4-iodophenyl) propanoate **27** (0.14 mmol, 1.0 equiv.) in dry and degassed toluene (3 ml) was added under argon CuI (0.018 mmol, 10% mol.), 1,10-Phenanthroline (0.033 mmol, 20 % mol.) and potassium thioacetate KSCoCH<sub>3</sub> (0.2 mmol, 1. equiv.). The mixture was allowed to stir vigorously at 100 °C under Argon for 24 hours. The resulting mixture was cooled to room temperature. DCM (10 ml) and water (10 ml) were added and the mixture was stirred. The organic layer was separated and aqueous layer was extracted with DCM (3 x 10 ml). Organic layers were combined, washed with NaHCO<sub>3</sub>, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. Crude extract was adsorbed on silica and further purified through flash chromatography (hexanes/ EtOAc, 8:2) affording the title compound **28** as dark pink oil (22 mg; 25 %),  $R_f = 0.63$  (PE/EtOAc, 2:1),  $[\alpha]^{20}_D = +63$  (c = 1, EtOAc).

Analytical data :

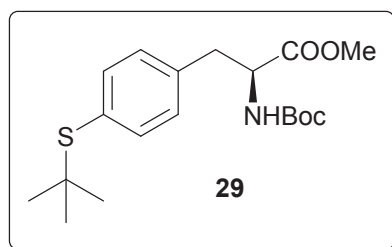
<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O),  $\delta$  = 7.34 (d,  $J_{Har} = 7.8$ Hz, 2H, H<sub>ar</sub>); 7.18 (d,  $J_{Har} = 7.6$  Hz, 2H, H<sub>ar</sub>); 4.99 (d,  $J = 7.2$  Hz, 1H, CHNH); 4.59 (bd,  $J = 6.4$  Hz, 1H, CHNH); 3.71 (s, 3H, OCH<sub>3</sub>); 3.10 (m, 2H, CH<sub>2</sub>CH); 2.41 (s, 3H, COCH<sub>3</sub>); 1.42 (s, 9H, 3CH<sub>3</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>),  $\delta$  = 194, 172, 155 (CO); 137.7 (C<sub>ar</sub>); 134.5 (CH<sub>ar</sub>); 130.1 (CH<sub>ar</sub>); 126.7 (C<sub>ar</sub>); 80.1 (C(CH<sub>3</sub>)<sub>3</sub>); 54.3 (NHCH); 52.3 (OCH<sub>3</sub>); 38.2 (CH<sub>2</sub>CH); 30.2 (COCH<sub>3</sub>); 28.4 (C(CH<sub>3</sub>)<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI+) : [M + Na]<sup>+</sup>     $C_{17}H_{23}NaNSO_5$      $m/z$  theoretical : 376.1189  
 $m/z$  measured : 376.1188

[M + H]<sup>+</sup>     $C_{17}H_{24}NSO_5$      $m/z$  theoretical : 354.1370  
 $m/z$  measured : 354.1371

**Methyl (S)-2-(tert-butoxycarbonylamino)-3-(4-tert-butylsulfanylphenyl) propanoate (29)**  
(Rudolf, 2013)



$C_{19}H_{29}NO_4S$   
Exact Mass: 367,18  
MW: 367,50 g.mol<sup>-1</sup>

The tris(di-benzylideneacetone)dipalladium(0)-chloroform adduct (24  $\mu$ mol, 0.019 equiv.) was added to dry DMF (4 ml) followed by 1,1'-Bis(diphenylphosphino)ferrocene (94  $\mu$ mol, 0.073 equiv.) and the mixture was allowed to stir at rt for 10 min. A mixture of methyl (S)-2-(tert-butoxycarbonylamino)-3-(4-iodophenyl) propanoate **28** (1.28 mmol, 1 equiv.) and Et<sub>3</sub>N (2.94 mmol, 2.3 equiv.) in dry DMF (4 ml) was added dropwise and the resulting mixture was allowed to stir at rt for 10 min before addition of 2-methyl-2-propanethiol (1.66 mmol, 1.3 equiv.). Reaction was stirred at 75 °C under argon and monitored by TLC (PE/EtOAc, 8:2) until full disappearance of starting material (approx. 45min). The mixture was diluted with EtOAc wash with H<sub>2</sub>O. Aqueous phase was extract with EtOAc (2 x 10 ml). Organic layers were combined, washed with NaHCO<sub>3</sub> (3 x 10ml), brine (3 x 10 ml), dried over MgSO<sub>4</sub> and concentrated. Crude extract was further purified through flash chromatography (hexanes/ EtOAc, 9:1) affording the title compound **29** as thick yellow oil (438 mg; 93 %),  $R_f$  = 0.68 (EP/EtOAc, 2:1),  $[\alpha]^{20}_D$  = + 44 (c = 1, DCM).

Analytical data :

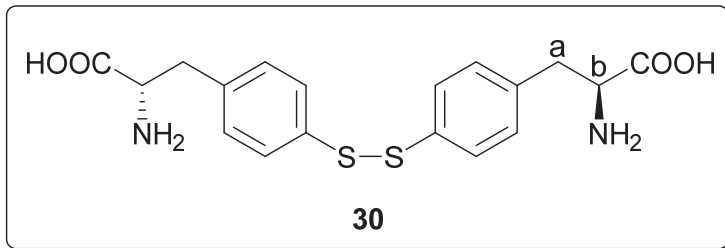
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>),  $\delta$  = 7.45 (d,  $J_{\text{Har}}$  = 8Hz, 2H, H<sub>ar</sub>); 7.09 (d,  $J_{\text{Har}}$  = 8 Hz, 2H, H<sub>ar</sub>); 4.99 (d,  $J$  = 7.5 Hz, 1H, CHNH); 4.62-4.55 (m, 1H, CHNH); 3.69 (s, 3H, OCH<sub>3</sub>); 3.07 (ddd,  $J$  = 39.7, 13.7, 6.0 Hz, 2H, CH<sub>2</sub>CH); 1.41 (s, 9H, 3CH<sub>3</sub>); 1.27 (s, 9H, 3CH<sub>3</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>),  $\delta$  = 172.3 (COOCH<sub>3</sub>); 155.1 (COOC(CH<sub>3</sub>)<sub>3</sub>); 137.7 (CH<sub>ar</sub>); 136.9 (C<sub>ar</sub>); 131.5 (C<sub>ar</sub>); 129.55 (CH<sub>ar</sub>); 80.2 (SC(CH<sub>3</sub>)<sub>3</sub>); 54.5 (NHCH); 52.4 (OCH<sub>3</sub>); 45.9 (OC(CH<sub>3</sub>)<sub>3</sub>); 38.4 (CH<sub>2</sub>CH); 31.1 (SC(CH<sub>3</sub>)<sub>3</sub>); 28.4 (OC(CH<sub>3</sub>)<sub>3</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + Na]<sup>+</sup>     $C_{19}H_{29}NNaO_4S$      $m/z$  theoretical : 390.1710  
 $m/z$  measured : 390.1708

[M + H]<sup>+</sup>     $C_{19}H_{30}NO_4S$      $m/z$  theoretical : 368.1890  
 $m/z$  measured : 368.1890

**Bis-(4-mercapto-L-Phenylalanine) Disulfide (30)** (Rudolf, 2013)



$C_{18}H_{20}N_2O_4S_2$   
Exact Mass: 392,09  
MW: 392,49 g.mol<sup>-1</sup>

Concentrated HCl (37 %, 5 ml) was added to *N*-(tert-butoxycarbonyl)-4-(*S*-tert-butylthio)-L-Phenylalanine methyl ester **29** (91 mg, 0.25 mmol) and the mixture was allowed to stir at reflux until full disappearance of starting material (3h). After cooling to rt, the solution was diluted with H<sub>2</sub>O (20ml) and washed with EtOAc (3 x 20 ml). The aqueous phase was lyophilized to yield the title compound **30** as slightly brown solid (32.1 mg; 65 %), *R*<sub>f</sub> = 0.19 (disulfide), 0.38 (thiol) (ACN/H<sub>2</sub>O, 6:1), [ $\alpha$ ]<sup>20</sup><sub>D</sub> = - 5 (c = 1, MeOH).

Analytical data :

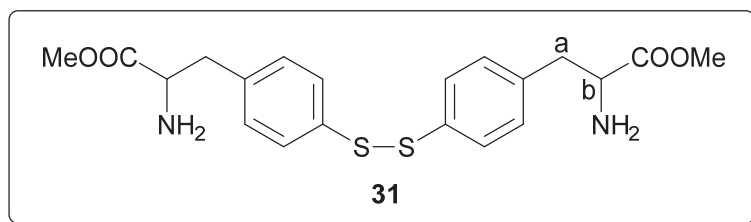
<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>), δ = 8.49 (s, 6H, CHNH<sub>2</sub>); 7.49 (d, *J*<sub>Har</sub> = 8.2 Hz, 4H, H<sub>ar</sub>); 7.32 (d, *J*<sub>Har</sub> = 8.2 Hz, 4H, H<sub>ar</sub>); 4.14 (m, 2H, C<sub>b</sub>HNH<sub>2</sub>); 3.14 (d, *J*<sub>a/b</sub> = 6.3 Hz, 4H, C<sub>a</sub>H<sub>2</sub>C<sub>b</sub>H), assignments were confirmed by qCOSY.

<sup>13</sup>C NMR (125 MHz, DMSO-d<sub>6</sub>), δ = 170.16 (COOCH); 134.8-134.9 (C<sub>ar</sub>); 130.4-127.1 (CH<sub>ar</sub>); 52.7 (CH<sub>2</sub>C<sub>b</sub>H); 34.78 (C<sub>a</sub>H<sub>2</sub>CH), assignments were confirmed by qHSQC.

HRMS (ESI+) : [M + H]<sup>+</sup>      C<sub>18</sub>H<sub>21</sub>N<sub>2</sub>O<sub>4</sub>S<sub>2</sub>      *m/z* theoretical : 393.0936  
*m/z* measured : 393.0937

$[M + 2H]^{2+}$  C<sub>18</sub>H<sub>22</sub>N<sub>2</sub>O<sub>4</sub>S<sub>2</sub>  $m/z$  theoretical : 197.0505  
 $m/z$  measured : 197.0508



Thiotyrosine methyl ester disulfide (**31**)

Exact Mass: 420,12

MW : 420,54 g.mol<sup>-1</sup>

Title compound **31** was obtained according the general procedure 4 starting with Thiotyrosine **30** (150 mg, 0.4 mmol). Purification through flash chromatography (PE/EtOAc, 1:1) yielded tittle **31** compound as uncoloured oil (4 mg, R = 2.4 %),  $R_f$  = 0.18 (thiol), 0.52 (disulfide) (ACN/H<sub>2</sub>O, 6:1).

Analytical data :

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>),  $\delta$  = 7,46 (d,  $J_{\text{Har}}$  = 8 Hz, 2H, H<sub>ar</sub>); 7,20 (d,  $J_{\text{Har}}$  = 8 Hz, 2H, H<sub>ar</sub>); 3,72 (t,  $J_{a/b}$  = 8 Hz, 1H, C<sub>b</sub>H); 3,68 (s, 3H, -OCH<sub>3</sub>); 2,98 (dd,  $J_{a/b}$  = 8 Hz,  $J_{b/b''}$  = 12 Hz, 2H, C<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gCOSY.

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>),  $\delta$  = 176 (COOCH<sub>3</sub>); 138,2 (C<sub>ar</sub>); 136,6 (C<sub>ar</sub>); 131,2 (CH<sub>ar</sub>); 129,3 (CH<sub>ar</sub>); 56,6 (C<sub>b</sub>H); 52,4 (OCH<sub>3</sub>); 41,2 (C<sub>a</sub>H<sub>2</sub>), assignments were confirmed by gHSQC.

HRMS (ESI<sup>+</sup>) : [M + H]<sup>+</sup>    C<sub>20</sub>H<sub>23</sub>O<sub>4</sub>S<sub>2</sub> $m/z$  theoretical : 421.1250 $m/z$  measured : 421.1252

## Thioglycosylation assays of DtGly E159Q with Thiotyrosine

### Reaction conditions

Thioglycosylation reaction were conducted using pNPGlc as donor, unprotected Thiotyrosine (Ttyr) or Thiotyrosine methyl ester (TtyrME) as glycoside acceptor and Dithiothreitol (DTT) as reducing agent. 1 ml solution reaction was prepared with the system donor / acceptor / reducing agent in the ratio 1:2:20 (with 1 mM pNPGlc, 2 mM Ttyr or TtyrME and 20 mM DTT) in phosphate buffer 20 mM pH 8. The mixture was allowed to stir at 37 °C for 1h before addition of the biocatalyst DtGly E159Q (45 nM). Then reaction was incubated at 37 °C under stirring for 20h. Aliquots (100 µl) were withdrawn after  $t = 1\text{h}$ , 2h, 4h and 20h and reaction was stopped by adding quenching solution (100 µl) ACN/HCO<sub>2</sub>H (1:10). Three control solutions were prepared in the same conditions containing : i) 1 mM donor; ii) 2mM acceptor / DTT in the ratio 1:10; iii) 1 mM donor / acceptor / DTT in the ratio 1:2:20. Solutions were incubated at 37 °C under stirring for 20h and aliquots were withdrawn in the same conditions as described previously. All aliquots were filtrated before LC-MS analysis.

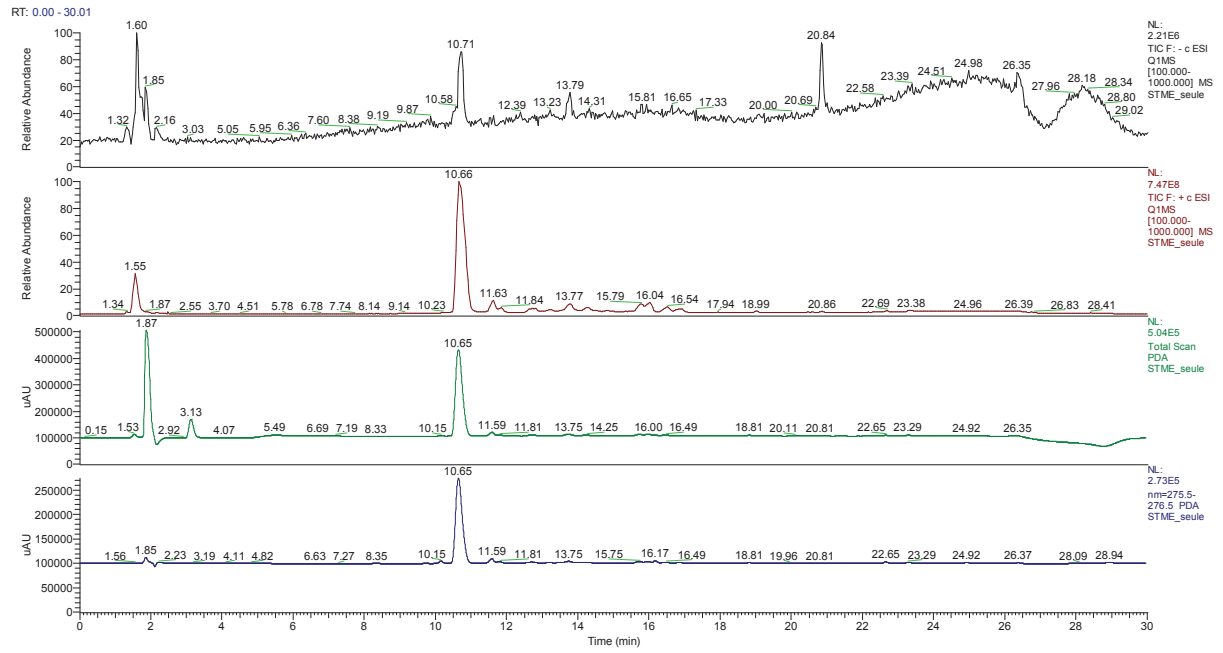
### LC-MS analysis

An Agilent 1100 (Palto Alto, CA, USA) system was coupled to a Micromass Quatro Ultima triple-quadrupole mass spectrometer (Waters, Saint Quentin, France) equipped with a Z-spray dual orthogonal ESI source. Data were performed using Mass Lynx 1.4.1 software (Waters,Saint Quentin,France).

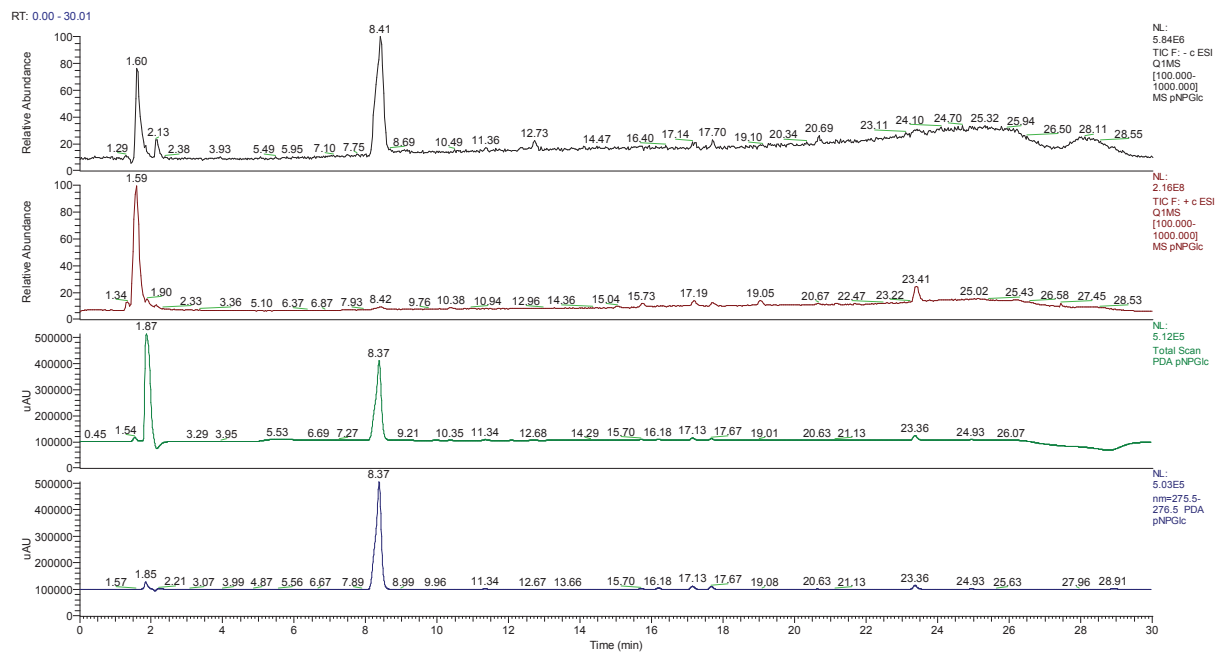
HPLC separation was performed with an Alltima C18 column (150x 4.6 mm; 5µ). Separation was done at a flow rate of 1 mL/min. The mobile phase employed for the separation was composed of water (solvent A) and acetonitrile (solvent B) both acidified with 0.1% formic acid and it was used according to the following elution gradient program: 0-2 min: 5% B; 2-22 min: 5-70% B; 24-26 min: 30-95% B. The injection volume was 20 µL. The diode-array detector recorded the UV spectra from 190 to 300 nm. The 1 mL/min flow from the HPLC-DAD device was then split to obtain a flow rate of 0.3 mL/min directed to the MS system. Measurements were performed in positive and negative ion modes with a 4 kV capillary voltage, a 35 V cone voltage, a 300°C desolvation temperature, a 150°C source temperature, a cone gas (nitrogen) flow rate of 86L/h and a desolvation gas (nitrogen) flow rate of 64 L/h. MS acquisitions were done in full-scan mode from  $m/z$  100 to  $m/z$  1000.

- *The following pictures represent the Mass (TIC) and UV chromatograms obtained for control 1, control 2 and the thioglycosylation reaction performed with DtGly E159Q with TtyrME (Figure 84 –88) :*

## Experimental part

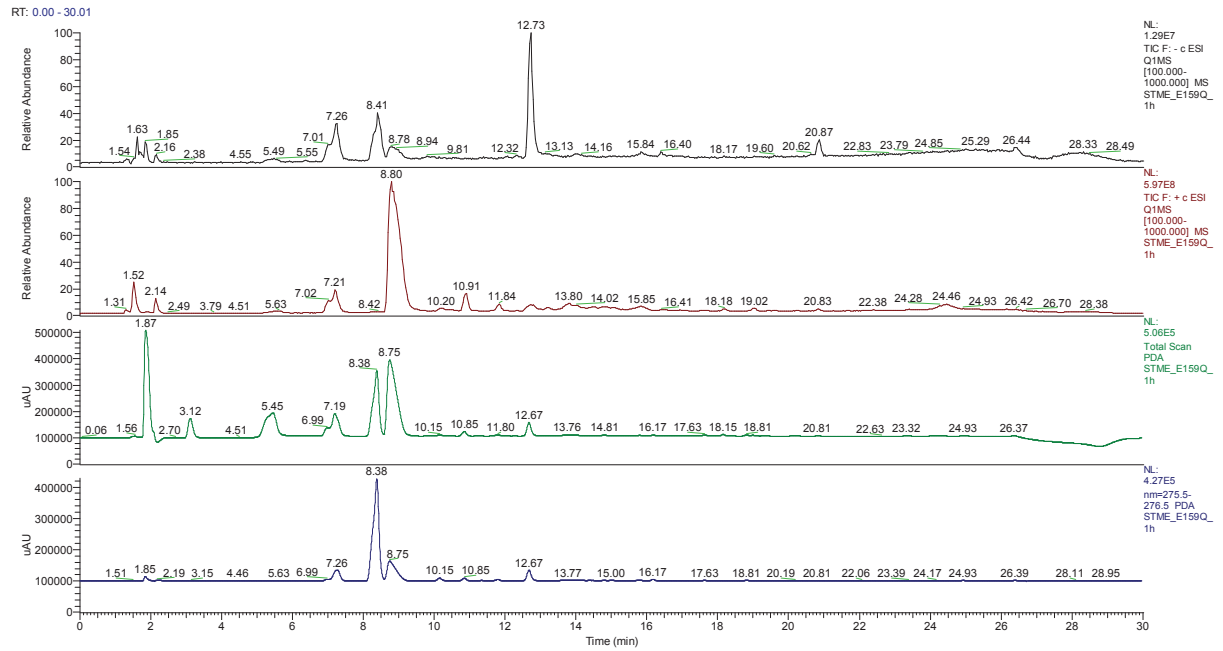


**Figure 84 :** TIC and UV chromatograms of control 1 (TtyrME) after 20h incubation at 37 °C.

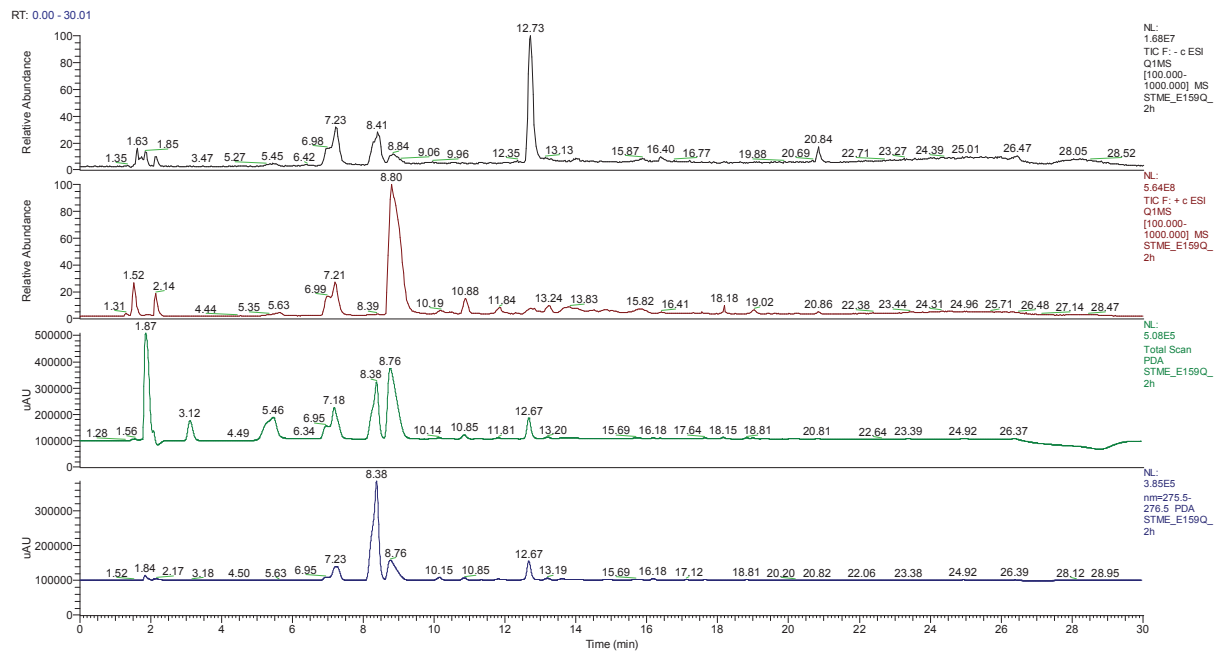


**Figure 85 :** TIC and UV chromatograms of control 2 (pNPGlc) after 20h incubation at 37 °C.

## Experimental part

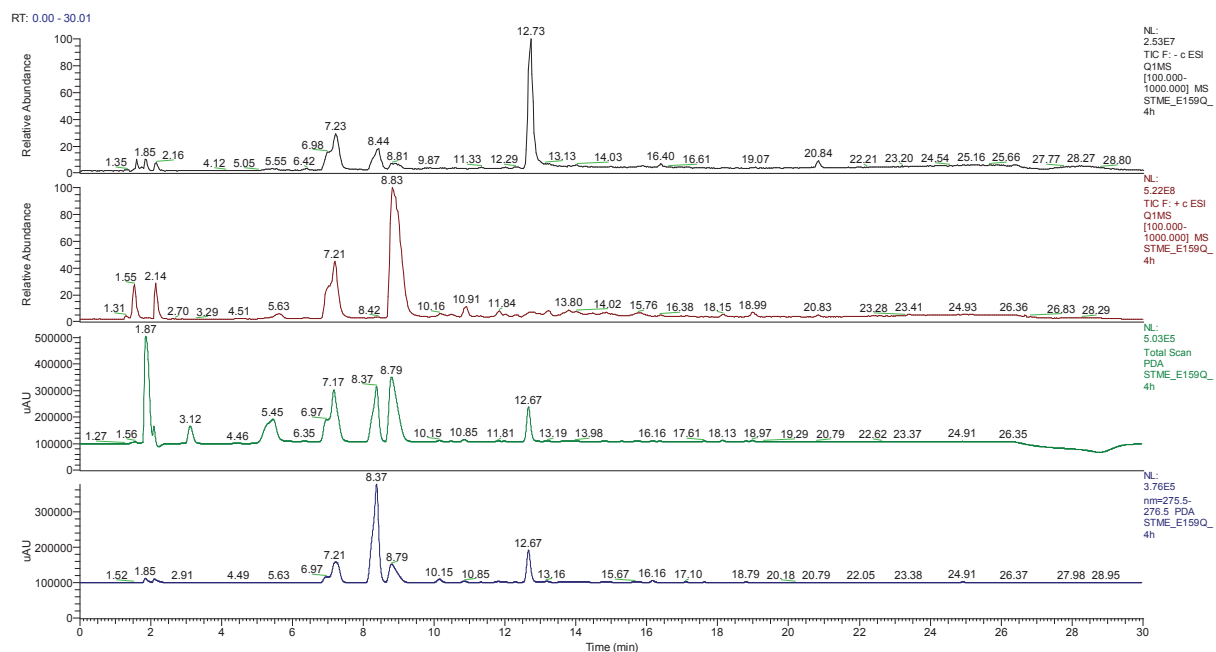


**Figure 86 :** LC-MS analysis of enzymatic thioglucosylation of TtyrME with DtGly E159Q 45 nM after 1h incubation at 37 °C.



**Figure 87 :** LC-MS analysis of enzymatic thioglucosylation of TtyrME with DtGly E159Q 45 nM after 2h incubation at 37 °C.

## Experimental part



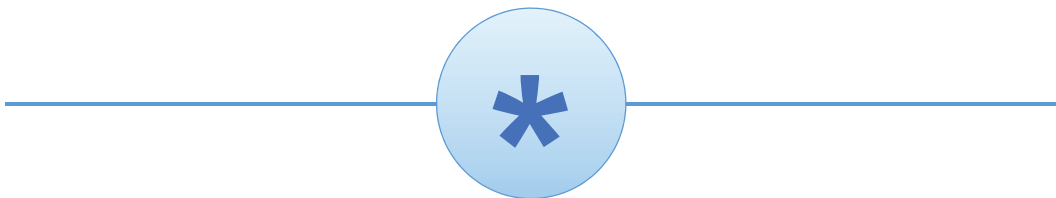
**Figure 88 :** LC-MS analysis of enzymatic thioglucosylation of TtyrME with DtGly E159Q 45 nM after 4h incubation at 37 °C.

## Bibliographie

- Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T.-C., and Zwart, P. H.; *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **(2010)**, 66, 213-221
- Brachet, E., Brion, J. D., Alami, M., and Messaoudi, S.; *Chem. Eur. J.* **(2013)**, 19, 15276-15280
- Bradford, M. M.; *Anal. Biochem.* **(1976)**, 72, 248-254
- Compton, S. J., and Jones, C. G.; *Anal. Biochem.* **(1985)**, 151, 369-374
- El Abdellaoui, S., Destandau, E., Toribio, A., Elfakir, C., Lafosse, M., Renimel, I., Andre, P., Cancellieri, P., and Landemarre, L.; *Anal. Bioanal. Chem.* **(2010)**, 398, 1329-1338
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K.; *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **(2010)**, 66, 486-501
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R.; *Nucleic Acids Res.* **(2010)**, 38, W695-W699
- Hancock, S. M., Corbett, K., Fordham-Skelton, A. P., Gatehouse, J. A., and Davis, B. G.; *Chembiochem* **(2005)**, 6, 866-875
- Kabsch, W.; *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **(2010)**, 66, 125-132
- Kurashima, K., Fujii, M., Ida, Y., and Akita, H.; *J. Mol. Catal. B: Enzym.* **(2003)**, 26, 87-98
- Kurashima, K., Fujii, M., Ida, Y., and Akita, H.; *Chem. Pharm. Bull.* **(2004)**, 52, 270-275
- Li, W., Koike, K., Asada, Y., Yoshikawa, T., and Nikaido, T.; *Carbohydr. Res.* **(2003)**, 338, 729-731
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J.; *J. Appl. Crystallogr.* **(2007)**, 40, 658-674
- Meyer, F. M., Liras, S., Guzman-Perez, A., Perreault, C., Bian, J. W., and James, K.; *Org. Lett.* **(2010)**, 12, 3870-3873
- Novoa, A., Eierhoff, T., Topin, J., Varrot, A., Barluenga, S., Imberty, A., Roemer, W., and Winssinger, N.; *Angew. Chem., Int. Ed.* **(2014)**, 53, 8885-8889
- Ooi, Y., Hashimoto, T., Mitsuo, N., and Satoh, T.; *Chem. Pharm. Bull.* **(1985)**, 33, 1808-1814
- Rudolf, J. D., and Poulter, C. D.; *ACS Chem. Biol.* **(2013)**, 8, 2707-2714
- Samoshin, A. V., Dotsenko, I. A., Samoshina, N. M., Franz, A. H., and Samoshin, V. V.; *Int. J. Carbohydr. Chem.* **(2014)**, 2014, 1-8
- Sasaki, S., Kitamura, S., Negoro, N., Suzuki, M., Tsujihata, Y., Suzuki, N., Santou, T., Kanzaki, N., Harada, M., Tanaka, Y., Kobayashi, M., Tada, N., Funami, M., Tanaka, T., Yamamoto, Y., Fukatsu, K., Yasuma, T., and Momose, Y.; *J. Med. Chem.* **(2011)**, 54, 1365-1378
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J., and Higgins, D. G.; *Mol. Syst. Biol.* **(2011)**, 7, 539
- Soria-Castro, S. M., and Penénory, A. B.; *Beilstein J. Org. Chem.* **(2013)**, 9, 467-475
- Stevenson, D. E., Stanley, R. A., and Furneaux, R. H.; *Biotechnol. Bioeng.* **(1993)**, 42, 657-666

- Tailford, L. E., Money, V. A., Smith, N. L., Dumon, C., Davies, G. J., and Gilbert, H. J.; *J. Biol. Chem.* **(2007)**, 282, 11291-11299
- Tyl, C., Felsing, S., and Brecker, L.; *J. Mol. Catal. B: Enzym.* **(2004)**, 28, 55-63

## ANNEXES







## LISTE DES ANNEXES

**Annexe 01** : Alignement des séquences peptidiques de B5AYN4 et de la mannosidase cristallisée BtMan2A.

**Annexe 02** : Alignement des séquences ADN du clone pET28-a(+) *ditch\_1692* clone (sens F et anti-sens R) et du gène *ditch\_1692*.

**Annexe 03** : Alignement des séquences ADN du clone pET28-a(+) *ditch\_0289* clone (sens F et anti-sens R) et du gène *ditch\_0289*.

**Annexe 04** : Alignement des séquences ADN du clone pET28-a(+) *ditch\_0359* clone (sens F et anti-sens R) et du gène *ditch\_0359*.

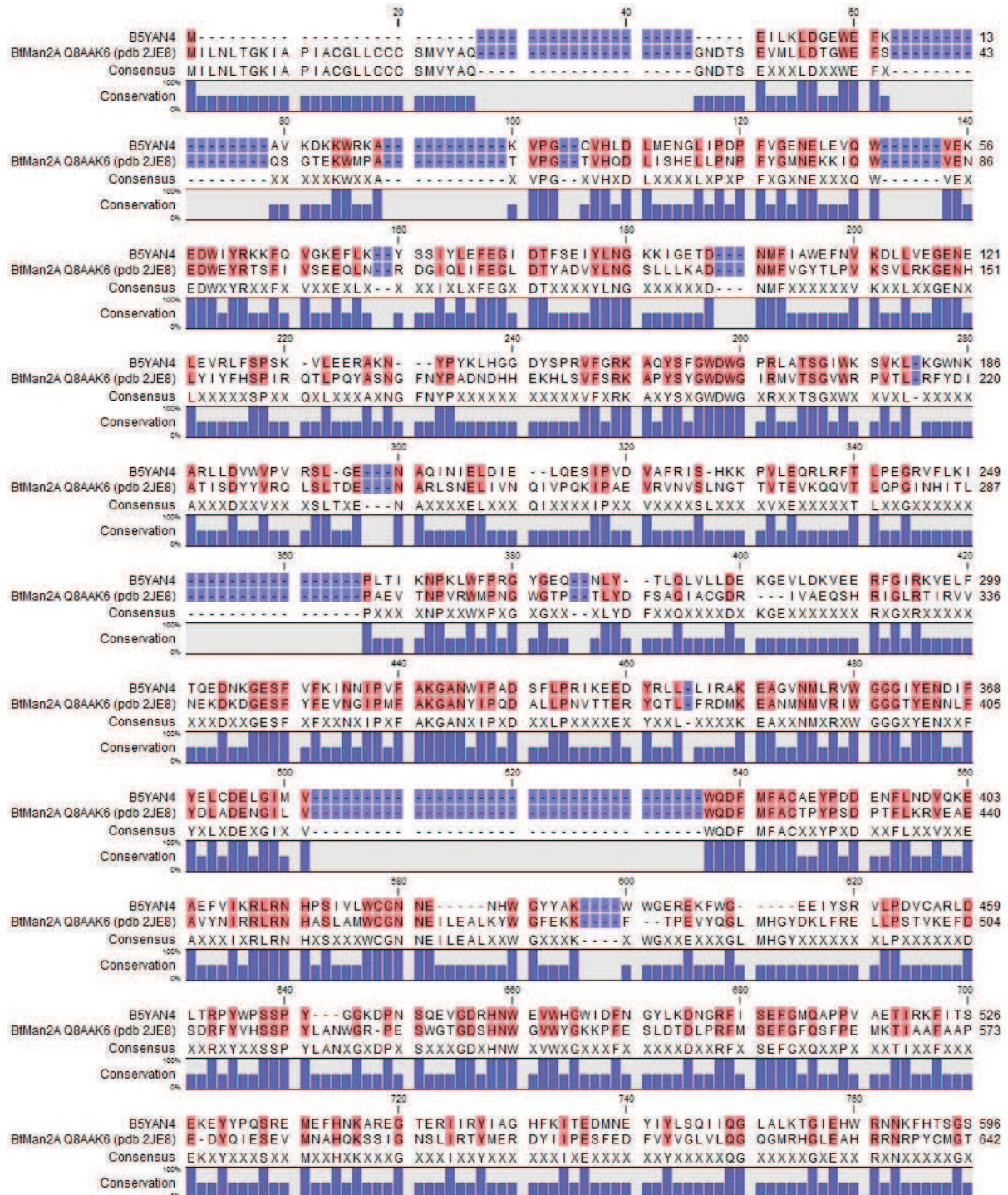
**Annexe 05** : Alignement des séquences protéique de DtGly (B5YCI2) et de glycosidases de structures connues de GH1.

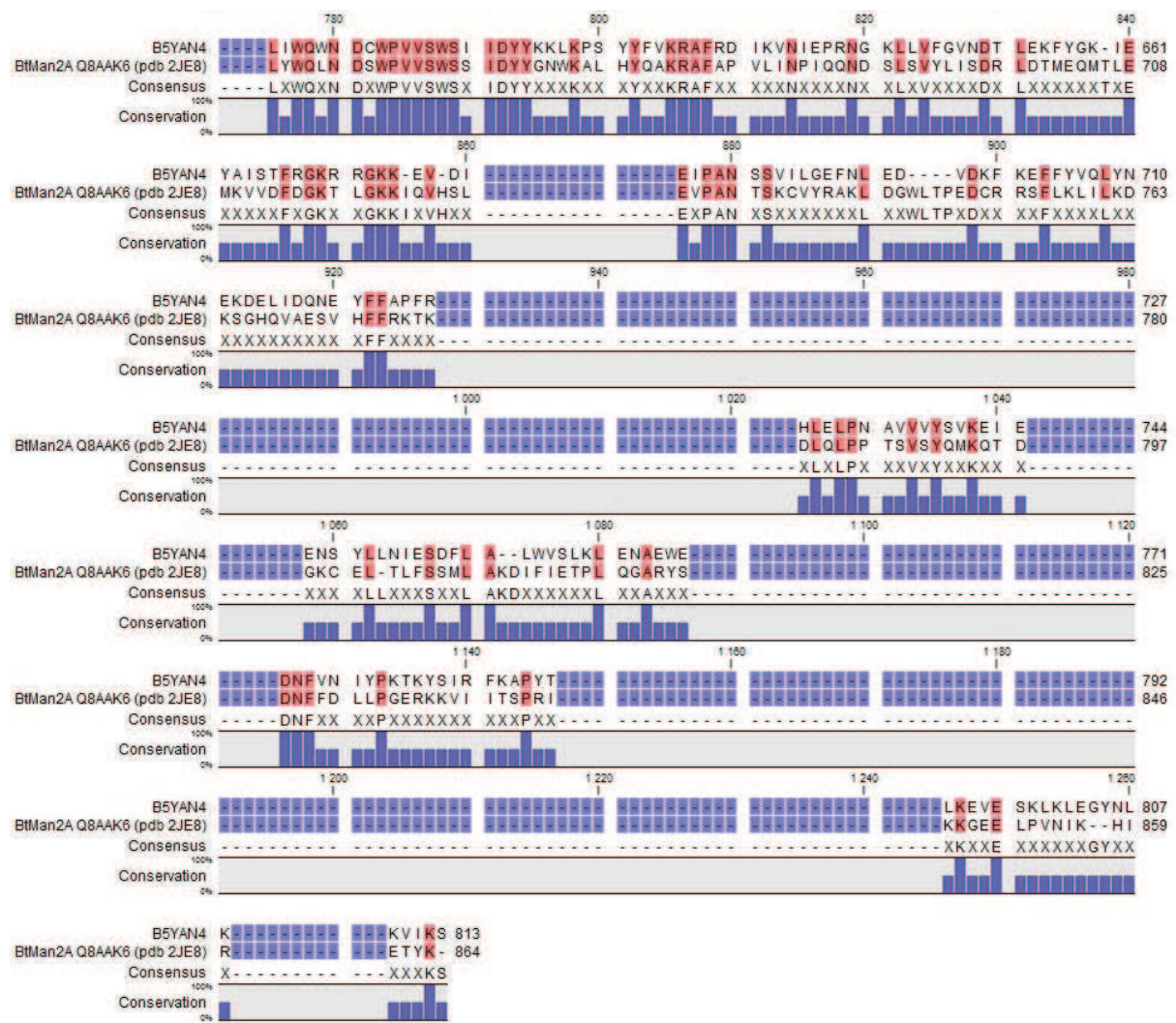
**Annexe 06** : Alignement des séquences ADN de DtGly *wt* et des mutants DtGly E159Q et E159A.

**Annexe 07** : Alignement des séquences peptidiques de DtMan (B5YAN4) et de glycosidases de structures connues de GH2.

**Annexe 08** : Alignement des séquences ADN de DtMan *wt* et des mutants DtMan E425X.

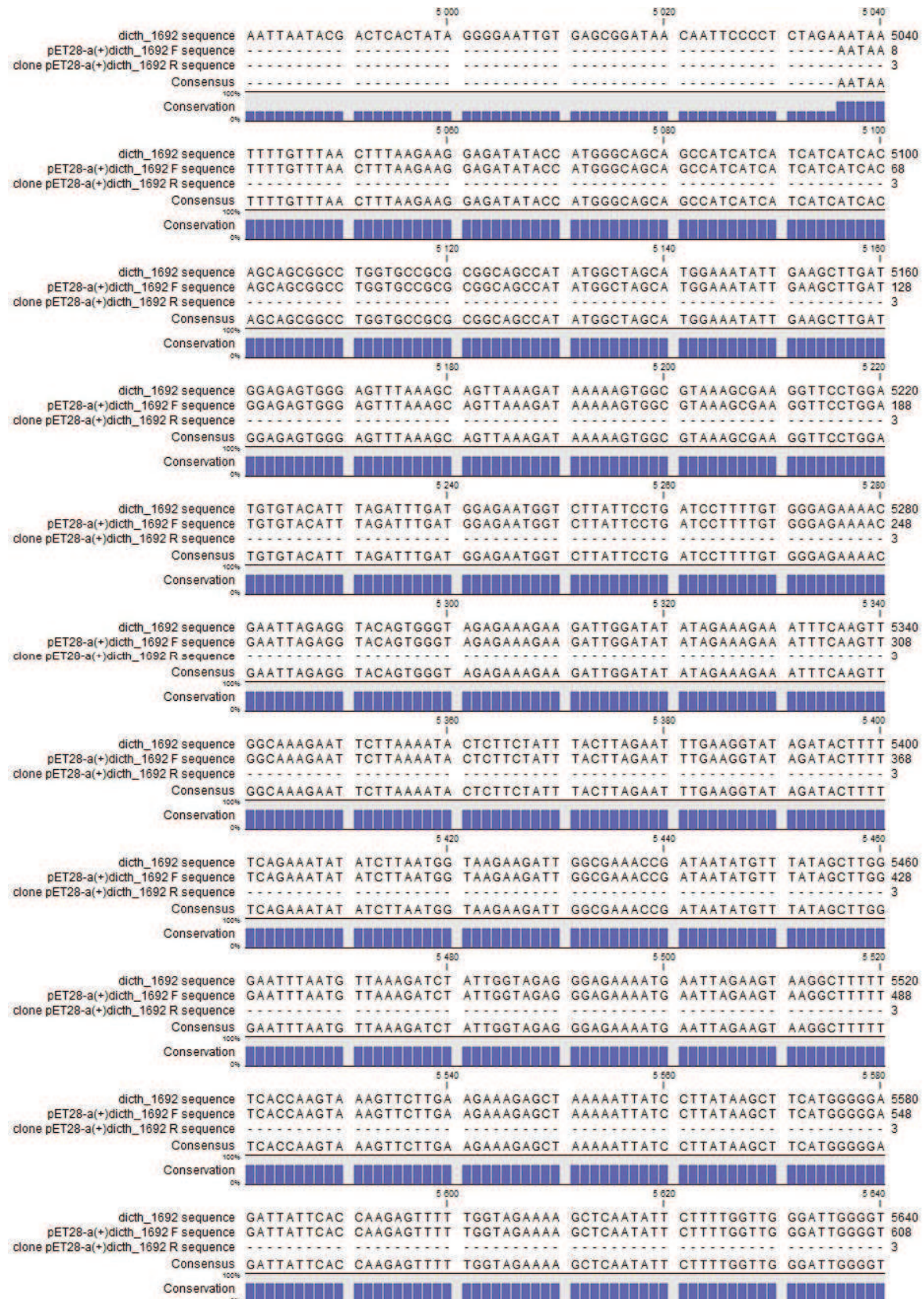
**Annexe 01 : Alignement des séquences peptidiques de B5AYN4 et de la mannosidase cristallisée BtMan2A.** Les protéines sont identifiées selon leur code Uniprot ainsi que leur code PDB donné entre parenthèses. L'homologie des résidu est indiqué suivant le gradient bleu –rouge (bleu = résidus non homologues, rouge = résidus homologues).





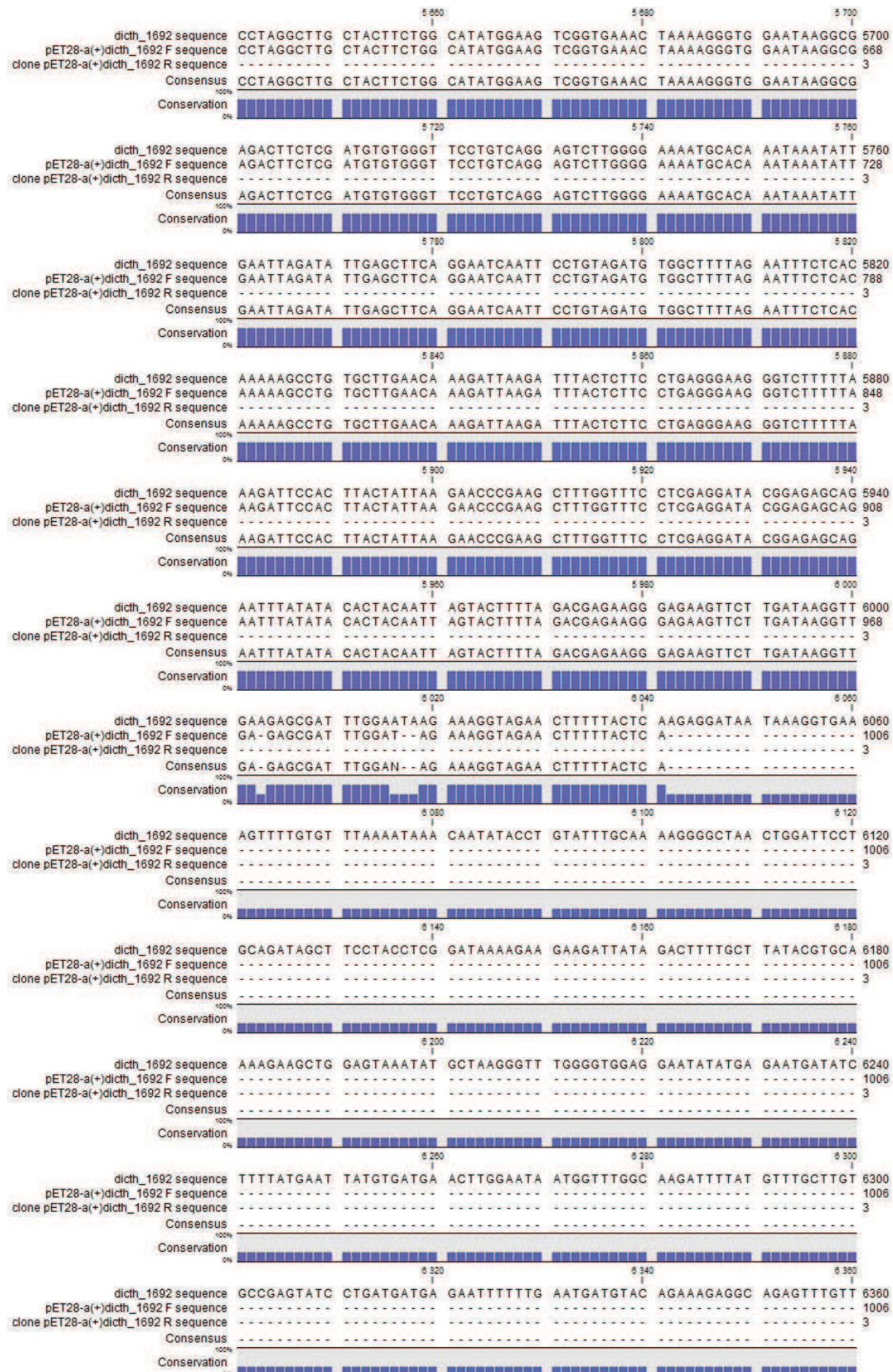


## Annexe 02 : Alignement des séquences ADN du clone pET28-a(+) *ditch\_1692* clone (sens F et anti-sens R) et du gène *ditch\_1692*.



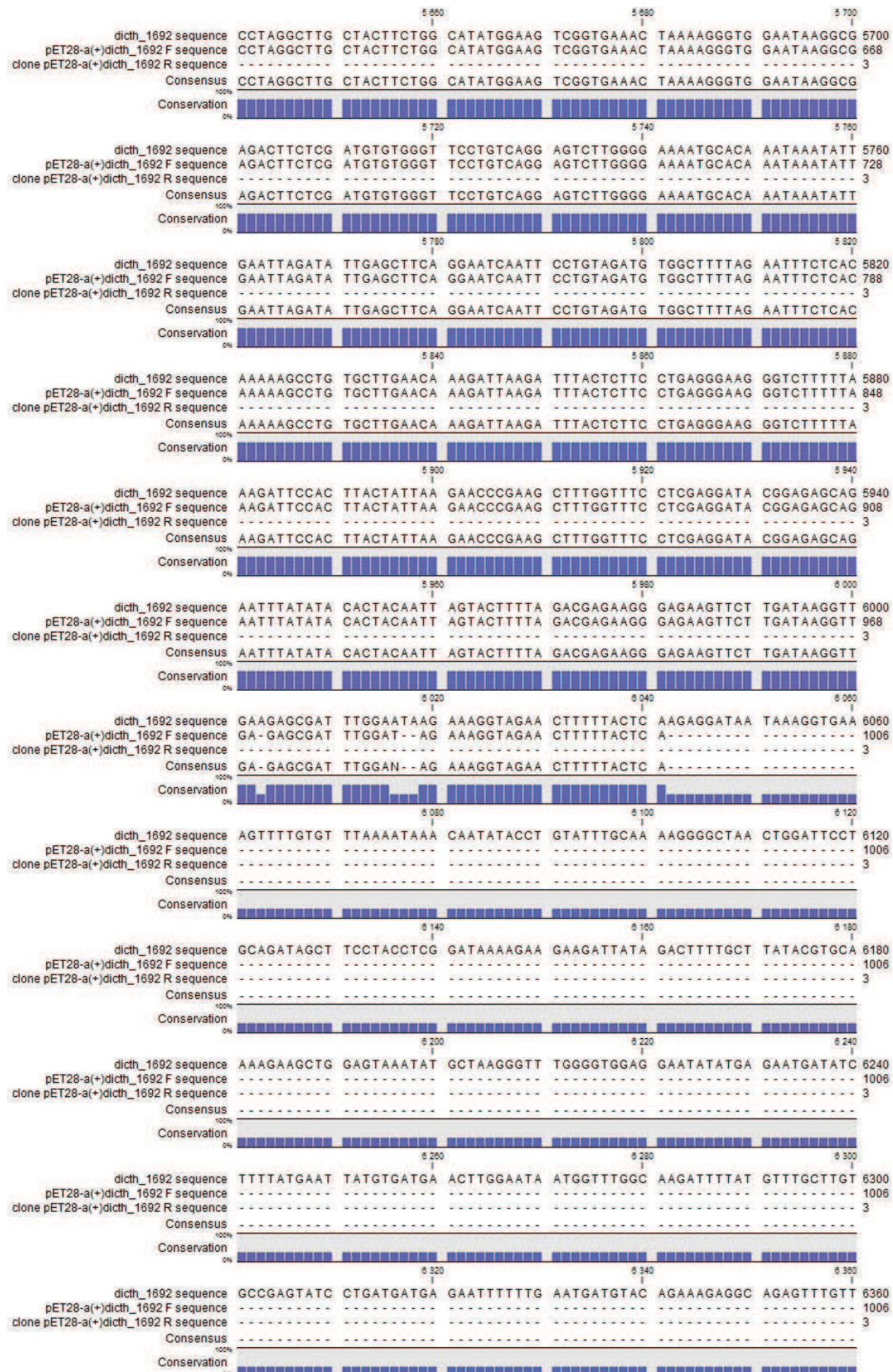


## Annexes



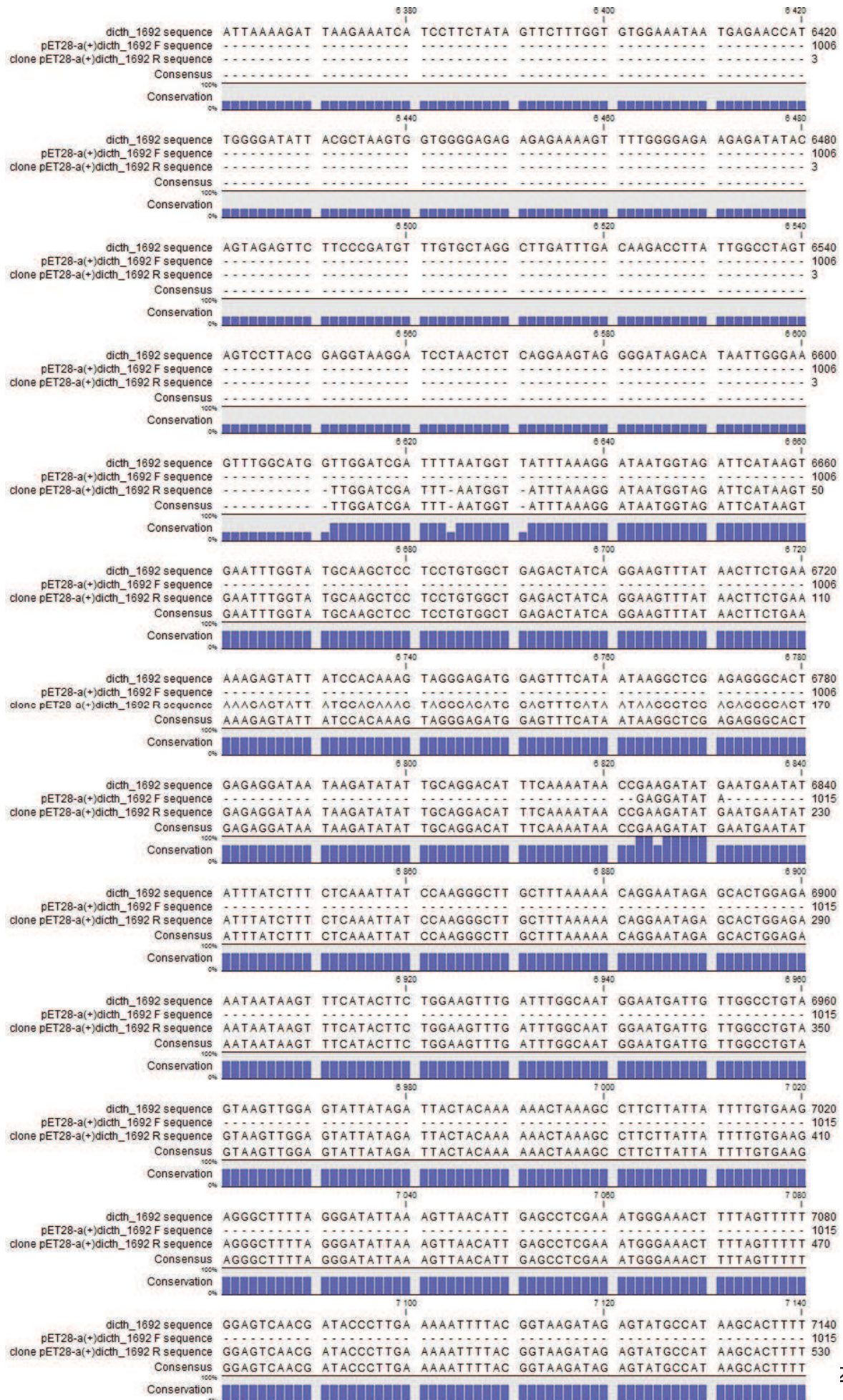


## Annexes



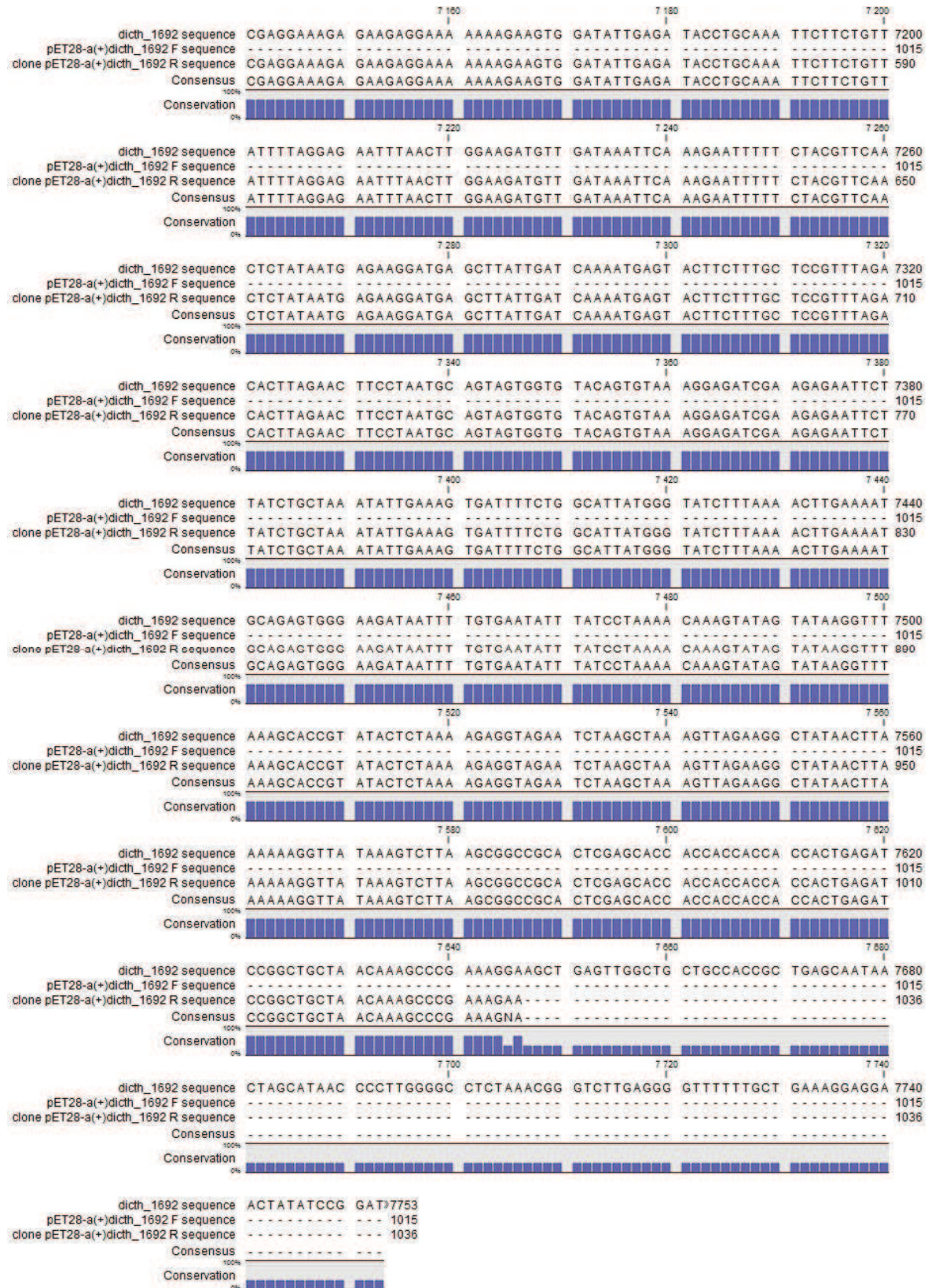


# Annexes



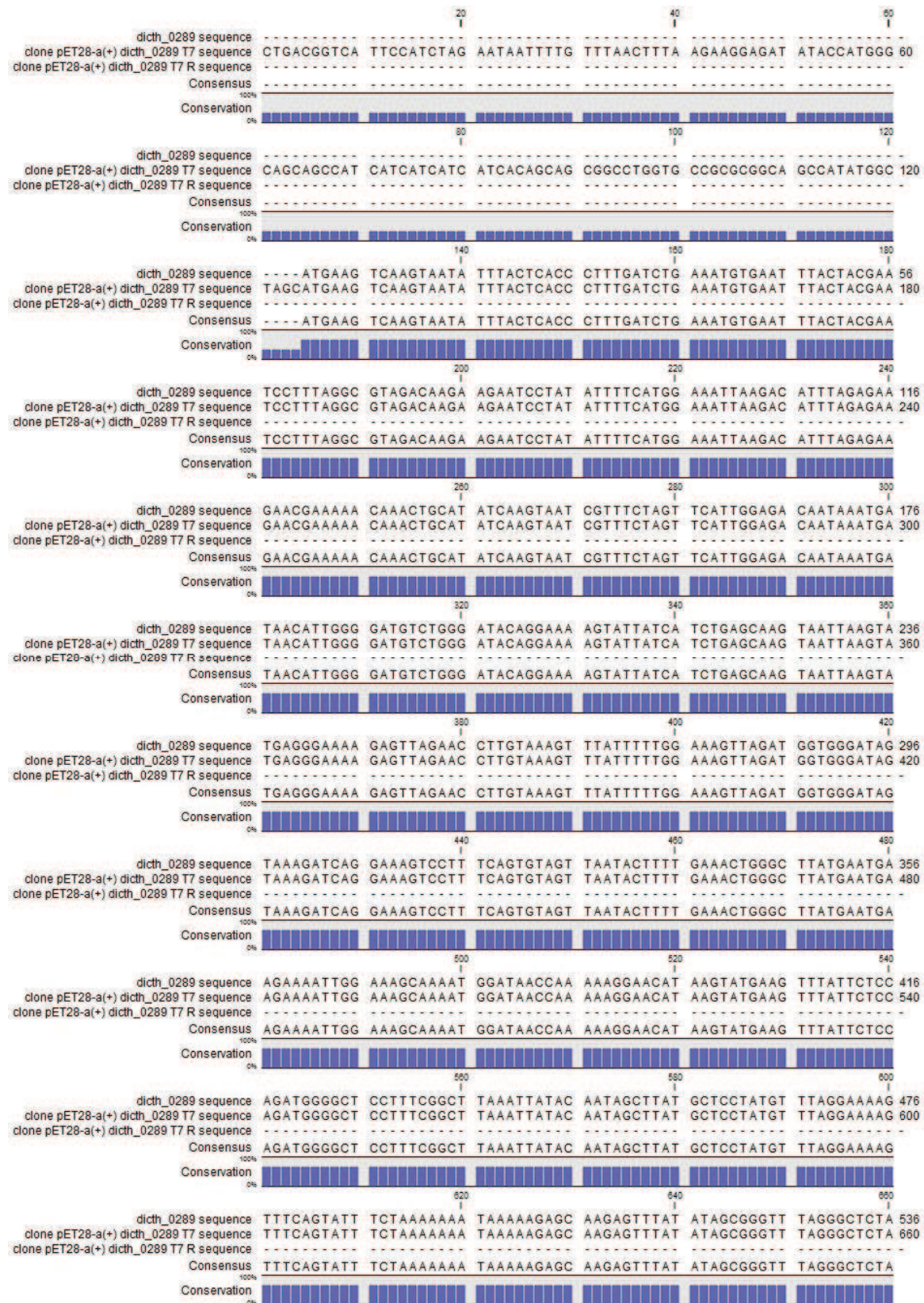


# Annexes



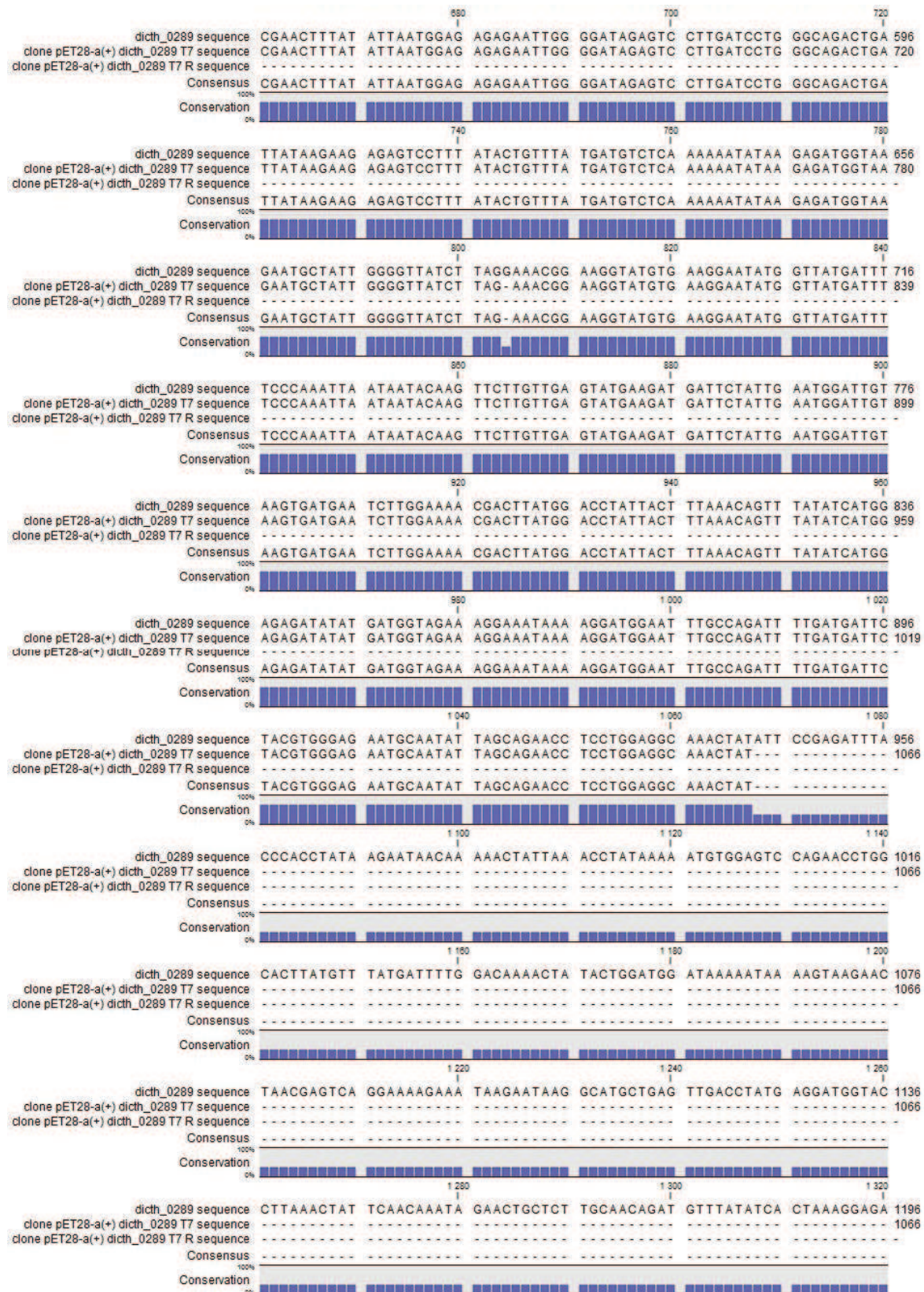


### Annexe 03 : Alignement des séquences ADN du clone pET28-a(+) *ditch\_0289* clone (sens F et anti-sens R) et du gène *ditch\_0289*.



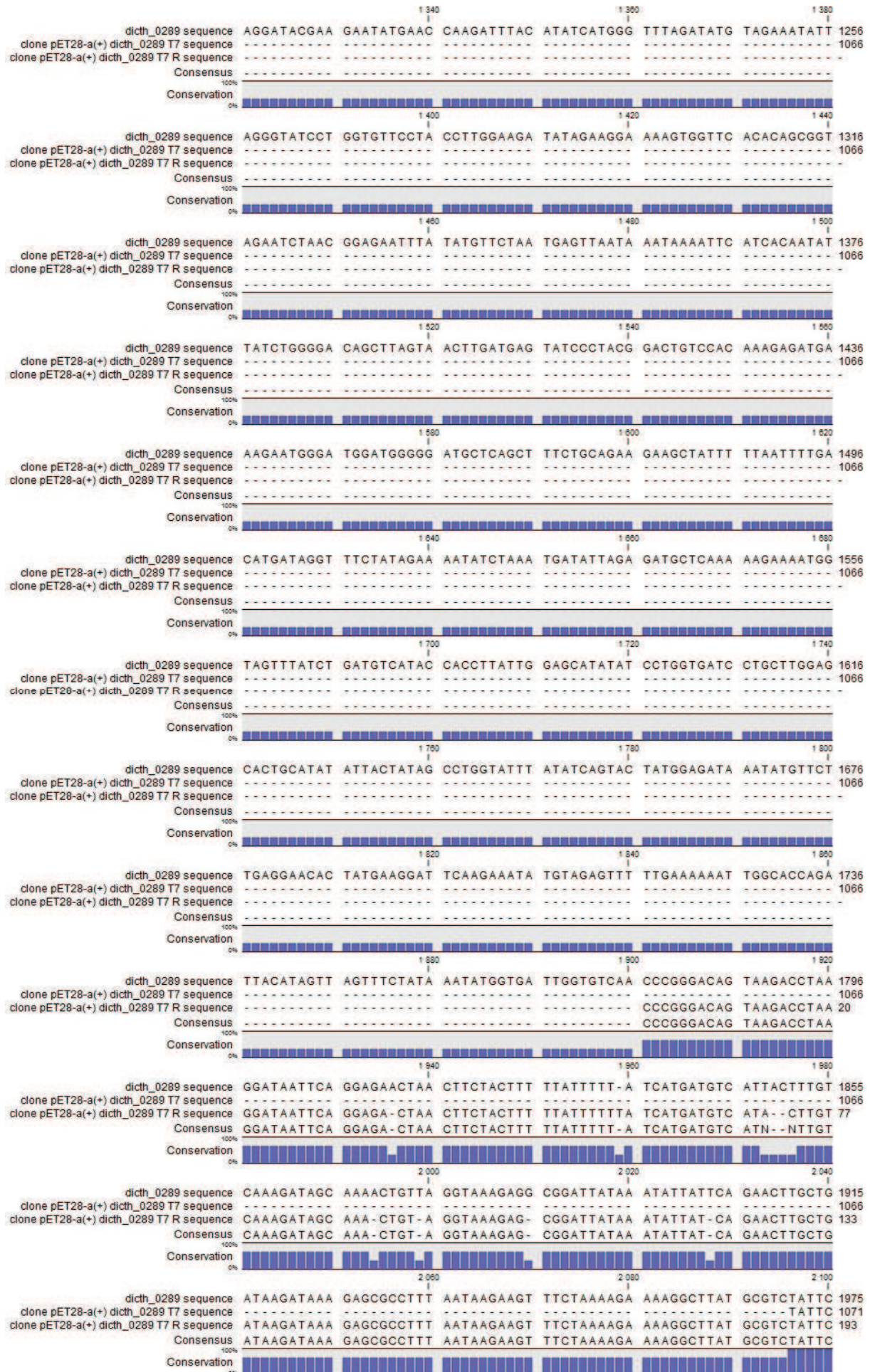


# Annexes





# Annexes



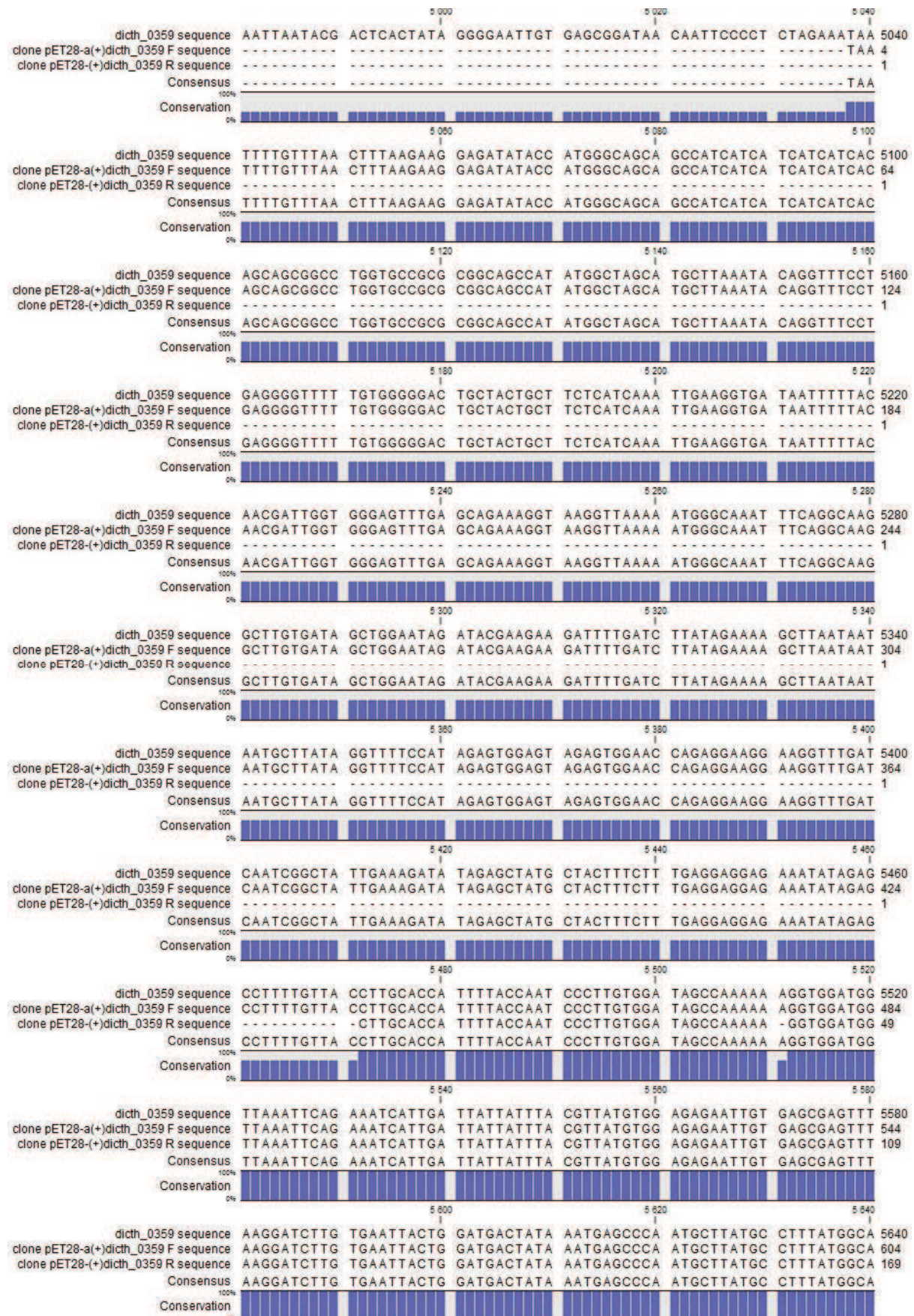


# Annexes

dich_0289 sequence	CAACAGAGTT	GAGTGAAGAA	AATGTAAAAG	CACTCTTAGA	AAAGTATCCT	GAAGATATAA	2035
clone pET28-a(+) dich_0289 T7 sequence	CGA-----	GAGTGAAGAA	AATGTAAAAG	CACTCTTAGA	AAAGTATCCT	GAAGATATAA	1074
clone pET28-a(+) dich_0289 T7 R sequence	CAACAGAGTT	GAGTGAAGAA	AATGTAAAAG	CACTCTTAGA	AAAGTATCCT	GAAGATATAA	253
Consensus	CAACAGAGTT	GAGTGAAGAA	AATGTAAAAG	CACTCTTAGA	AAAGTATCCT	GAAGATATAA	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	AAGACTTTCT	GAGGCAACAA	TTTACTATTT	TATCGTCTTT	AGGGATGTTT	ACAAGTCAGA	2095
clone pET28-a(+) dich_0289 T7 sequence	AAGACTTTCT	GAGGCAACAA	TTTACTATTT	TATCGTCTTT	AGGGATGTTT	ACAAGTCAGA	1074
clone pET28-a(+) dich_0289 T7 R sequence	AAGACTTTCT	GAGGCAACAA	TTTACTATTT	TATCGTCTTT	AGGGATGTTT	ACAAGTCAGA	313
Consensus	AAGACTTTCT	GAGGCAACAA	TTTACTATTT	TATCGTCTTT	AGGGATGTTT	ACAAGTCAGA	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	CTTTGAATAC	TCTGCCTCTA	TATTTAAATC	TTGTTCTCTGA	AGATAAGGTG	CAAGATGTTT	2155
clone pET28-a(+) dich_0289 T7 sequence	CTTTGAATAC	TCTGCCTCTA	TATTTAAATC	TTGTTCTCTGA	AGATAAGGTG	CAAGATGTTT	1074
clone pET28-a(+) dich_0289 T7 R sequence	CTTTGAATAC	TCTGCCTCTA	TATTTAAATC	TTGTTCTCTGA	AGATAAGGTG	CAAGATGTTT	373
Consensus	CTTTGAATAC	TCTGCCTCTA	TATTTAAATC	TTGTTCTCTGA	AGATAAGGTG	CAAGATGTTT	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	TAAAGACATT	ATTAGAGGAT	ATTATTATTA	GACATGATTA	TCATCTTGAT	ACAGGAATTG	2215
clone pET28-a(+) dich_0289 T7 sequence	TAAAGACATT	ATTAGAGGAT	ATTATTATTA	GACATGATTA	TCATCTTGAT	ACAGGAATTG	1074
clone pET28-a(+) dich_0289 T7 R sequence	TAAAGACATT	ATTAGAGGAT	ATTATTATTA	GACATGATTA	TCATCTTGAT	ACAGGAATTG	433
Consensus	TAAAGACATT	ATTAGAGGAT	ATTATTATTA	GACATGATTA	TCATCTTGAT	ACAGGAATTG	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	TTGCTACAAG	ATATATATTT	GATGTTCTAA	CTAGTTATGG	GTATGACGAG	GTGGCTTATA	2275
clone pET28-a(+) dich_0289 T7 sequence	TTGCTACAAG	ATATATATTT	GATGTTCTAA	CTAGTTATGG	GTATGACGAG	GTGGCTTATA	1074
clone pET28-a(+) dich_0289 T7 R sequence	TTGCTACAAG	ATATATATTT	GATGTTCTAA	CTAGTTATGG	GTATGACGAG	GTGGCTTATA	493
Consensus	TTGCTACAAG	ATATATATTT	GATGTTCTAA	CTAGTTATGG	GTATGACGAG	GTGGCTTATA	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	AAATAGTTAA	TCAAAAAACA	TATCCGAGTT	TTGGATATAT	GATCGAAGAG	GGTGCTACAA	2335
clone pET28-a(+) dich_0289 T7 sequence	AAATAGTTAA	TCAAAAAACA	TATCCGAGTT	TTGGATATAT	GATCGAAGAG	GGTGCTACAA	1074
clone pET28-a(+) dich_0289 T7 R sequence	AAATAGTTAA	TCAAAAAACA	TATCCGAGTT	TTGGATATAT	GATCGAAGAG	GGTGCTACAA	553
Consensus	AAATAGTTAA	TCAAAAAACA	TATCCGAGTT	TTGGATATAT	GATCGAAGAG	GGTGCTACAA	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	CTTTATGGGA	AAGATGGGAG	AAATTAACAT	CTACAGGTAT	GAATTCTCAT	AATCATATTA	2395
clone pET28-a(+) dich_0289 T7 sequence	CTTTATGGGA	AAGATGGGAG	AAATTAACAT	CTACAGGTAT	GAATTCTCAT	AATCATATTA	1074
clone pET28-a(+) dich_0289 T7 R sequence	CTTTATGGGA	AAGATGGGAG	AAATTAACAT	CTACAGGTAT	GAATTCTCAT	AATCATATTA	613
Consensus	CTTTATGGGA	AAGATGGGAG	AAATTAACAT	CTACAGGTAT	GAATTCTCAT	AATCATATTA	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	TGTTTGGTAG	TGTAGATGCT	TGGTTCTATA	GAGTTATTGC	TGGAGTAAGA	GTAGGAGAGC	2455
clone pET28-a(+) dich_0289 T7 sequence	TGTTTGGTAG	TGTAGATGCT	TGGTTCTATA	GAGTTATTGC	TGGAGTAAGA	GTAGGAGAGC	1074
clone pET28-a(+) dich_0289 T7 R sequence	TGTTTGGTAG	TGTAGATGCT	TGGTTCTATA	GAGTTATTGC	TGGAGTAAGA	GTAGGAGAGC	073
Consensus	TGTTTGGTAG	TGTAGATGCT	TGGTTCTATA	GAGTTATTGC	TGGAGTAAGA	GTAGGAGAGC	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	CAGGGTGGAA	TAAAAATAA	TTTGAGCCTC	ATCCAGTTGG	AGATCTAAAG	TACGCCAAGG	2515
clone pET28-a(+) dich_0289 T7 sequence	CAGGGTGGAA	TAAAAATAA	TTTGAGCCTC	ATCCAGTTGG	AGATCTAAAG	TACGCCAAGG	1074
clone pET28-a(+) dich_0289 T7 R sequence	CAGGGTGGAA	TAAAAATAA	TTTGAGCCTC	ATCCAGTTGG	AGATCTAAAG	TACGCCAAGG	733
Consensus	CAGGGTGGAA	TAAAAATAA	TTTGAGCCTC	ATCCAGTTGG	AGATCTAAAG	TACGCCAAGG	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	CGAGATTAAA	TACCATTAAA	GGGGAAGTGG	AGATAAATTG	GCAGAAAAC	GAAAAATATAT	2575
clone pET28-a(+) dich_0289 T7 sequence	CGAGATTAAA	TACCATTAAA	GGGGAAGTGG	AGATAAATTG	GCAGAAAAC	GAAAAATATAT	1074
clone pET28-a(+) dich_0289 T7 R sequence	CGAGATTAAA	TACCATTAAA	GGGGAAGTGG	AGATAAATTG	GCAGAAAAC	GAAAAATATAT	793
Consensus	CGAGATTAAA	TACCATTAAA	GGGGAAGTGG	AGATAAATTG	GCAGAAAAC	GAAAAATATAT	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	TTAGTATGAG	GATATCTGTA	CCAGTGAAC	CTGAGGGAGA	AGTGCATGTA	CCTAAGTTGT	2635
clone pET28-a(+) dich_0289 T7 sequence	TTAGTATGAG	GATATCTGTA	CCAGTGAAC	CTGAGGGAGA	AGTGCATGTA	CCTAAGTTGT	1074
clone pET28-a(+) dich_0289 T7 R sequence	TTAGTATGAG	GATATCTGTA	CCAGTGAAC	CTGAGGGAGA	AGTGCATGTA	CCTAAGTTGT	853
Consensus	TTAGTATGAG	GATATCTGTA	CCAGTGAAC	CTGAGGGAGA	AGTGCATGTA	CCTAAGTTGT	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	TTGAGAGATT	TGTAGTGAAA	GAAGGAGATA	ATATAATTTA	TGAAAAAAG	GGTGATCTTG	2695
clone pET28-a(+) dich_0289 T7 sequence	TTGAGAGATT	TGTAGTGAAA	GAAGGAGATA	ATATAATTTA	TGAAAAAAG	GGTGATCTTG	1074
clone pET28-a(+) dich_0289 T7 R sequence	TTGAGAGATT	TGTAGTGAAA	GAAGGAGATA	ATATAATTTA	TGAAAAAAG	GGTGATCTTG	913
Consensus	TTGAGAGATT	TGTAGTGAAA	GAAGGAGATA	ATATAATTTA	TGAAAAAAG	GGTGATCTTG	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	AAGAAAAATGA	GAAGTATATA	GTCATAAGAG	TTGGATCCGG	ATCATACAAC	TTTTATATGG	2755
clone pET28-a(+) dich_0289 T7 sequence	AAGAAAAATGA	GAAGTATATA	GTCATAAGAG	TTGGATCCGG	ATCATACAAC	TTTTATATGG	1074
clone pET28-a(+) dich_0289 T7 R sequence	AAGAAAAATGA	GAAGTATATA	GTCATAAGAG	TTGGATCCGG	ATCATACAAC	TTTTATATGG	973
Consensus	AAGAAAAATGA	GAAGTATATA	GTCATAAGAG	TTGGATCCGG	ATCATACAAC	TTTTATATGG	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	AAAAGATATA	A-----	-----	-----	-----	-----	2766
clone pET28-a(+) dich_0289 T7 sequence	AAAAGATATA	AGTCGACAAG	CTTGCGGCCG	CACTCGAGCA	CCAGCACCAC	CACCACTGAG	1074
clone pET28-a(+) dich_0289 T7 R sequence	AAAAGATATA	A-----	-----	-----	-----	-----	1033
Consensus	AAAAGATATA	A-----	-----	-----	-----	-----	
Conservation	100%	100%	100%	100%	100%	100%	
dich_0289 sequence	-----	-----	-----	-----	-----	-----	2766
clone pET28-a(+) dich_0289 T7 sequence	ATCCGGCTGC	TAACAAAAGCC	CGAAAAGAGCG	AATTGCGAC	-----	-----	1074
clone pET28-a(+) dich_0289 T7 R sequence	ATCCGGCTGC	TAACAAAAGCC	CGAAAAGAGCG	AATTGCGAC	-----	-----	1071
Consensus	ATCCGGCTGC	TAACAAAAGCC	CGAAAAGAGCG	AATTGCGAC	-----	-----	
Conservation	100%	100%	100%	100%	100%	100%	

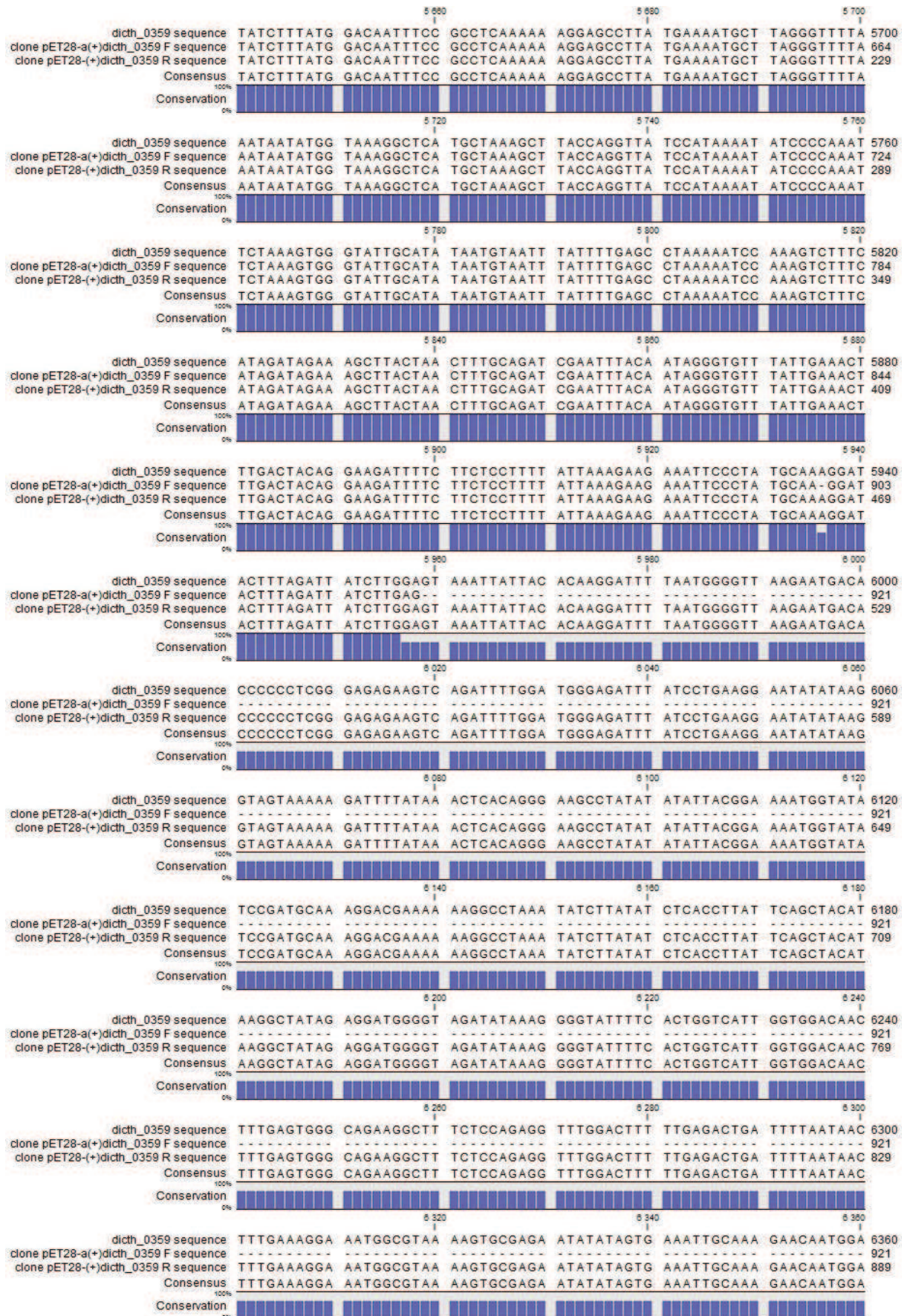


# **Annexe 04 : Alignement des séquences ADN du clone pET28-a(+) *ditch\_0359* clone (sens F et anti-sens R) et du gène *ditch\_0359*.**

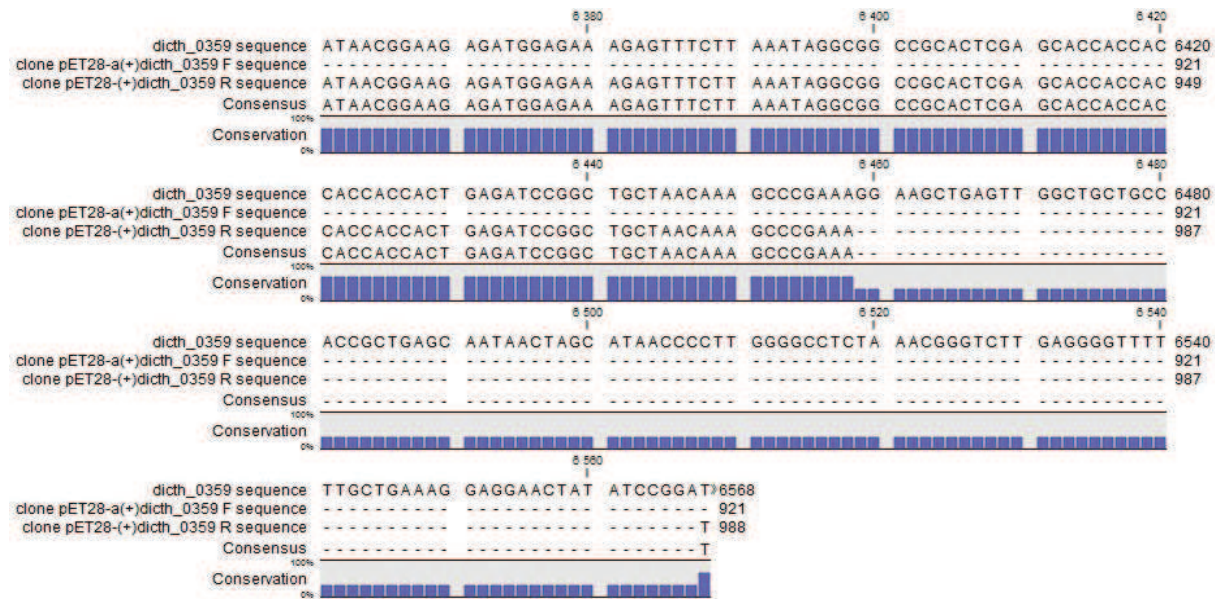




# Annexes

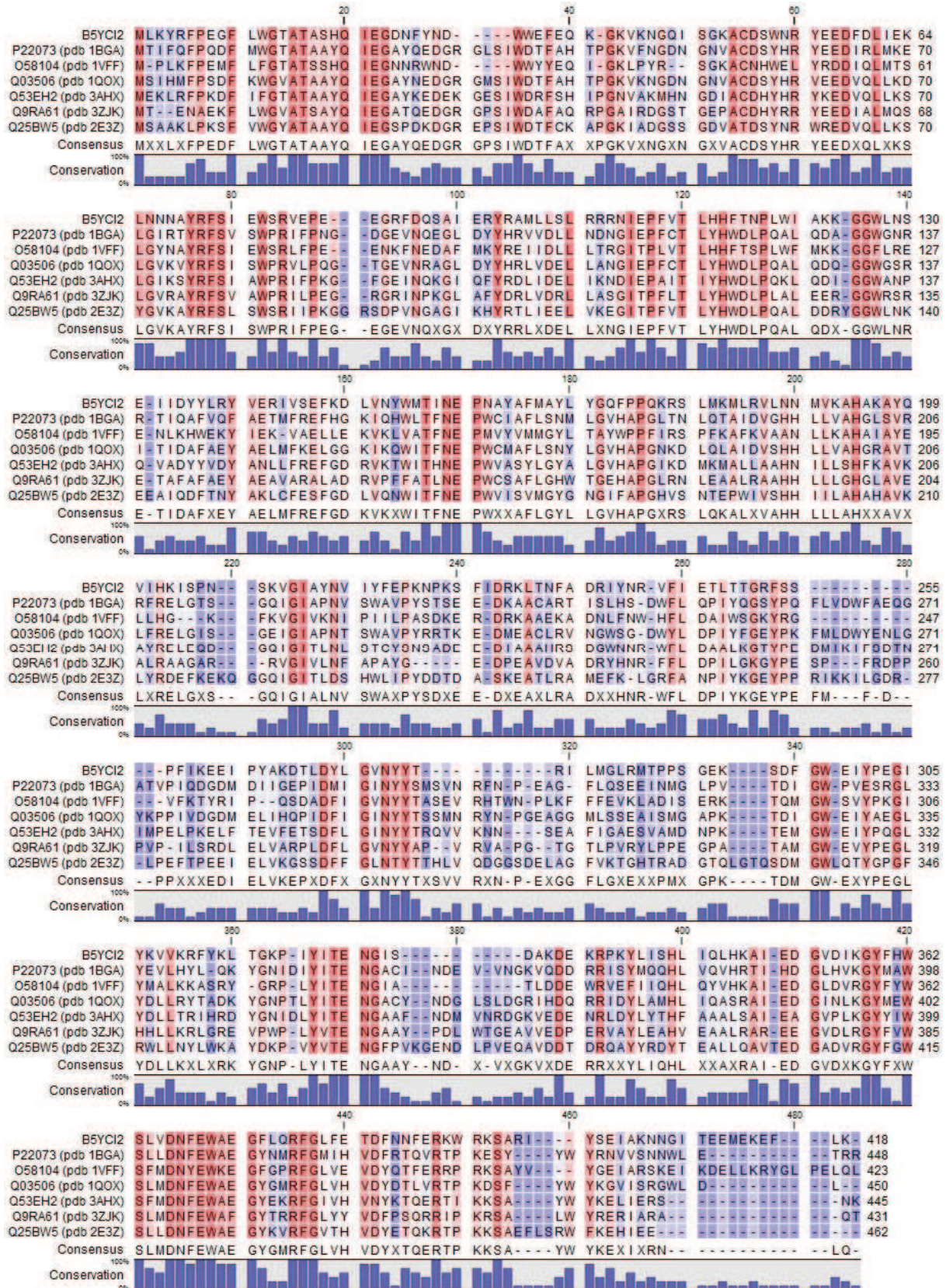


## Annexes



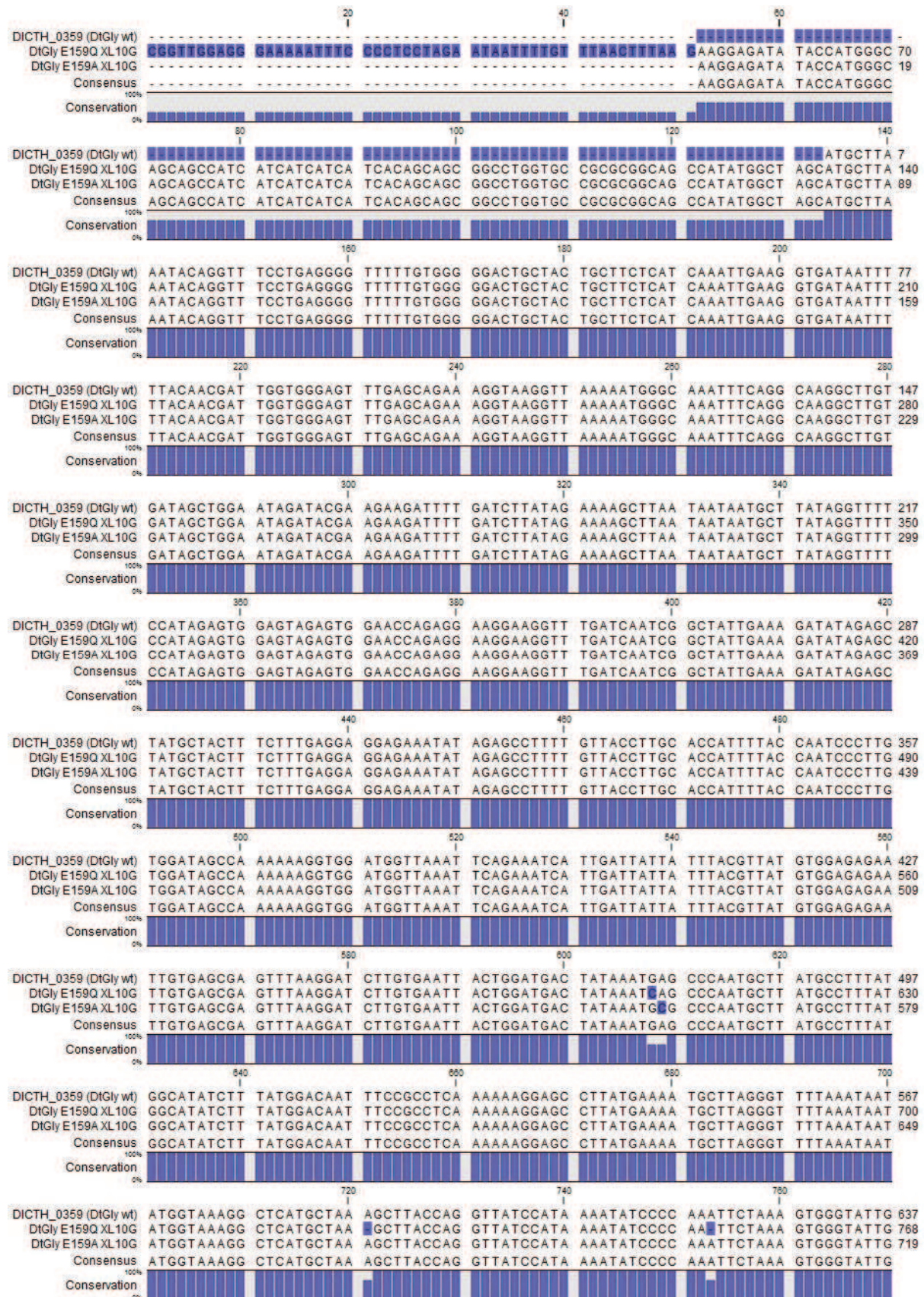


**Annexe 05 : Alignement des séquences protéique de DtGly (B5YCI2) et de glycosidases de structures connues de GH1.** Les protéines sont identifiées selon leur code Uniprot ainsi que leur code PDB donné entre parenthèses. L'homologie des résidus est indiquée suivant le gradient bleu –rouge (bleu = résidus non homologues, rouge = résidus homologues).

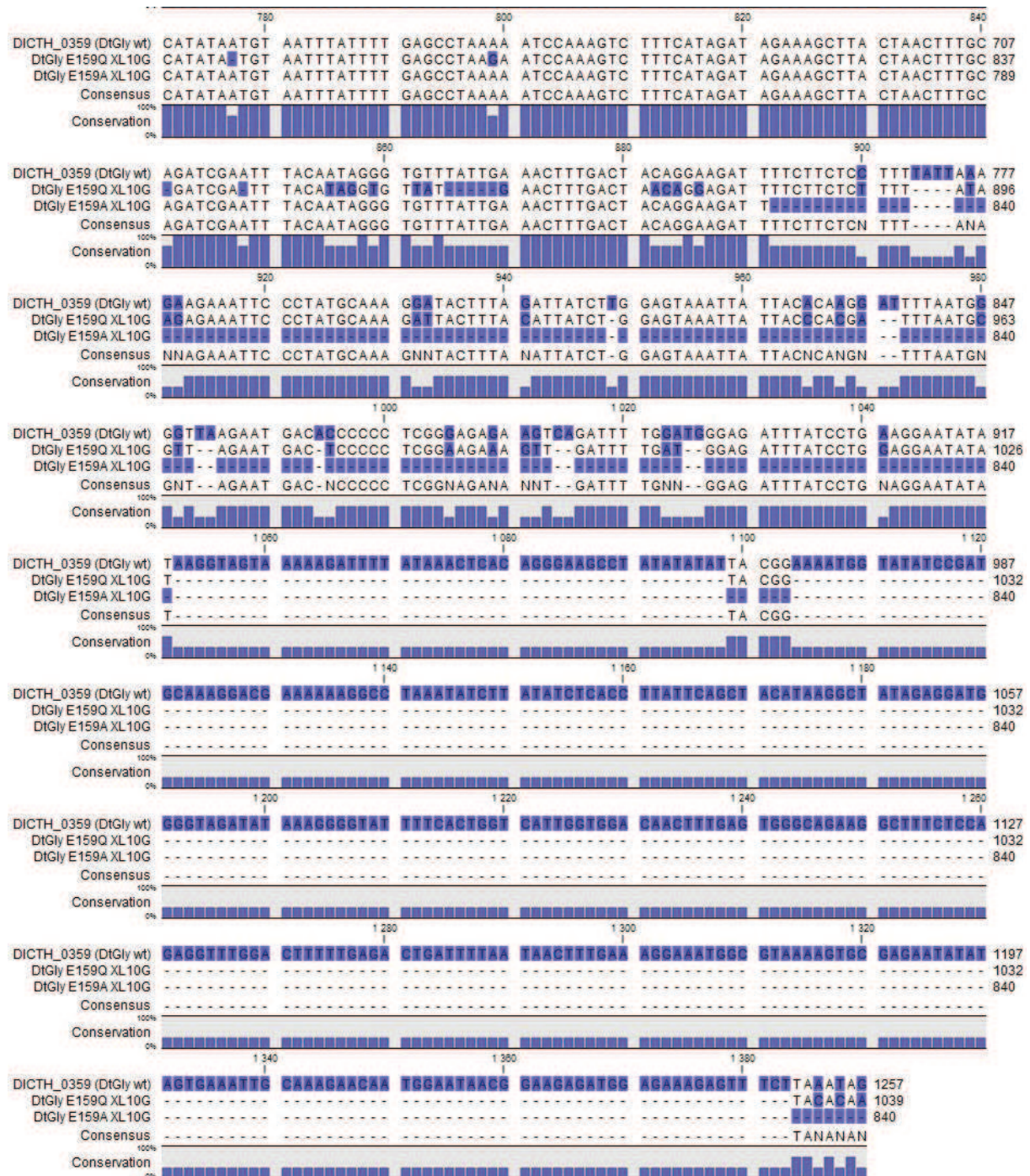




**Annexe 06 : Alignement des séquences ADN de DtGly wt et des mutants DtGly E159Q et E159A.** Les nucléotides non homologues sont indiqués sur fond bleu.

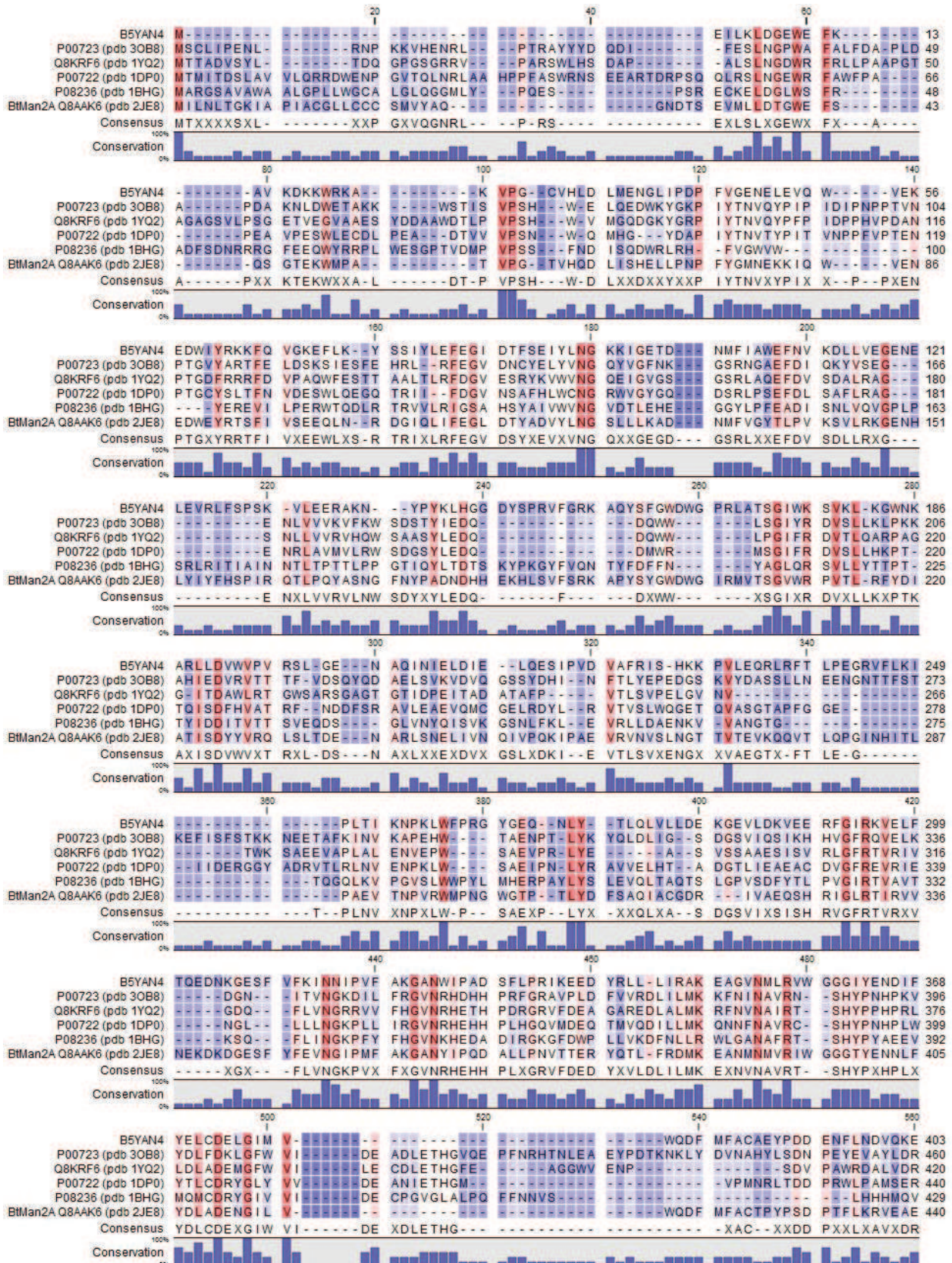






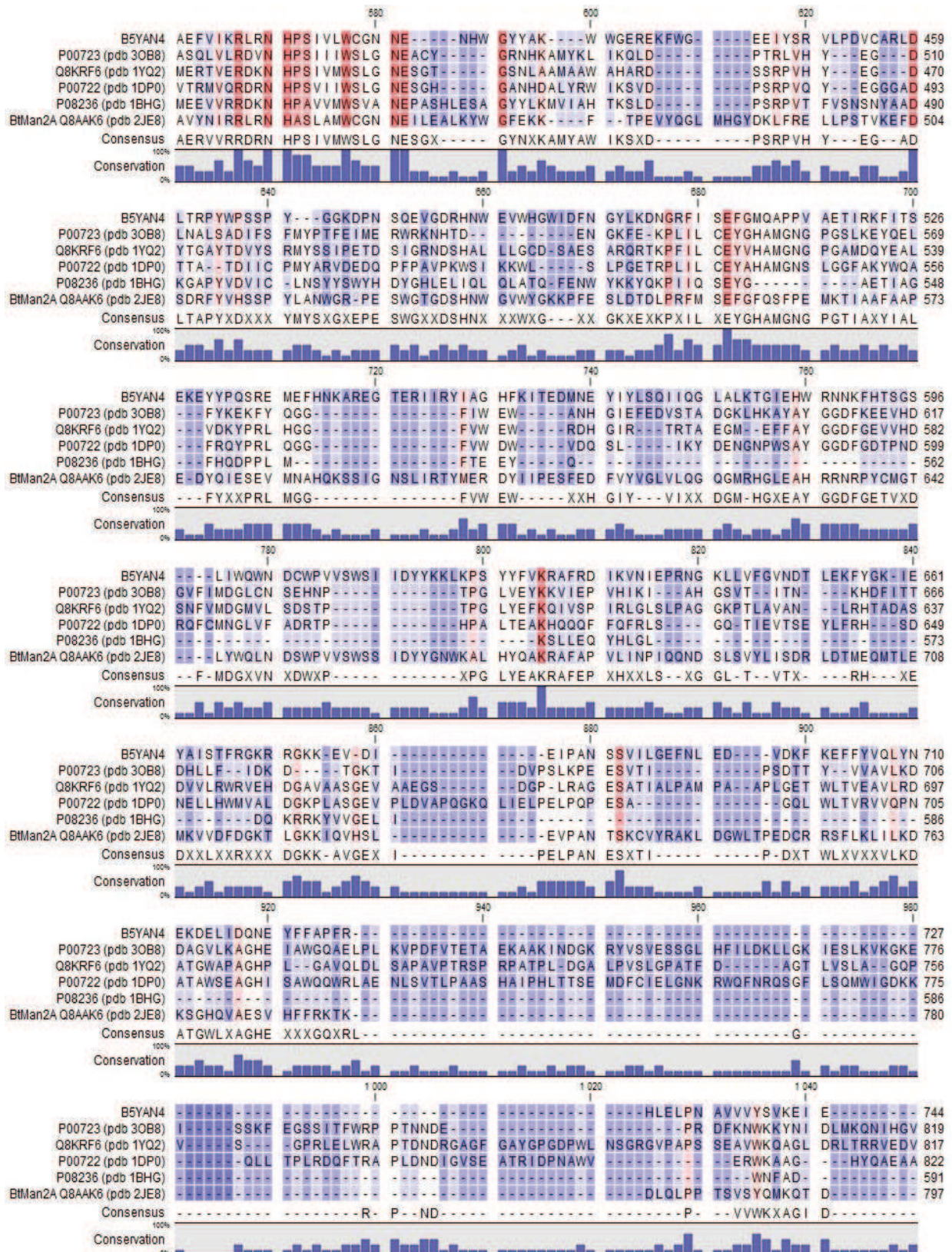


**Annexe 07 : Alignement des séquences peptidiques de DtMan et de glycosidases de structures connues de GH2.** Les protéines sont identifiées selon leur code Uniprot ainsi que leur code PDB donné entre parenthèses. L'homologie des résidu est indiqué suivant le gradient bleu –rouge (bleu = résidus non homologues, rouge = résidus homologues).

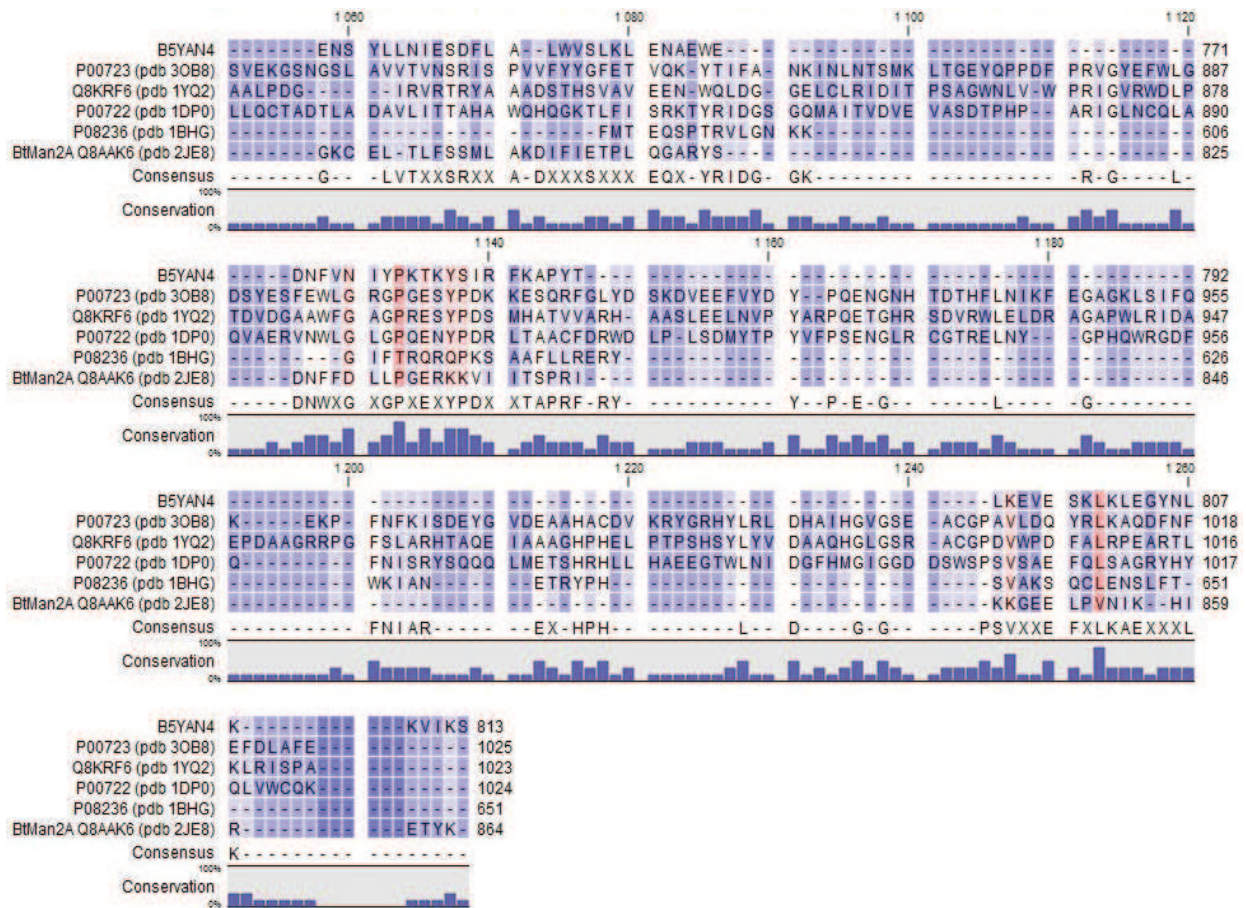




# Annexes



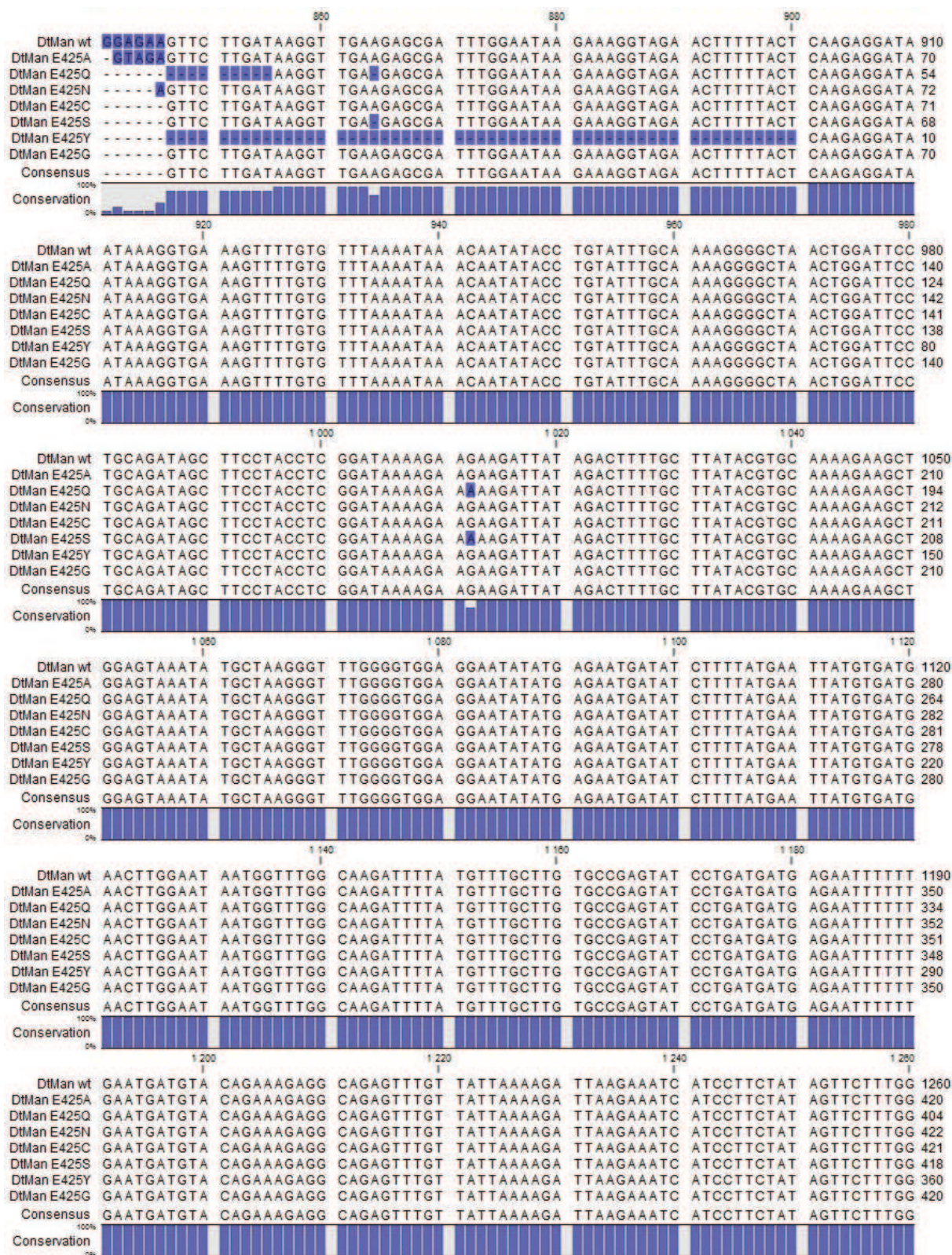






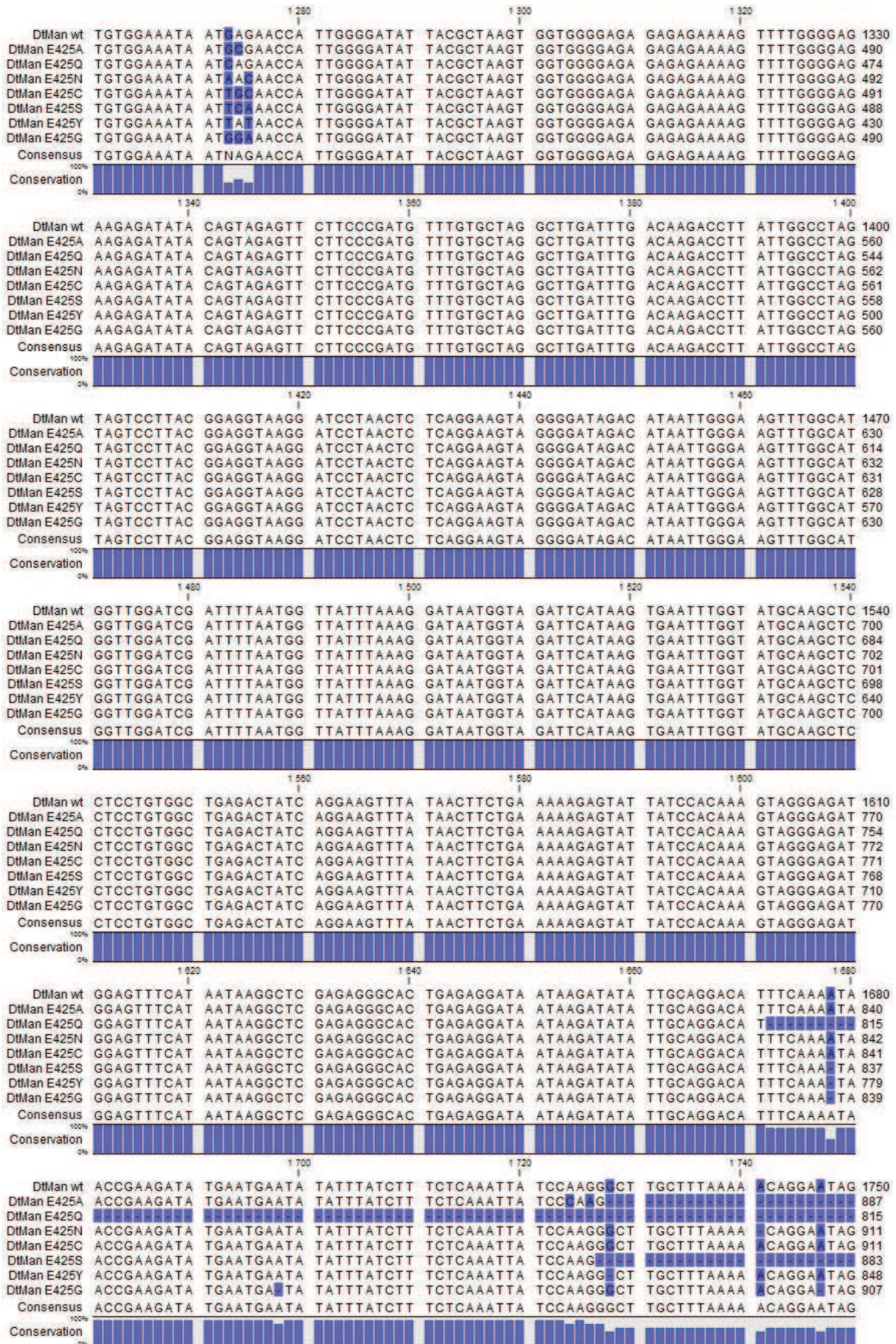
## Annexe 08 : Alignement des séquences ADN de DtMan wt et des mutants DtMan E425X.

Les nucléotides non homologues sont indiqués sur fond bleu.



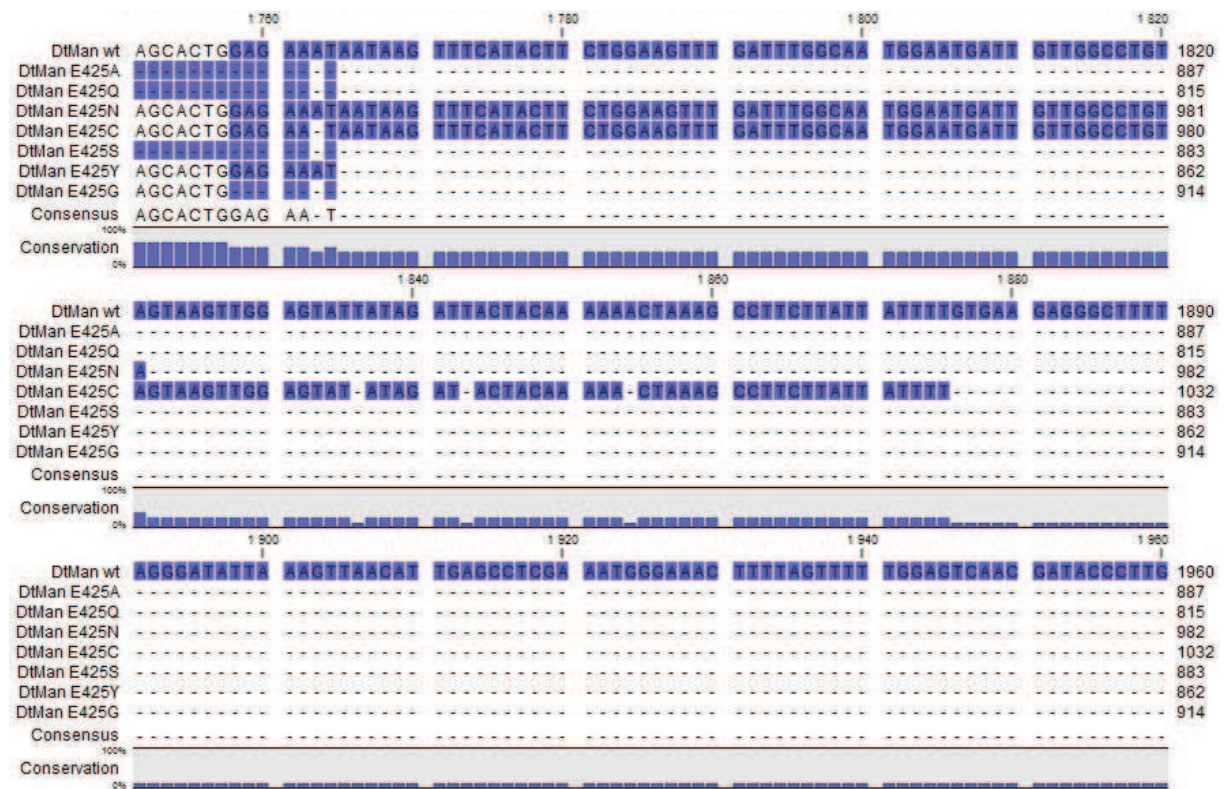


# Annexes





# Annexes





## Développement d'outils enzymatiques pour la synthèse de protéines S-glycosylées

### Résumé :

Dans la communauté scientifique, la glycosylation suscite un vif intérêt tant les relations existantes entre les sucres et les protéines sont étroites et sont impliquées dans de nombreux processus biologiques. De nombreuses méthodologies de synthèse ont été développées pour permettre la production de glycoprotéines sous forme homogènes nécessaires pour leur étude. Les formes S-glycosidiques sont d'un intérêt particulier grâce notamment à la réactivité de l'atome de soufre qui permet le couplage spécifique sucre – acide aminé et qui confère à la liaison une relative résistance à l'hydrolyse acido/basique et enzymatique. Dans le cadre de ce projet nous avons souhaité utiliser des outils biocatalytiques pour proposer une nouvelle voie d'accès aux protéines S-glycosylées en s'appuyant sur le concept des thioglycoligases. Pour répondre à ce challenge, nous nous sommes intéressés à la production d'une banque de glycosidases natives à partir du génome de la bactérie thermophile *Dictyoglomus thermophilum*. Trois protéines originales ont pu être produites, caractérisées et criblées en activité transglycosylase. Parmi ces dernières la  $\beta$ -glycosidase DtGly s'est révélée être un biocatalyseur efficace en favorisant la synthèse d'une variété de glycosides d'alkyle et de glycosyl glycérol. Par la suite, neuf mutants thioglycoligases ont été produits par mutagenèse dirigée à partir des glycosidases natives. A l'issue de leur criblage en activité thioligase, le mutant de DtGly E159Q a été retenu pour la synthèse de S-glycoconjugués. Une étude de relation structure/activité menée sur l'enzyme mutée nous a conduit à la synthèse d'un acide aminé non-naturel, la Thiotyrosine. L'essai de glycosylation de l'analogue estérifié de cette Thiotyrosine par DtGly E159Q a été suivi par LC-MS et a permis de mettre en évidence la formation du produit de thioglycosylation. Ce premier résultat préliminaire s'avère très prometteur pour la validation du concept de S-glycosylation d'un acide aminé catalysée par une thioglycoligase.

Mots clés : biocatalyse, glycoconjugués, thioglycoligase, acide aminé S-glycosylé

## Novel enzymatic tools for S-glycosylated proteins synthesis

### Summary :

Glycosylation, one of the most complex co- and post-translational modifications of proteins, is of utmost importance for the scientific community as carbohydrates and proteins are closely related and involved in numerous diseases. The need to access homogeneous glycoproteins allows the development of a wide range of synthetic methods. Among them S-glycosidic forms have been attractive thanks to the reactivity of sulfur atom which allows the specific sugar – amino acid conjugation and confers a relative resistance through acido/basic or enzymatic hydrolysis. In this project we attempt to take advantage of natural tools to develop an enzymatic methodology to prepare S-glycosylated proteins through the thioglycoligase concept. In order to take up this challenge we first generate an enzymatic pool of wild type glycosidases from the thermophilic bacteria *Dictyoglomus thermophilum*. Activity screening of the three proteins produced and characterized allows the isolation of the  $\beta$ -glycosidase DtGly with good transglycosylation activity, thus affording a variety of alkyl glycosides and glyceroglycosides. Next we generate nine thioglycoligase mutants through site-directed mutagenesis and after screening of activity, the variant DtGly E159Q was selected to produce S-glycoconjugates. Finally, DtGly E159Q relation structure/activity studies led us to synthesize the unnatural amino acid Thiotyrosine. Enzymatic thioglycosylation assay catalyzed by DtGly E159Q was monitored by LC-MS and gave converging evidence of the successful formation of the S-glycosylated Thiotyrosine. This tremendous preliminary result is promising for further investigations of enzymatic S-glycosylation of amino acid using thioglycoligases as toolbox.

Keywords : biocatalysis, glycoconjugates, thioglycoligase, S-glycosylated amino acid



ICOA – Institut de Chimie Organique et Analytique  
UMR 7311 – Université d'Orléans Rue de Chartres – BP  
6759 45067 Orléans Cedex 2 - France

