Études structurales et fonctionnelles d'alpha-glucosidases bactériennes

Magali Dejob

▶ **To cite this version:**

Magali Dejob. Études structurales et fonctionnelles d'alpha-glucosidases bactériennes. Biochimie [q-bio.BM]. Université Claude Bernard - Lyon I, 2013. Français. ⟨NNT : 2013LYO10117⟩. ⟨tel-01172411⟩

HAL Id: tel-01172411

https://theses.hal.science/tel-01172411v1

Submitted on 7 Jul 2015

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

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

HAL Authorization

Numéro d'ordre : 117-2013 Année 2013

THÈSE DE L'UNIVERSITÉ DE LYON

Délivrée par

L'UNIVERSITÉ CLAUDE BERNARD LYON 1

ÉCOLE DOCTORALE INTERDISCIPLINAIRE SCIENCE-SANTÉ

DIPLÔME DE DOCTORAT

(arrêté du 7 août 2006)

soutenue publiquement le 15 Juillet 2013

par

Magali DEJOB

ETUDES STRUCTURALES ET FONCTIONNELLES D'α-GLUCOSIDASES BACTERIENNES

Directrice de thèse : Madame le Docteur N. AGHAJARI

JURY :	Madame le Docteur G. VERONESE,	Rapporteur
	Monsieur le Professeur P. COUTINHO,	Rapporteur
	Monsieur le Docteur G. MICHEL,	Rapporteur
	Madame le Professeur A. GIRARD-EGROT,	Examinateur
	Monsieur le Professeur B. DIJKSTRA,	Examinateur
	Madame le Docteur N. AGHAJARI,	Directrice de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. François-Noël GILLY**
Vice-président du Conseil d'Administration	M. le Professeur Hamda BEN HADID
Vice-président du Conseil des Etudes et de la Vie Universitaire	M. le Professeur Philippe LALLE
Vice-président du Conseil Scientifique	M. le Professeur Germain GILLET
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : M. le Professeur P. FARGE

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. le Professeur F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. FLEURY
Département Chimie Biochimie	Directeur : Mme le Professeur H. PARROT
Département GEP	Directeur : M. N. SIAUVE
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur A. GOLDMAN
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : Mme S. FLECK
Département Sciences de la Terre	Directeur : Mme la Professeure I. DANIEL
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. C. COLLIGNON
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. P. FOURNIER
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. C. VITON
Institut Universitaire de Formation des Maîtres	Directeur : M. A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Administrateur provisoire : M. N. LEBOISNE

REMERCIEMENTS:

Je tiens tout d'abord à remercier le Pr. Gilbert Deléage, Directeur de l'Institut de Biologie et Chimie des Protéines, de m'avoir accueillie au sein de son Unité de Recherche.

Je remercie sincèrement le Dr. Richard Haser qui m'a permis d'intégrer son équipe. Je tiens également à remercier le Dr. Nushin Aghajari de m'avoir encadrée durant ces années, pour ses conseils et sa rigueur. Je voudrais leur témoigner toute ma reconnaissance pour les moyens qu'ils m'ont donnés pour mener à bien ces travaux.

J'adresse toute ma gratitude aux Dr Gabrielle Veronese, Dr Gurvan Michel et Pr. Pedro Coutinho pour avoir accepté d'être les rapporteurs de cette thèse et aux Pr. Bauke Dijkstra et Pr. Agnès Girard-Egrot pour avoir accepter de juger mes travaux.

Je garderai en mémoire toutes ces années passées au LBC en compagnie des « permanents » : Richard, Nushin, Patrice, Christophe, Seb, Fred, Lionel et Xavier, des « étudiants » présents : Romain, Halima et Loic et passés : Laura, Élo, Marie, et des « anciens » : Allison, Moez, Alex et Jen. Merci pour leurs discussions scientifiques, leurs conseils et disponibilités. Je les remercie aussi pour tous les autres moments comprenant entre autres le café du matin (mais pas que), l'horoscope et l'info-pourrie du jour, les discussions (très) variées du midi, les craquages en tous genres (musicaux, canidés, et j'en passe…) et les pauses qui m'ont permis d'apprécier tous les jours ou même supporter les plus difficiles.

Une pensée pour les IBCPiens que j'ai pu côtoyés : p'tit Marco, Ju', Niko, Vale, Juliette, Dorothée, Vincent, Julie, Jerem, Lorena et Burcu. Je pense aussi à mes Stéphanois rencontrés grâce au DUOP : Yann et Guénaëlle (p'tite Savora).

Mention « décalée » pour la musique qui m'a accompagnée tout au long de cette thèse avec en vrac les découvertes, les coups de cœur, etc : SHK.PNK., Massive Attack, Daft Punk, S.O.A.D., Bloc Party, AaRON, Skrillex, Britney, Phoenix, Moderat, Muse, H. Zimmer, P. Kalkebrenner…

Merci à Aurélie qui a su être présente au bon moment.

Je pense à Allice avec qui les discussions de « 5 minutes » ont abouti à une amitié sincère.

Je tiens particulièrement à remercier Anne-Laure (Bio, ma Sœur) et Vince (BDV insurgé et métalleux) qui ont eu les mots et la patience pour me soutenir à renfort de messages, soirées et délires (chalet, mojito, darknious' et autre huile de canard…) tout au long de cette aventure. Je remercie aussi Julien, Emma&Guillaume, Gaz, Gan&Vincent, Adeline (t'en veux ?) et les STU, également impliqués dans les craquages de groupe (top à la vachette, tête de table, trop « swag » et tellement 2009). Je pense également à Christian, Morgan&Marion et Seb (je tends le bâton mais: mo-nu-men-tal).

Pour finir, je remercie Pounet, Nounette et Frérot qui ont toujours cru en moi, notamment quand je n'y croyais plus. Merci de cette confiance infaillible.

Dans les sciences, le chemin est plus important que le but. Les sciences n'ont pas de fin.

Erwin Chargaff

RÉSUMÉ

Il est reconnu, depuis des années, que la flore intestinale par son équilibre complexe et dynamique joue un rôle essentiel dans la santé humaine. Une des stratégies les plus prometteuses pour la maintenir ou l'améliorer consiste à moduler le microbiome par l'utilisation de bactéries probiotiques ou de sucres prébiotiques. C'est dans ce contexte que s'inscrivent les études structurales et fonctionnelles d'α-glucosidases bactériennes développées dans cette thèse. Ces enzymes hydrolysant les liaisons α-(1,4)-glucosidiques sont classées, selon la base de données CAZy, dans les familles de glycoside hydrolases (GH) 4, 13, 31, 63, 97 et 122.

Ces travaux de thèse, centrés sur trois α-glucosidases issues de *Lactobacillus bulgaricus* (11842aglu, GH31), *Lactococcus lactis* (1403aglu, GH13) et *Shewanella* sp. ANA-3 (SHWaglu, GH97), exposent la mise au point de leurs protocoles de surexpression et de purification. Ils présentent également des études bioinformatiques de 11842aglu et de 1403aglu, ainsi qu'une caractérisation enzymatique préliminaire de cette dernière. Une analyse structurale et fonctionnelle approfondie de SHWaglu a aussi été réalisée.

La résolution, par cristallographie aux rayons X, des structures de SHWaglu seule, en complexe avec différents ligands et de mutants, a participé à enrichir les connaissances, jusqu'à présent peu étendues, sur les enzymes de la famille GH97. Ainsi, un motif structural conservé au sein de cette famille a notamment été mis en évidence. Par ailleurs, ces informations structurales combinées aux études enzymatiques ont permis de révéler des déterminants moléculaires de l'activité de cette α-glucosidase et, par conséquent, d'établir les relations structure-fonction-activité de cette enzyme.

Ainsi, l'ensemble des données obtenues, couplé à des études d'ingénierie protéique, contribue à ouvrir de nouvelles perspectives industrielles, notamment en suggérant d'optimiser ou de conférer des activités enzymatiques modifiées dans certaines cibles de choix afin de leur faire synthétiser des sucres de type prébiotiques.

ABSTRACT

It is now generally accepted that the gut flora with its complex and dynamic nature plays a vital role in human health. One of the most promising strategies for maintaining or improving health is to modulate the microbiome by the use of probiotics and prebiotics as food supplements. The structure/function/activity relationship studies of bacterial α-glucosidases described in this thesis have been performed within this context. These α-(1,4)-glucosidic bond hydrolyzing enzymes are classified, according to the CAZy database, into glycoside hydrolases families (GH) 4, 13, 31, 63, 97 and 122.

This thesis work, has focused on three α-glucosidases from *Lactobacillus bulgaricus* (11842aglu, GH31), *Lactococcus lactis* (1403aglu, GH13) and *Shewanella* sp. ANA-3 (SHWaglu, GH97), and the development of their overexpression and purification protocols. It also presents a bioinformatics studies of 11842aglu and 1403aglu, as well as preliminary enzymatic characterization of the latter. As for SHWaglu, detailed structural and functional studies have been carried out.

The crystal structures of SHWaglu in its native state, in complex with different ligands as well as site directed mutants have contributed to increase our knowledge on enzymes from the GH97 family which to date remains relatively limited. Notably, a conserved structural motif in this family has been identified. Overall, the structural- and enzymatic studies and analyses have revealed molecular-and structural determinants governing the activity and broad substrate specificity of this α-glucosidase which is adapted to cold temperatures.

Apart from the insight gained from a fundamental research point of view, data described within this work, coupled with protein engineering studies may contribute to open up new industrial perspectives, in particular by suggesting optimized or altered enzyme activities in some attractive enzyme targets with the aim of synthesizing prebiotic compounds.

TABLE DES MATIERES

AVANT-PROPOS 19
INTRODUCTION BIBLIOGRAPHIQUE 23
Chapitre 1 : Le contexte probiotique 25
I. La flore intestinale 25
A. Un « organe » souvent oublié… 25
B. … et pourtant indispensable 25
1. Rôles dans le métabolisme 26
2. Rôles dans l'immunité 26
3. Des pathologies associées 27
C. Ses origines 27
D. Sa composition générale 27
E. Récapitulatif 29
II. Les probiotiques 30
A. Définition 30
B. Les effets des probiotiques 31
1. Sur la digestion 32
2. Sur l'immunité 32
3. Sur la balance microbienne 33
4. Sur la cholestérolémie 33
5. Sur la tension artérielle 34
6. Sur la carcinogenèse 34
7. Sur l'humeur 34
C. Les microorganismes couramment utilisés 35
III. Les prébiotiques 35
A. Définition 35
B. Les prébiotiques candidats 36
Chapitre 2 : Les enzymes actives sur les sucres 37
I. La base de données « CAZy » (Carbohydrate Active enZyme) 38
II. Les glycoside hydrolases 39
III. Les mécanismes d'hydrolyse 40
A. Inversion de la configuration du carbone anomérique 40
B. Rétention de la configuration du carbone anomérique 40
C. Les variations de mécanisme 42
D. Hydrolyse dépendante du NAD 42
IV. Les topologies du site actif 42
V. Le site actif et les sous-sites de fixation du substrat 43
VI. Les « Carbohydrate Binding Modules » - CBMs 44
Chapitre 3 : Les α-glucosidases 46
I. Description des différentes familles 46
A. Famille GH4 46
1. Activité 46
2. Description structurale 46
B. Famille GH13 47

1. Activité 47
2. Description structurale 48
C. Famille GH31 49
1. Activité 49
2. Description structurale 50
D. Famille GH63 51
1. Activité 51
2. Description structurale 52
E. Famille GH97 53
1. Activité 53
2. Description structurale 53
F. Famille GH122 55
G. Récapitulatif 55
II. Caractéristiques biochimiques 56
A. Spécificité de substrat 56
B. Conditions optimales de réaction 61
C. Récapitulatif des enzymes citées 63
III. Les inhibiteurs 64
A. Les thiosucres 65
B. Les iminosucres 65
1. Déoxynojirimycine 65
2. Castanospermine 66
C. Les carbasucres 66
D. Le Tris 67
Chapitre 4 : Les applications et la bio-ingénierie 68
Chapitre 5 : Les α-glucosidases du projet 70
I. Issue de *Shewanella* sp. ANA-3 70
A. Informations sur la souche 70
B. Informations sur l'enzyme 72
II. Issue de *Lactococcus lactis* IL1403 74
A. Informations sur la souche 74
B. Informations sur l'enzyme 75
III. Issue de *Lactobacillus bulgaricus* ATCC11842 77
A. Informations sur la souche 77
B. Informations sur l'enzyme 77
MATERIEL, METHODES ET PRINCIPES 79
Chapitre 1 : Techniques de biologie moléculaire 81
I. Plasmides 81
II. Techniques classiques de manipulation de l'ADN 84
A. Transformation de bactéries 84
B. Extraction et purification de l'ADN plasmidique 84
1. Miniprep 84
2. MidiPrep 84
C. Mesure de la concentration en ADN 85
D. Digestion de l'ADN par des endonucléases de restriction 85

E. Électrophorèse sur gel d'agarose 85
F. Extraction sur gel d'agarose 85
G. Ligature 86
H. Technique de PCR 86
I. Mutagenèse dirigée 87
III. Séquençage d'ADN 88
Chapitre 2 : Techniques de biochimie 90
I. Surproduction des protéines de fusion 90
A. Culture bactérienne et induction de la surexpression 90
1. Milieux de culture 90
2. Mise en culture et induction de la surexpression 90
B. Lyse cellulaire 90
II. Purification des protéines d'intérêt par chromatographie 91
A. La chromatographie d'affinité avec métaux immobilisés (IMAC) 92
1. Principe 92
2. Protocole 92
B. La chromatographie échangeuse d'ions 93
1. Principe 93
2. Protocole 94
C. La chromatographie d'exclusion 94
1. Principe 94
2. Protocole 95
III. Révélation des protéines 95
A. Électrophorèse en gel de polyacrylamide en condition dénaturante 95
1. Principe 95
2. Protocole 96
B. Révélation des protéines 97
1. Révélation au bleu de Coomassie colloïdal 97
2. Révélation au nitrate d'argent 97
3. Western-Blot 97
IV. Digestion par la thrombine 99
V. Analyse de l'état oligomérique des protéines 99
Chapitre 3 : Études fonctionnelles 100
I. Principe : 100
A. Modèle de Michaelis-Menten 100
B. Inhibition de l'activité enzymatique 103
II. Mise en application 103
A. Hydrolyse du *para*-nitrophényl-α-D-glucopyranoside (*p*NPG) 103
1. Détermination des conditions optimales de réaction 103
2. Détermination des paramètres cinétiques de l'α-glucosidase de Shewanella sp.ANA-3 pour le pNPG et les ligands 104
3. Détermination des paramètres cinétiques de l'α-glucosidase de Shewanella sp.ANA-3 pour les substrats naturels 105
B. Criblage des substrats de l'enzyme par chromatographie sur couche mince ... 105
1. Principe 105
2. Protocole 106

C. Détermination de l'anomérie de produit libéré 106
1. Principe 106
2. Protocole 106
III. Immobilisation d'enzyme 107
Chapitre 4 : Études structurales 108
I. Dichroïsme circulaire 108
A. Principe 108
B. Protocole 108
II. Cristallogenèse 109
A. Principe 109
B. Méthode 109
C. Trempage et co-cristallisation 110
1. Trempage 110
2. Co-cristallisation 111
D. Ensemencement 111
E. Cryoprotection des cristaux 111
F. Déshydratation 112
III. Cristallographie rayons X 112
A. Obtention et enregistrement des données de diffraction 112
1. Le principe de diffraction 112
2. Enregistrement des données 113
3. Le problème des phases 116
4. Affinement du modèle : vers la structure finale 118
5. Les cartes de densité électronique 120
IV. « Docking » 121
A. Méthode de « docking » 121
B. La fonction de score 122
C. Les logiciels majeurs 122
RESULTATS ET DISCUSSION 123
Chapitre 1 : Étude de l'α-glucosidase de *Lactobacillus bulgaricus* ATCC11842 125
I. Étude bioinformatique 125
A. Séquence complète 125
B. Région C-terminale 126
II. Biologie moléculaire 128
III. Expression et purification 129
A. Expression et extraction de la protéine soluble 129
B. Purification 130
Chapitre 2 : Étude de l'α-glucosidase de *Lactococcus lactis* IL1403 133
I. Étude bioinformatique 133
II. Biologie moléculaire 134
III. Expression et purification 135
IV. Caractérisation biochimique 137
A. Caractérisation préliminaire 137
1. Détermination du pH optimum 138
2. Détermination de la température optimale 139

3. Effets des ions divalents sur l'activité catalytique 139
4. Criblage des substrats naturels de 1403aglu 140
5. Stabilité de 1403aglu dans le temps 142
B. Détermination des constantes catalytiques 142
V. État oligomérique de la protéine en solution 142
VI. Étude structurale 143
A. Cristallogenèse 143
B. Dichroïsme circulaire 145
Chapitre 3 : Étude de l'α-glucosidase de *Shewanella* sp. ANA-3 147
I. Biologie moléculaire 147
II. Expression et purification de SHWaglu 148
III. État oligomérique de la protéine en solution 150
IV. Étude fonctionnelle 151
A. Caractérisation préliminaire 151
1. Impact de la concentration cellulaire dans les billes d'alginate 152
2. Détermination du pH optimal 152
3. Détermination de la température optimale 153
4. Effets des ions divalents sur l'activité catalytique 154
5. Le rôle des ions chlorure 155
6. Thermostabilité 156
7. Stabilité vis-à-vis du pH 157
8. Impacts de l'immobilisation sur l'activité catalytique 158
9. Criblage des substrats naturels de SHWaglu 159
B. Détermination des constantes catalytiques 161
1. Les substrats 161
2. Les inhibiteurs et autres ligands 162
C. Confirmation du mécanisme catalytique 163
V. Étude structurale 164
A. Cristallogenèse 164
B. Cristallographie aux rayons X 166
1. Résolution de la structure 166
2. Description de la structure de SHWaglu 168
3. Études structurales de SHWaglu native en complexe avec des inhibiteurs 178
4. Études structurales de SHWaglu native en complexe avec des substrats naturels 183
C. Comparaison de structures 188
D. Docking avec le panose 191
IV. Études structurales de mutants 193
A. Biologie moléculaire, expression et purification 193
B. Études cristallographiques des mutants E485Q et K434A 194
CONCLUSIONS & PERSPECTIVES 199
RÉFÉRENCES BIBLIOGRAPHIQUES 207
PUBLICATIONS 231

LISTE DES ABREVIATIONS

ADN :	Acide désoxyribonucléique
APS :	Ammonium PerSulfate
B. thetaiotaomicron :	*Bacteroides thetaiotaomicron*
BET :	Bromure d'éthidium
BLAST :	Basic Local Alignment Search Tool
C-terminal :	carboxy-terminal
Ca^{2+} :	ion calcium
CAZy :	Base de données Carbohydrate Active enZYmes
CCM :	Chromatographie sur Couche Mince
CCP4 :	Collaborative Computational Project N°4
Cl^- :	ion chlorure
Co^{2+} :	ion cobalt
dNTP :	désoxyribonucléotide
D.O. :	Densité Optique
DP :	Degré de Polymérisation
E. coli :	*Escherichia coli*
EC :	Enzyme Commission
EDTA :	Acide Ethylène Diamine Tétraacétique
Fe^{2+} :	ion fer
FPLC :	Fast Performance Liquid Chromatography
GH :	Glycoside Hydrolase
HEPES :	Acide 4-(2-hydroxyéthyl)-1-Pipérazine Ethane Sulfonique
IPTG :	IsoPropyl β-D-1-ThioGalactopyranoside
IMAC :	Immobilized Metal Ion Affinity Chromatography
IUPAC-IUBMB :	International Union of Pure and Applied Chemistry-International Union of Biochemistry and Molecular Biology
kb :	kilopaire de base
k_{cat}:	constante catalytique
kDa :	kiloDalton
KEGG :	Kyoto Encyclopedia of Genes and Genomes
K_m :	constante de Michaelis
L. bulgaricus :	*Lactobacillus bulgaricus* ATCC11842
L. lactis :	*Lactococcus lactis* IL1403

LB :	Luria Broth
Li^+ :	ion lithium
MES :	2-(N-morpholino)ethanesulfonic acid
Mg^{2+} :	ion magnesium
MME :	MonoMethyl Ether
Mn^{2+} :	ion manganèse
MOPS :	3-(N-morpholino)propanesulfonic acid
N-terminal :	amino-terminal
NAD :	nicotinamide adénine dinucléotide
Ni^{2+} :	ion nickel
NTA :	acide nitriloacétique
***p*NPG** :	*p*-NitroPhényl-α-D-Glucopyranoside
PAGE :	PolyAcrylamide Gel Electrophoresis
pb :	paire de base
PBS :	Phosphate Buffer Saline
PCR :	Polymerase Chain Reaction
PDB :	Banque de données Protein Data Bank
PEG :	PolyEthylène Glycol
pI :	point Isoélectrique
***p*NPG** :	*para*-nitrophenyl-α-D-glucopyranoside
qsp :	quantité suffisante pour
RMSD :	Root Mean Square Deviation
rpm :	rotations par minute
SDS :	Sodium Dodecyl Sulfate
Shewanella :	*Shewanella* sp. ANA-3
TEMED :	TEtraMethylEthyleneDiamine
Tris :	Trishydroxyméthylaminométhane
UV :	UltraViolet
X-Gal :	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Zn^{2+} :	ion zinc

Avant-propos

Ce travail de thèse a débuté en septembre 2009. Il a été réalisé sous la direction du Dr. Nushin Aghajari au sein de l'équipe de Biocristallographie et de Biologie Structurale des Cibles Thérapeutiques (BBSCT) à l'Institut de Biologie et Chimie des Protéines (IBCP, FR3302) dans l'Unité « Bases Moléculaires et Structurales des Systèmes Infectieux », (BMSSI, UMR5086).

Mon principal sujet de recherche a été l'étude structurale et fonctionnelle de trois α-glucosidases issues de souches bactériennes : *Lactobacillus bulgaricus* (11842aglu), *Lactococcus lactis* (1403aglu) et *Shewanella* sp. ANA-3 (SHWaglu). Ces trois projets parallèles ont été menés avec des degrés d'avancement différents même si, dans tous les cas, les protocoles d'expression et de purification ont été établis. Ces travaux ont permis notamment de résoudre par remplacement moléculaire et à haute résolution la structure de SHWaglu seule (1,77 Å), en complexe avec différents ligands (entre 1,85 et 2 Å) et deux mutants (2,25 et 2,7 Å). Ce manuscrit décrit également les études fonctionnelles de SHWaglu et l'établissement de relations structures-fonction-activité pour cette dernière.

Au cours de ces années, j'ai également eu l'opportunité de m'impliquer dans un second sujet d'étude portant sur une β-galactosidase issue de *Streptococcus thermophilus* et qui a fait l'objet d'une publication (Rhimi *et al.*, 2010). Ce sujet est issu d'une collaboration entre l'équipe de BBSCT et l'équipe « Interactions des Firmicutes avec l'Environnement » de l'UMR1319 « MICALIS-Microbiologie de l'Alimentation au Service de la Santé Humaine » de l'INRA.

Les β-galactosidases sont décrites depuis plusieurs années comme ayant un fort potentiel industriel. Elles sont très couramment employées dans le secteur agro-alimentaire, essentiellement pour réduire la concentration en lactose des produits laitiers afin de surmonter l'intolérance au lactose. De plus, l'utilisation de β-galactosidases immobilisées en tant que bioconvertisseurs efficaces à l'échelle industrielle a déjà été rapportée. L'hydrolyse du lactose dans les produits laitiers fermentés nécessite une β-galactosidase efficace à des pH acides et sur une large gamme de températures. En outre, le petit-lait, qui est une matière première peu coûteuse avec un pH acide et riche en lactose, est une source attractive pour la production d'édulcorants hypocaloriques. Dans ce cadre, la β-galactosidase issue de *Streptococcus thermophilus*, qui est une souche probiotique déjà utilisée pour la production des yaourts, est une cible de choix.

Ce second sujet ne sera pas développé dans ce manuscrit. Néanmoins, il m'a apporté les bases pour mener à bien l'étude enzymatique d'une glycoside hydrolase. Plus encore, ces travaux m'ont permis de m'initier aux techniques de biologie moléculaire et de biochimie.

Publications :

Dejob M., Rhimi M., Ilhammami R., Haser R. & Aghajari N. (2013) *Shewanella* sp. ANA-3 GH97 α-glucosidase is active at low temperatures and chloride dependent. *Soumis.*

Rhimi M., Boisson A., Dejob M., Boudebouze S., Maguin E., Haser R. & Aghajari N. (2010) Efficient bioconversion of lactose in milk and whey: immobilization and biochemical characterization of a β-galactosidase from the dairy *Streptococcus thermophilus* LMD9 strain. Research in Microbiology, **161**, 515-525.

Introduction bibliographique

Chapitre 1 : Le contexte probiotique

I. La flore intestinale

A. Un « organe » souvent oublié…

Il est reconnu, depuis des décennies, que le corps humain contient dix fois plus de cellules bactériennes que de cellules humaines (Savage, 1977). Ces microorganismes colonisent pratiquement toute surface exposée à l'environnement extérieur (comme la peau ou encore les voies respiratoires). Toutefois, l'un des sites du corps le plus notable est le tractus gastro-intestinal (TGI) avec une surface d'environ 400 m² (Shanahan, 2002) ce qui en fait de loin l'organe le plus colonisé, avec une masse totale en microorganismes proche de 1,5 kilogrammes (Xu & Gordon, 2003) et 500 à 1000 espèces bactériennes distinctes recensées (Xu *et al.*, 2003).

Avec une communauté microbienne de plusieurs milliards de microorganismes, notre intestin est le site d'un équilibre extraordinairement complexe et dynamique, basé sur le mutualisme (Hooper & Gordon, 2001). Ainsi, le microbiote intestinal (ou flore intestinale) joue un rôle primordial dans les processus métaboliques, nutritionnels, physiologiques et immunologiques du corps humain (Gerritsen *et al.*, 2011). Collectivement, les membres de la flore ont une activité métabolique comparable à un organe virtuel (Bocci, 1992; Rastall, 2004; Backhed *et al.*, 2005). Plus encore, l'ensemble de leurs génomes est supérieur au génome nucléaire humain de plus de cent fois (Backhed *et al.*, 2005; Ley *et al.*, 2008; Qin *et al.*, 2010) avec neuf millions de gènes uniques (Zoetendal *et al.*, 2008; Yang *et al.*, 2009). Le terme de « microbiome » a été proposé par le prix Nobel Joshua Lederberg pour décrire ce génome collectif, l'idée étant que la vision génétique d'ensemble de *Homo sapiens* comme forme de vie devait inclure les gènes de ce microbiome (Hooper & Gordon, 2001). Par extension, le microbiote désigne aujourd'hui l'ensemble des microorganismes vivant dans un environnement donné.

B. … et pourtant indispensable

Des études sur des animaux non colonisés par des microbes de l'environnement extérieur (appelés « *germ-free* ») ont montré une influence très importante du microbiote sur l'organisme hôte. En effet, il a été constaté une réduction du renouvellement des cellules de la muqueuse (Savage *et al.*, 1981), de l'activité des enzymes digestives, de l'angiogenèse, ainsi que de l'épaisseur de la paroi musculaire ou bien encore de la motilité (Shanahan,

2002). Il a été démontré que la flore microbienne modulait l'expression de gènes de l'hôte impliqués dans des fonctions diverses et physiologiques (Hooper *et al.*, 2001), mais aussi qu'elle permettait de réduire l'apport calorique nécessaire pour maintenir la masse corporelle (Wostmann *et al.*, 1983). Il a également été rapporté le rôle émergent de la flore commensale sur l'humeur et le comportement, ceci impliquant une communication bidirectionnelle selon un axe microbiome-intestin-cerveau. En effet, la colonisation bactérienne joue un rôle majeur dans la maturation des systèmes endocrine et immunitaire, revêtant une importance toute particulière dans le développement du système sérotoninergique (Clarke *et al.*, 2012).

1. Rôles dans le métabolisme

Les premières études métagénomiques ont démontré que, comparativement au génome humain, le microbiome intestinal humain était fortement enrichi pour les catégories de gènes impliqués dans le métabolisme ; gènes trouvés dans les noyaux de groupes orthologues (COGs et la base de données KEGG) (Gill *et al.*, 2006; Kurokawa *et al.*, 2007; Turnbaugh *et al.*, 2009). Aussi, les voies métaboliques de certains sucres, acides aminés, vitamines, nucléotides, cofacteurs ou de biosynthèse de métabolites secondaires sont fortement représentées dans le microbiome humain. Ces voies sont non seulement nécessaires aux microorganismes mais ont également des effets sur l'hôte lui-même (Gerritsen *et al.*, 2011). Au final, les interactions métaboliques entre les microorganismes et l'hôte sont bénéfiques pour les deux parties.

2. Rôles dans l'immunité

Les contributions les plus importantes et les plus étudiées du microbiote sont celles liées au développement et à la fonction du système immunitaire (Maslowski *et al.*, 2009; Lathrop *et al.*, 2011; Hooper *et al.*, 2012). Un système immunitaire plus robuste permet de mieux combattre les pathogènes microbiens et, au final, de fournir une niche plus saine pour les bactéries commensales. Cette vision implique que les mécanismes de l'hôte et de la flore ont évolué conjointement afin de développer des stratégies synergiques contre les agents infectieux (Lee & Mazmanian, 2010).

Cependant, le défi que le système immunitaire de l'hôte doit relever est de distinguer les microorganismes bénéfiques des pathogènes car leurs motifs moléculaires sont similaires et reconnus par le système inné. Il semblerait donc que la discrimination se fasse grâce au système immunitaire adaptatif et à une stimulation de la tolérance de la muqueuse (Lee & Mazmanian, 2010; Round & Mazmanian, 2010).

3. Des pathologies associées

Il a été démontré que des anormalités dans la composition de la flore microbienne engendraient des pathologies telles que les maladies inflammatoires chroniques intestinales (comme la maladie de Crohn ou colite ulcéreuse) ou entérocolite nécrosante (Fell, 2005; Xavier & Podolsky, 2007). Ces pathologies auraient pour origine une activation inappropriée du système immunitaire, déterminant une perte de tolérance aux bactéries commensales avec, comme conséquence, une réponse dirigée contre le microbiote intestinal (Hooper & Gordon, 2001; Sharon *et al.*, 2013). De plus, des perturbations de l'équilibre microbien peuvent affecter le métabolisme de l'hôte, le prédisposant ainsi à l'obésité ou au diabète (Musso *et al.*, 2011).

C. Ses origines

Généralement considéré comme stérile, le tractus intestinal fœtal est immédiatement colonisé dès la naissance par un transfert maternel et est influencé par le mode d'accouchement, le régime alimentaire de l'enfant, le degré d'hygiène ou bien encore la médication (Murgas Torrazza & Neu, 2011; Costello *et al.*, 2012; Nicholson *et al.*, 2012). Une étude, montrant la présence de bactéries dans le méconium, a toutefois évoqué que l'apparition de la flore bactérienne pouvait être encore plus précoce, ce qui suggère un début de colonisation intra-utérine expliquant l'origine des premiers colonisateurs de l'intestin. L'évolution de la composition de la flore commensale sera ensuite conditionnée par le mode d'alimentation du nourrisson : le lait maternel (Jiménez *et al.*, 2008), contrairement au lait infantile, est enrichi en oligosaccharides capables de stimuler sélectivement la croissance des bifidobactéries (Zivkovic *et al.*, 2011). Ces dernières ainsi que les entérobactéries semblent être des colonisateurs précoces (Mountzouris *et al.*, 2002). Ces bactéries pionnières peuvent moduler l'expression de gènes chez l'hôte en vue de créer un environnement qui leur est favorable et de prévenir la croissance d'autres bactéries introduites plus tard dans l'écosystème (Xu & Gordon, 2003). Des changements rapides et d'envergure de la communauté microbienne ont ainsi été constatés pendant la colonisation intestinale chez l'enfant (Sharon *et al.*, 2013).

D. Sa composition générale

Il apparait que chaque individu porte une flore intestinale ayant un caractère unique (Human Microbiome Project, 2012), stable dans le temps chez les hôtes sains et possédant une spécificité d'hôte et de lieu (Rajilic-Stojanovic *et al.*, 2007; Zoetendal *et al.*, 2008).

Toutefois, une étude a montré que presque 40% des gènes du microbiome d'un organisme donné sont retrouvés chez la moitié des individus de la cohorte analysée (Qin *et al.*, 2010).

Des études de métagénomique démontrent qu'il existe des variations de communautés bactériennes sur toute la longueur du tractus intestinal. Ainsi, le microbiome de l'intestin grêle semble être moins complexe que celui du côlon avec une diversité d'espèces plus réduite (Zoetendal *et al.*, 2012). Sont retrouvés aussi bien des microorganismes anaérobies facultatifs comme les lactobacilles, les streptocoques et les entérobactéries que des anaérobies strictes comme les Bifidobacteria, les Bacteroidetes et les Clostridia (Rastall, 2004; Zoetendal *et al.*, 2012). La flore du côlon, quant à elle, est dominée par les bactéries anaérobies strictes comme les Bacteroidetes et Clostridia (Figure 1), l'une des plus connues étant *Bacteroides thetaiotaomicron*. Autre distinction notable, les bactéries anaérobies facultatives sont 1000 fois moins nombreuses que les anaérobies strictes (Rastall, 2004; Yang *et al.*, 2009).

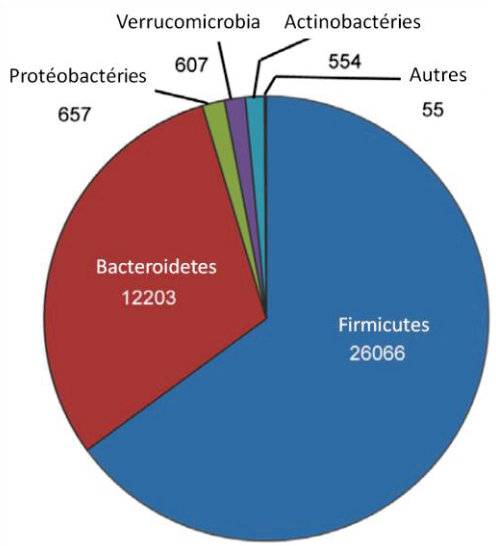


Figure 1 : Composition bactérienne du côlon (Yang *et al.*, 2009). Les Clostridium sont classés dans les Firmicutes. Les valeurs correspondent au nombre de séquences génomiques.

Même si le microbiote intestinal est relativement stable, des fluctuations intra-individuelles peuvent survenir à cause de changement environnementaux, d'événements pathologiques ou encore en fin de vie. En outre, l'utilisation d'antibiotiques affecte non seulement les pathogènes, mais aussi les bactéries bénéfiques. Il est donc possible de rompre l'équilibre, ceci conduisant à des perturbations de la composition microbienne (Gerritsen *et al.*, 2011). Problème collatéral : il a été rapporté une augmentation de résistance aux antibiotiques résultant d'un transfert de gènes de résistance entre les microorganismes (Jernberg *et al.*, 2011).

E. Récapitulatif

Le microbiote intestinal humain, fondé sur un équilibre complexe et dynamique, est un atout pour la santé de l'hôte. Cet écosystème est caractérisé par des interactions réciproques et dynamiques avec l'hôte, retrouvées à différents niveaux tant métabolique qu'immunitaire. Les bactéries le composant sont principalement du genre Firmicutes et Bacteroidetes anaérobies strictes chez lesquelles sont retrouvées des souches pathogènes, neutres ou bénéfiques pour l'organisme colonisé (Figure 2).

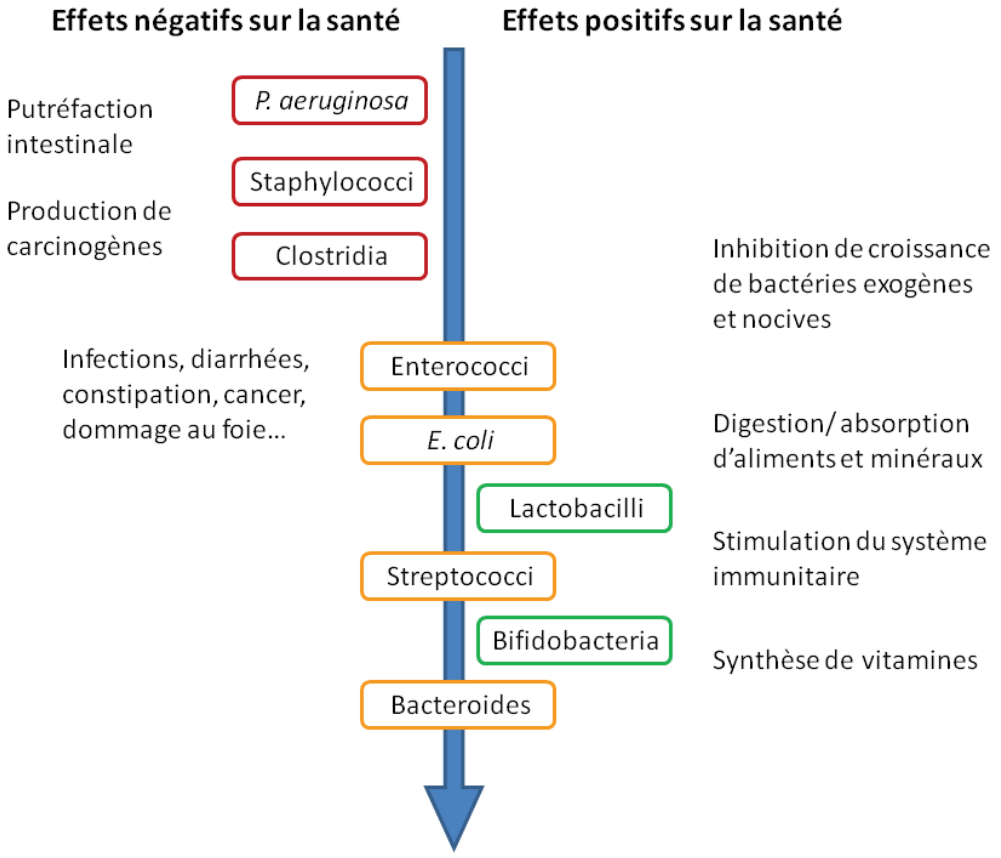


Figure 2 : Vue d'ensemble du microbiote humain (adapté de Rastall, 2004). *Les genres bactériens sont énumérés selon les effets sur la santé (positifs, négatifs ou variables).*

II. Les probiotiques

Depuis qu'il est reconnu que la flore intestinale joue un rôle important dans la santé humaine et certaines pathologies, les manipulations de ces microorganismes par des antibiotiques, probiotiques et prébiotiques sont devenues des approches séduisantes pour améliorer ou maintenir la santé (Preidis & Versalovic, 2009; Gareau *et al.*, 2010).

Cependant, comme dit précédemment, les antibiotiques touchent non seulement les pathogènes, mais aussi les communautés microbiennes commensales intestinales ; ce qui n'en font pas des composés idéaux. Une stratégie très prometteuse (Gerritsen *et al.*, 2011) consiste en la modulation du microbiome de façon plus ciblée par l'utilisation de probiotiques et/ou prébiotiques (l'association étant appelée « synbiotique »).

A. Définition

Le terme de « probiotique » a pour la première fois été utilisé en 1965 pour décrire « des substances sécrétées par un microorganisme stimulant la croissance d'un autre » (Lilly & Stillwell, 1965), en opposition au terme d' « antibiotique ». Parker les définit ensuite comme des « organismes et substances contribuant à la balance microbienne intestinale » (Parker, 1974), affirmation améliorée en 1989 comme suit : « complément alimentaire microbien vivant affectant de façon bénéfique l'animal hôte par une amélioration de la balance microbienne intestinale » (Fuller, 1989). Au cours des années, la définition s'est de plus en plus affinée (comme développé dans Schrezenmeir & de Vrese, 2001) pour, au final, être formulée par l'Organisation Mondiale de la Santé comme suit : « des microorganismes vivants qui, lorsqu'ils sont administrés en quantité suffisante, confèrent un avantage pour la santé de l'hôte » (FAO/WHO, 2002) ; ceci expliquant l'intérêt grandissant de la communauté scientifique pour ces microorganismes (Figure 3).

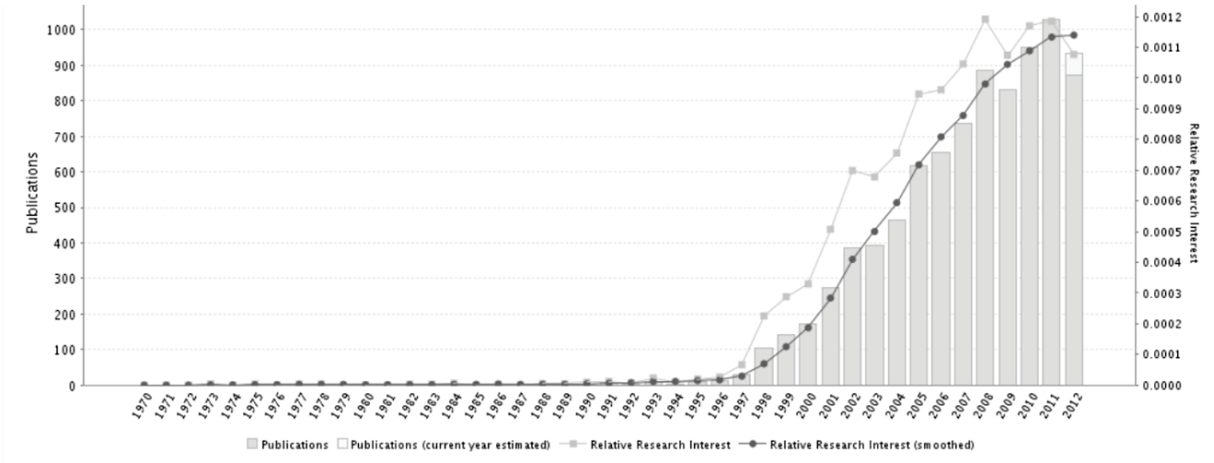


Figure 3: Nombre d'articles publiés par an par la communauté scientifique concernant les probiotiques (généré par GoPubMed®).

B. Les effets des probiotiques

De nombreux effets bénéfiques pour la santé ont été attribués à ces microorganismes (Ng *et al.*, 2009; Iannitti & Palmieri, 2010). Généralement, les probiotiques sont classés en trois niveaux d'action (Gerritsen *et al.*, 2011) :

- (1) Une action directe dans le tractus intestinal, par exemple par interaction directe avec le microbiote ou par des activités enzymatiques.
- (2) Une interaction directe avec la couche de mucus intestinal et l'épithélium, qui influence la fonction de barrière intestinale et le système immunitaire muqueux.
- (3) Ils peuvent avoir des effets en dehors du TGI, par exemple sur le système immunitaire systémique et sur les autres organes comme le foie et le cerveau.

Bien que des données sur des observations *in vivo* commencent à émerger, la majorité des études mécanistiques ont été réalisées *in vitro*, *ex vivo* ou à l'aide de modèles animaux. De ce fait, des réserves peuvent être émises sur la réelle efficacité clinique *in vivo* (Gerritsen *et al.*, 2011). Des études sur des mélanges de différents probiotiques commencent à pointer les avantages d'une action additive ou synergique en vue d'une meilleure efficacité (Timmerman *et al.*, 2004; Chapman *et al.*, 2011). Chapman a démontré plus tard que les souches probiotiques étaient capables d'inhiber au moins une autre de ces souches quand elles sont utilisées dans une préparation multi-souches, avec toutefois des statistiques montrant une meilleure efficacité sur les pathogènes avec un mélange qu'avec des souches seules (Chapman *et al.*, 2012).

1. Sur la digestion

Une amélioration des symptômes de l'intolérance au lactose a été observée lors de consommation de bactéries lactiques. Ceci serait dû aux activités lactases bactériennes qui compenseraient la déficience en β-galactosidase de l'hôte (de Vrese *et al.*, 2001). Mieux encore, une étude récente montre que les effets peuvent persister au moins trois mois après une consommation de quatre semaines de la suspension de probiotiques (Almeida *et al.*, 2012).

Une amélioration au niveau du transit intestinal a également été constatée, plus particulièrement chez les sujets dont le transit est ralenti (Marteau *et al.*, 2002; Waller *et al.*, 2011) sans pour autant que cet effet ne soit expliqué à ce jour.

Les probiotiques peuvent également réduire les problèmes de diarrhées, qu'elles soient causées par un rotavirus (Dubey *et al.*, 2008; Narayanappa, 2008) ou un traitement antibiotique (Hickson *et al.*, 2007; Ruszczynski *et al.*, 2008; Szymanski *et al.*, 2008; Johnston *et al.*, 2011).

2. Sur l'immunité

Les probiotiques ont des propriétés immunomodulatoires. Ils peuvent réduire la prévalence des allergies chez les sujets sensibles ou encore inhiber des réponses inflammatoires dans l'intestin (Tsai *et al.*, 2012). Des études ont suggéré que les lactobacilles pouvaient être utilisés pour stimuler le système immunitaire en première ligne de défense contre les pathogènes envahissants (Goldin, 1998). Ces espèces sont capables d'améliorer l'immunité chez la souris de façon dose- et souche-dépendantes (Galdeano & Perdigon, 2004; Paturi *et al.*, 2007). Les effets sont constatés à différents niveaux :

Tout d'abord sur l'immunité innée, qui est la première ligne de défense mais aussi qui joue un rôle crucial dans le développement des réponses immunitaires adaptatives ultérieures. Dans les modèles murins, plusieurs lactobacilles améliorent les réponses innée et adaptative par l'induction de la maturation des cellules dendritiques, mais aussi par la stimulation des cellules immunitaires sécrétant des molécules pro-inflammatoires (Christensen *et al.*, 2002; Cross *et al.*, 2002).

Ensuite, des effets bénéfiques ont également été rapportés sur l'immunité muqueuse intestinale. Les lactobacilles peuvent stimuler la sécrétion des immunoglobulines A dans la lumière de l'intestin par les cellules immunitaires (Vizoso Pinto *et al.*, 2009) mais aussi moduler l'immunité muqueuse déterminée par les tissus lymphoïdes secondaires chez la souris (Vinderola *et al.*, 2006).

D'autres études ont montré que plusieurs types de bactéries lactiques avaient des effets antiallergiques chez la souris. Il a été constaté une inhibition de la synthèse d'immunoglobulines E antigène-spécifiques, capables de déclencher des réactions de type allergique comme l'asthme, certaines rhinites ou l'urticaire chronique (Nonaka *et al.*, 2008; Enomoto *et al.*, 2009).

La production en grande quantité d'acétate par des bifidobactéries permet d'induire des effets anti-inflammatoires et/ou anti-apoptotiques au niveau des cellules épithéliales du côlon, contribuant ainsi à protéger l'hôte d'infections létales causées par une souche d'*E. coli* (Fukuda *et al.*, 2011).

3. Sur la balance microbienne

L'équilibre de la flore intestinale résulte d'interactions microbiennes sous la forme de compétitions ou de modifications de l'écosystème.

La prise de probiotiques peut mener à la création de conditions microbiologiques intestinales permettant de supprimer les microorganismes néfastes (Mountzouris *et al.*, 2007) et de favoriser les bactéries bénéfiques. L'effet de barrière des bactéries intestinales a une fonction de protection contre les pathogènes. Ainsi, les mécanismes potentiels de protection sont une adhésion compétitive à la muqueuse *via* le mucus (Candela *et al.*, 2008; Van Tassell & Miller, 2011; Turpin *et al.*, 2012), la prévention de lésions de la barrière cellulaire intestinale (Johnson-Henry *et al.*, 2008; Hummel *et al.*, 2012) et la production de substances antimicrobiennes létales pour les pathogènes comme les bactériocines (Dobson *et al.*, 2012).

4. Sur la cholestérolémie

La réduction du taux de cholestérol par les bactéries lactiques a été démontrée chez l'homme, la souris et le porc (Kawase *et al.*, 2000; Haberer *et al.*, 2003). Les acides biliaires (également connus sous le nom de sels biliaires) sont formés par des dérivés du cholestérol et par des stéroïdes acides secrétés par le foie. Les effets d'abaissement peuvent être dus en partie à la déconjugaison des sels biliaires par des souches de bactéries produisant une hydrolase adaptée. Une fois déconjugués, ils sont plus facilement excrétés dans les fèces comparativement aux sels conjugués, déterminant ainsi une réduction du taux de cholestérol libre. En outre, elles peuvent également diminuer la solubilité du cholestérol et ainsi réduire son absorption (Nguyen *et al.*, 2007). Plus simplement, il a également été proposé que les bactéries assimilent directement le cholestérol. Une assimilation significative de ce composé en présence de bile a été démontrée *in vitro* pour une souche de

L. acidophilus (Gilliland *et al.*, 1985). Cependant, ces résultats restent controversés et n'ont pas été à ce jour confirmés.

5. Sur la tension artérielle

Certaines souches de lactobacilles peuvent produire des peptides bioactifs dont certains inhibent l'activité de l'enzyme de conversion de l'angiotensine, responsable de l'hypertension artérielle par une action vasoconstrictrice. Ces peptides, possédant un motif proline-proline en C-terminal résistant aux peptidases, sont suffisamment petits pour être stables et absorbés (Hata *et al.*, 1996). D'autres mécanismes semblent être impliqués comme la stimulation de l'activité bradykinine (Perpetuo *et al.*, 2003) ou encore le mécanisme de détente vasculaire (Hernandez-Ledesma *et al.*, 2007).

6. Sur la carcinogenèse

Des études sur des modèles animaux ont montré des effets probiotiques provenant de bactéries lactiques sur la prévention du cancer du côlon avec une action inhibitrice sur le développement de lésions précancéreuses et des tumeurs (Pool-Zobel *et al.*, 1996; Brady *et al.*, 2000). De plus, une récente étude chez le rat a montré que les probiotiques pouvaient prévenir les cancers colorectaux causés par les maladies inflammatoires chroniques intestinales. En effet, ils moduleraient les réponses immunes au niveau de la muqueuse en atténuant l'inflammation chronique par des activités anti-inflammatoires et anti-carcinogéniques (Bassaganya-Riera *et al.*, 2012).

7. Sur l'humeur

De plus en plus de preuves suggèrent une interaction entre la flore intestinale, l'intestin et le système nerveux central, l'ensemble étant reconnu comme un axe microbiome-intestin-cerveau. Des études sur les rongeurs ont mis en évidence une dérégulation de cet axe dans les désordres fonctionnels de l'intestin, incluant le syndrome de l'intestin irritable (Bercik *et al.*, 2010; Cryan & O'Mahony, 2011).

Il a été démontré que l'absence et/ou la modification de la flore intestinale chez la souris affectaient la réponse de l'axe hypothalamique-pituitaire-surrénal face au stress (Sudo *et al.*, 2004; Neufeld *et al.*, 2011) et aux comportements d'anxiété (Lyte *et al.*, 2006; Diaz Heijtz *et al.*, 2011). Ceci est important étant donné la haute co-morbidité entre les désordres fonctionnels intestinaux et les désordres psychiatriques liés au stress comme l'anxiété et la dépression (Gros *et al.*, 2009). Des bactéries pathogènes chez les rongeurs peuvent induire

des comportements proches de l'anxiété médiés par des nerfs afférents vagaux (Goehler *et al.*, 2005; Lyte *et al.*, 2006). A l'inverse, des études sur des souches de probiotiques démontrent une régulation positive du comportement émotionnel *via* le nerf vague (Bravo *et al.*, 2011; Perez-Burgos *et al.*, 2012).

Plusieurs études cliniques suggèrent que les probiotiques pourraient moduler la réponse au stress et améliorer l'humeur et les symptômes d'anxiété chez les patients atteints de fatigue chronique et du syndrome de l'intestin irritable (Rao *et al.*, 2009).

C. Les microorganismes couramment utilisés

La plupart des probiotiques sont des bactéries similaires à celles retrouvées naturellement dans le tractus intestinal de l'homme et, plus particulièrement, à celles retrouvées chez les enfants nourris au sein, mode alimentaire connu pour déterminer une protection naturelle contre plusieurs maladies (Iannitti & Palmieri, 2010). Généralement, les microorganismes proviennent de deux groupes : les lactobacilles et les bifidobactéries. Chaque groupe implique différentes espèces comprenant plusieurs souches comme *Lactobacillus acidophilus*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus casei*, *Lactobacillus sporogenes*, *Bifidobacterium bifidum*, *Bifidobacteria longum*, *Bifidobacteria infantis*, *Streptococcus thermophilus*.

III. Les prébiotiques

Depuis que la notion de probiotique est apparue, un concept associé a émergé : les prébiotiques (Gibson & Roberfroid, 1995). Il s'agit d'une solution complémentaire séduisante permettant d'éviter les difficultés rencontrées lors de l'utilisation des probiotiques. En effet, ces derniers doivent être capables de résister aux conditions exigentes lors des processus industriels, mais également survivre au passage dans l'estomac avant d'atteindre leurs destinations finales.

A. Définition

Ils ont été initialement définis comme « des ingrédients alimentaires non digestibles qui affectent de façon bénéfique l'hôte en stimulant sélectivement la croissance et/ou l'activité d'une ou d'un nombre limité de bactéries dans le côlon et ainsi améliorer la santé de l'hôte » (Gibson & Roberfroid, 1995).

Depuis cette première définition, le concept de prébiotique attire de plus en plus l'attention de la communauté scientifique (Figure 4) mais aussi industrielle. Cependant, de

nombreux composés, principalement des oligo- et polysaccharidiques, ont été revendiqués en tant que tel sans pour autant répondre aux critères (Roberfroid, 2007). En effet, ces molécules doivent résister à l'acidité gastrique, aux enzymes mammifères et à l'absorption gastro-intestinale. Ils doivent également être fermentés par le microbiote intestinal et stimuler sélectivement la croissance et/ou l'activité de bactéries qui contribuent à la santé et au bien-être (Gibson *et al.*, 2004). Bien que le premier critère soit la résistance, ceci n'implique pas nécessairement que le prébiotique doive être complètement non digestible : il faut qu'il présente la garantie d'une quantité significative disponible dans l'intestin pour servir de substrat de fermentation. Le troisième critère de stimulation sélective est le plus difficile à remplir (Roberfroid, 2007).

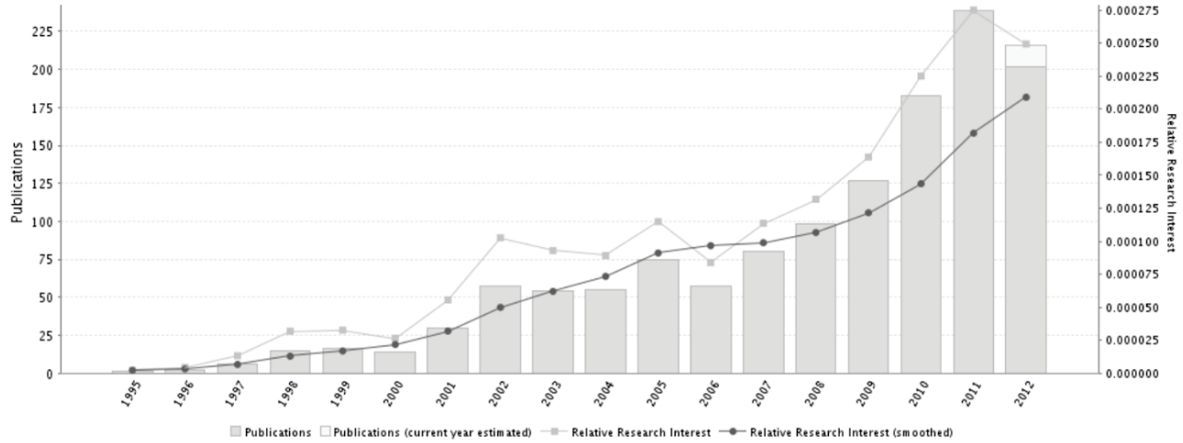


Figure 4 : Nombre d'articles publiés par an par la communauté scientifique sur les prébiotiques (généré par GoPubMed®).

B. Les prébiotiques candidats

L'inuline (Nazzaro *et al.*, 2012), les fructooligosaccharides (FOS) (Gustaw *et al.*, 2011; Barnes *et al.*, 2012), les galactooligosaccharides (GOS) (Macfarlane *et al.*, 2008), le lactulose et le polydextrose sont reconnus comme des prébiotiques établis (Stowell, 2007). Le mannitol et les maltodextrines ont démontré un effet prébiotique potentiel (Yeo & Liong, 2010). Les glucooligosaccharides comme le nigerose ou le kojibiose (Kang *et al.*, 2009), les isomaltooligosaccharides (IMO) comme le panose (Makelainen *et al.*, 2009; Andersen *et al.*, 2012), les xylooligosaccharides (XOS) (Kondepudi *et al.*, 2012) et le lactitol sont pressentis comme de possibles prébiotiques émergents mais ne peuvent pas être encore classés en tant que tels (Jie *et al.*, 2000; Gibson *et al.*, 2004; Patel & Goyal, 2012).

Actuellement, la technologie de production d'oligosaccharides est limitée à l'extraction de sources végétales, à l'hydrolyse acide ou enzymatique de polysaccharides ou la synthèse par réaction de transglycosylation (Patel & Goyal, 2012).

Chapitre 2 : Les enzymes actives sur les sucres

Les liaisons glycosidiques, en particulier entre deux glucoses, sont les liaisons covalentes les plus stables chimiquement de tous les biopolymères naturels. Les glycoside hydrolases clivent ces liaisons 10^{17} fois plus rapidement qu'une hydrolyse spontanée, ce qui en font les catalyseurs connus les plus efficaces (Wolfenden, 1999).

Les sucres et les glycoconjugués, retrouvés en quantité immense dans la nature, nécessitent une très grande variété de glycoside hydrolases (GH) et de glycosyltransférases (GT) en vue de leur hydrolyse ou de leur synthèse. Les glycoside hydrolases, au sens large du terme, combinent les *O*-, *N*- et *S*-hydrolases. Cependant, les *N*-hydrolases (EC 3.2.2.x) sont principalement des enzymes du métabolisme des acides nucléiques et la seule *S*-glycoside hydrolase (EC 3.2.3.1) connue a été réévaluée en 2001 en *O*-glycoside hydrolase selon une nouvelle classification (EC 3.2.1.147, IUBMB Enzyme Nomenclature, www.chem.qmul.ac.uk/iubmb/enzyme.html, 2012).

Les *O*-glycoside hydrolases sont classées sous la nomenclature EC 3.2.1.x, avec x représentant la spécificité de substrat. La nomenclature biochimique « traditionnelle » des glycosidases est basée exclusivement sur les spécificités de substrat/produit et tient rarement compte de leur mécanisme moléculaire de catalyse. Le principal avantage de cette classification est qu'elle est relativement simple. Jusqu'à maintenant, des activités enzymatiques de EC 3.2.1.1 à EC 3.2.1.175 ont été recensées (http://www.enzyme-database.org).

Cette classification, basée sur la spécificité de substrat ou de produit, pose cependant un problème intrinsèque. En effet, elle n'est pas appropriée pour décrire les enzymes qui agissent sur différents substrats comme les GH qui catalysent l'hydrolyse de polysaccharides distincts. Souvent, les glycoside hydrolases présentent une activité transglycosidase additionnelle (2.4.x.x).

De plus, les caractères structuraux spécifiques des GH ne sont pas considérés dans cette classification enzymatique. En effet, celle-ci ne reflète pas les caractéristiques structurales tridimensionnelles des enzymes apparentées et encore moins des ressemblances structurales, évolutionnaires et mécanistiques qui sont pourtant explorées depuis plusieurs années (Davies & Henrissat, 1995).

I. La base de données « CAZy » (Carbohydrate Active enZyme)

Les enzymes actives sur les sucres nécessitent donc une autre classification se basant conjointement sur les caractéristiques structurales et mécanistiques. L'accumulation de données sur les structures primaires des glycoside hydrolases (et des enzymes actives sur les sucres en général) a rendu possible l'élaboration d'un nouveau système de classification basé sur la comparaison de séquences en acides aminés de leurs domaines catalytiques. Ces enzymes sont ainsi répertoriées dans la base de données CAZy (www.cazy.org - Cantarel *et al.*, 2009) qui recense et décrit les modules catalytiques et fonctionnels des enzymes dégradant, modifiant ou créant des liaisons glycosidiques. Selon cette classification, des enzymes avec une spécificité de substrat différente sont parfois retrouvées dans une même famille, ceci tendant à prouver une évolution divergente pour acquérir de nouvelles spécificités (Aleshin *et al.*, 2003; Coutinho *et al.*, 2003; Giacopuzzi *et al.*, 2012). D'un autre coté, des enzymes hydrolysant le même substrat peuvent être retrouvées dans différentes familles de GH.

Il est ainsi possible d'avoir accès à des informations sur les glycoside hydrolases (GH), les glycosyltransférases (GT), les polysaccharide lyases (PL) et les carbohydrate estérases (CE). Nous ne développerons ici que les glycoside hydrolases. Sont aussi retrouvés les modules associés (appelés « carbohydrate-binding modules » ou CBM), également répertoriés en familles et rattachés aux modules catalytiques. Du fait du dynamisme des recherches entourant les CAZYmes (« Carbohydrates-Active Enzymes »), le classement est régulièrement actualisé. De plus, cette base de données propose des liens vers des banques externes de séquences protéiques et nucléiques et de structures tridimensionnelles.

II. Les glycoside hydrolases

Les glycoside hydrolases sont des enzymes qui catalysent l'hydrolyse sélective de liaisons inter-glycosidiques. Elles interviennent dans une multitude de réactions chez les êtres vivants. La détermination de la structure tridimensionnelle de ces enzymes en complexe avec leur substrat, analogues de substrat, inhibiteurs et/ou produits de la réaction catalytique, représente un enjeu majeur pour la compréhension de leurs modes de fonctionnement. A ce jour, il existe 127 familles de GH (annotées de GH1 à GH132 avec 5 familles supprimées), dont 103 possèdent au moins un représentant structural. Les familles GH21, 40, 41, 60 et 69 ont été retirées car des recherches plus poussées ont montré que les enzymes de ces familles étaient dépourvues de toute activité glycosidase.

Le critère pour rattacher une protéine à une famille de glycoside hydrolase est la présence dans sa séquence d'une région étendue (minimum 100 acides aminés) homologue à une séquence d'au moins un membre connu de cette famille (Henrissat, 1991). Étant donné qu'il existe une relation directe entre les similarités de séquence et de repliement structural, cette classification reflète plus les caractéristiques structurales que la spécificité de substrat propre. De plus, elle permet de révéler des relations évolutives entre ces enzymes au sein d'une famille alors même qu'elles proviennent d'organismes éloignés et fournit également un outil de poids pour obtenir des informations mécanistiques (Henrissat & Bairoch, 1993).

Il est admis que le repliement des protéines est beaucoup mieux conservé que leur séquence. Fort de ce fait, certaines familles sont réunies sous forme de « clans » lorsque de nouvelles séquences peuvent être apparentées à plusieurs familles ou encore lorsque des déterminations structurales mettent en avant des ressemblances entre des membres de familles différentes (Henrissat & Bairoch, 1996). A ce jour, quatorze superfamilles ont ainsi été répertoriées (de GH-A à GH-N). Il est supposé que toutes les glycosidases appartenant aux familles du même clan ont une origine commune d'évolution, une structure 3D similaire, un arrangement des résidus catalytiques conservé et le même mécanisme d'hydrolyse de la liaison glycosidique (Gebler *et al.*, 1992). Une exception est à noter pour la famille GH97 dans laquelle sont retrouvés les deux mécanismes d'action les plus classiques (Gloster *et al.*, 2008).

III. Les mécanismes d'hydrolyse

Deux mécanismes d'actions sont retrouvés le plus couramment chez les GH et ont été décrits pour la première fois par Koshland (Koshland, 1953). Dans les deux cas, le site actif contient une paire de groupements carboxyliques qui jouent un rôle fondamental dans la réaction. Toutefois, des variations de ces mécanismes et même un autre fondamentalement différent ont été découverts depuis lors.

A. Inversion de la configuration du carbone anomérique

L'hydrolyse avec inversion de la configuration du carbone anomérique se fait en une seule étape en passant par un état de transition de type ion oxocarbénium (Tanaka *et al.*, 1994). La protonation de l'oxygène de la liaison glycosidique par le résidu catalytique « acide », l'attaque sur le carbone anomérique d'une molécule d'eau déprotonée par le résidu « base/nucléophile » et le départ du produit aglycone se produisent de manière concomitante et concertée. Il est nécessaire que la molécule d'eau vienne se placer entre la base et la liaison osidique à hydrolyser. Ainsi, la distance entre les deux résidus catalytiques est comprise entre 6,5 et 11 Å (McCarter & Withers, 1994) (Figure 5).

À terme, l'hydrolyse d'une liaison α-glycosidique libère un produit dont le carbone anomérique est en configuration β.

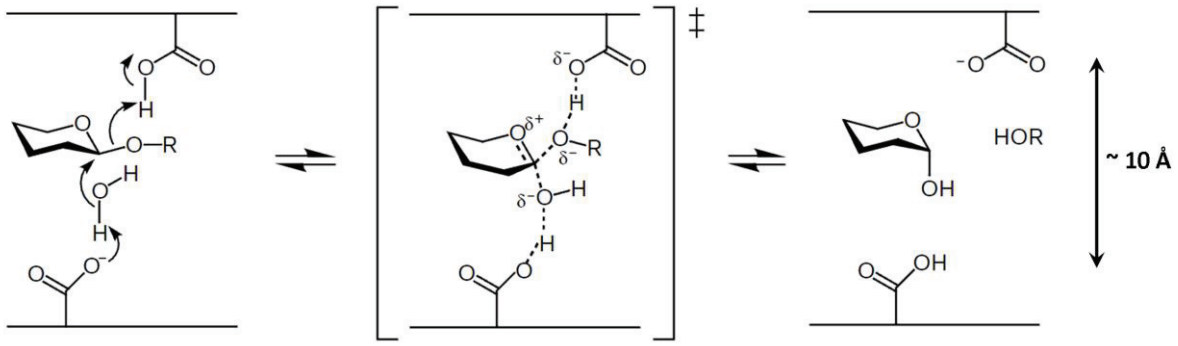


Figure 5 : Schéma de l'hydrolyse d'une liaison glycosidique (ici en configuration β) selon un mécanisme d'inversion de la configuration du carbone anomérique (adapté de Rye & Withers, 2000).

B. Rétention de la configuration du carbone anomérique

La caractéristique majeure de ce mécanisme est qu'il comporte deux étapes distinctes. La première consiste en la formation transitoire d'un intermédiaire glycosyl-enzyme covalent ; la seconde, en la rupture de cette liaison covalente, correspondant à la « déglycosylation » et la « régénération » de l'enzyme. A chaque étape, la configuration du

carbone anomérique est inversée, en passant par un état de transition de type ion oxocarbénium (Tanaka *et al.*, 1994), afin de conserver globalement la stéréochimie du substrat à la fin de la réaction.

Dans un premier temps, l'atome d'oxygène de la liaison osidique subissant l'hydrolyse est protoné par le résidu catalytique « acide/base » agissant comme un acide. Le second résidu catalytique agit simultanément comme nucléophile/base sur le carbone anomérique, provoquant la formation d'une liaison covalente amenant à une glycosyl-enzyme. Cette étape est souvent appelée « étape de glycosylation » avec départ du produit du coté aglycone de la liaison osidique.

Dans un second temps, une molécule d'eau joue le rôle de nucléophile après avoir été déprotonée par le résidu catalytique « acide/base » devenu base lors de la première étape. Il y a alors formation d'un second état de transition avec, à terme, la « déglycosylation » de l'enzyme conduisant à la libération du sucre (Figure 6).

Compte tenu du mécanisme décrit, les résidus catalytiques doivent être proches dans l'espace pour que le résidu catalytique « base/nucléophile » soit à proximité du carbone anomérique qui va être attaqué. Ainsi, on estime que la distance moyenne entre les fonctions carboxyliques impliquées est de l'ordre de 5,5 Å (McCarter & Withers, 1994).

Il est possible que la molécule d'eau qui joue le rôle de nucléophile puisse être remplacée par un sucre accepteur, on parle alors de transglycosylation. L'efficacité de cette réaction dépend toutefois de la compétition entre le sucre et la molécule d'eau (Wang & Huang, 2009).

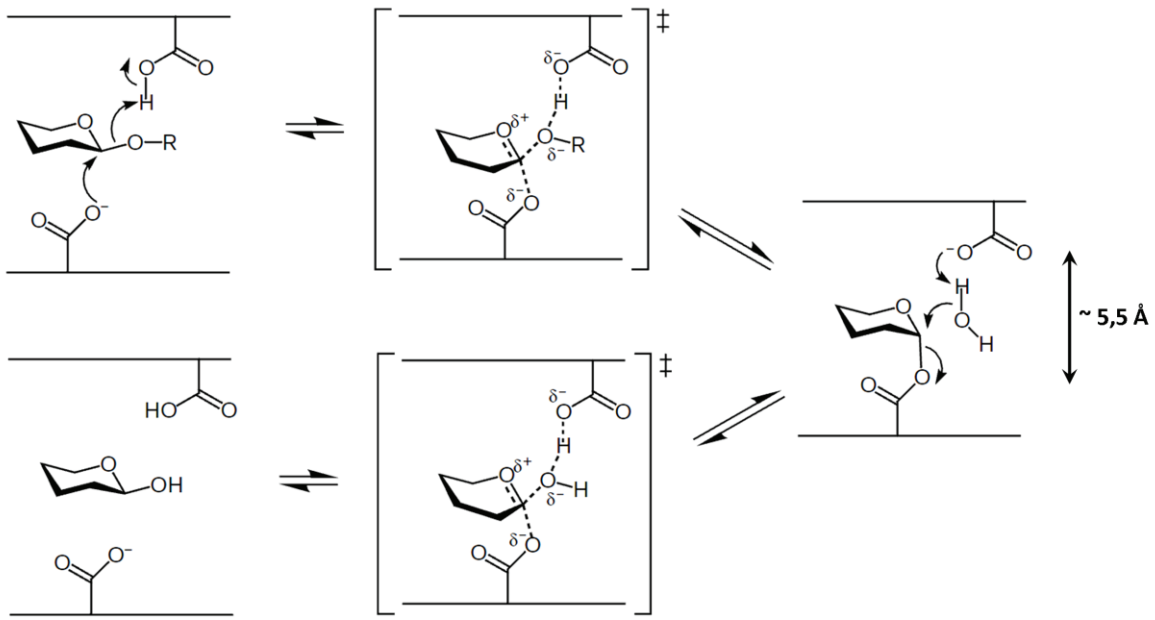


Figure 6 : Schéma de l'hydrolyse d'une liaison glycosidique (ici en configuration β) selon un mécanisme de rétention de la configuration du carbone anomérique (adapté de Rye & Withers, 2000).

C. Les variations de mécanisme

Depuis l'établissement de ces deux principaux mécanismes, il a été rapporté d'autres modes d'hydrolyse. Il est possible de retrouver dans certaines familles (GH18, 20, 25, 56, 84 et 85) la participation d'un groupement voisin du site actif où n'est présent que le résidu catalytique acide et non le résidu nucléophile. Le groupement nucléophile est apporté par le substrat lui-même et le groupement du résidu voisin joue un rôle de stabilisateur de charge lors de l'état de transition (Terwisscha van Scheltinga *et al.*, 1995; Mark *et al.*, 2001).
Dans le cas de la famille GH1, certaines enzymes voient leur acide glutamique « acide/base » remplacé par une glutamine. L'accomplissement de la réaction requiert alors un co-enzyme pour remplacer le groupement basique manquant et achever la deuxième étape d'hydrolyse (Burmeister *et al.*, 2000).

Chez les familles GH33 et 34, on note la présence dans le site actif d'un résidu catalytique nucléophile alternatif représenté par une tyrosine activée par un résidu adjacent « base » (Watson *et al.*, 2003).

D. Hydrolyse dépendante du NAD

Un autre mécanisme tout à fait différent, détaillé en 2004, met en jeu un cofacteur, le NAD, qui reste étroitement lié tout au long de la catalyse. Le mécanisme passe par des états de transition anioniques au lieu du « classique » ion oxocarbénium (Rajan *et al.*, 2004; Yip *et al.*, 2004). Ce mécanisme est retrouvé dans les familles GH4 et GH109.

IV. Les topologies du site actif

Indépendamment du mécanisme catalytique, il existe trois catégories de topologie de site actif (Davies & Henrissat, 1995) (Figure 7).

La topologie en poche (ou cratère) est optimale pour la reconnaissance de l'extrémité non-réductrice d'un saccharide et est retrouvée chez les monosaccharidases et les exo-polysaccharidases. Cette topologie n'est pas du tout adaptée aux substrats fibreux tels que la cellulose, par exemple.

La crevasse (ou sillon) est une structure ouverte qui permet la fixation de plusieurs unités de sucre des substrats polysaccharidiques. Cette topologie est retrouvée chez les endo-polysaccharidases, comme les α-amylases.

La dernière topologie dérive de la précédente, lorsque la protéine présente de longues boucles recouvrant une partie de la crevasse. Dans ce cas, le produit enzymatique peut être libéré bien que la chaine polysaccharidique reste fermement liée, ceci permettant à l'enzyme de réaliser plusieurs cycles d'hydrolyse consécutifs, se définissant sous le terme

de « processivité » (Breyer & Matthews, 2001). Ce processus est probablement un facteur clé pour l'efficacité enzymatique sur des substrats insolubles comme la cellulose microcristalline.

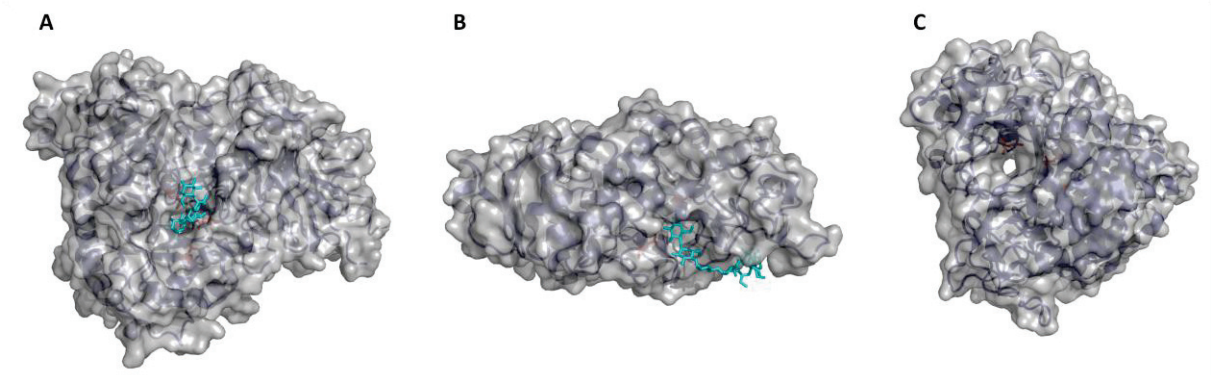


Figure 7 : Différentes topologies du site actif retrouvées chez les glycoside hydrolases. Les résidus catalytiques sont colorés en rouge, les ligands dans le site actif en cyan. A : Site actif en « poche » de l'α-glucosidase de *B. thetaiotaomicron* (Kitamura *et al.*, 2008). B : Site actif en « crevasse » de l'α-amylase de *H. vulgare* (Robert *et al.*, 2003). C : Site actif en « tunnel » de la cellulase de *C. cellulolyticum* (Parsiegla *et al.*, 1998).

V. Le site actif et les sous-sites de fixation du substrat

Avant 1997, la nomenclature des sous-sites n'était pas cohérente pour toutes les glycoside hydrolases, ceci rendant la comparaison entre différentes enzymes difficile voire impossible. Il a alors été proposé une nouvelle nomenclature prenant en compte deux critères primordiaux : la position relative des sous-sites par rapport au point d'hydrolyse ainsi que la conservation de la numérotation des sous-sites du coté réducteur et non-réducteur lors de nouveaux complexes. La cartographie des sous-sites adoptée est donc une nomenclature de type *–n* à *+n* représentant respectivement l'extrémité non-réductrice et réductrice, la liaison hydrolysée se situant par définition entre les sous-sites -1 et +1 (Davies *et al.*, 1997) (Figure 8).

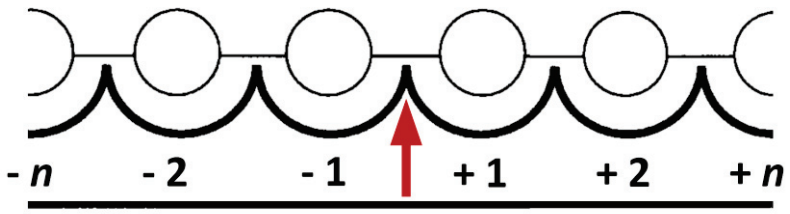


Figure 8 : Représentation schématique des sous-sites de fixation (adapté de Davies *et al.*, 1997). L'extrémité réductrice du substrat est positionnée du coté des sous-sites « positifs ». L'hydrolyse de la liaison glycosidique se produit entre les sous-sites -1 et +1.

VI. Les « Carbohydrate Binding Modules » - CBMs

Les glycoside hydrolases possèdent très souvent une structure modulaire avec, au minimum, un domaine catalytique adjoint à un ou plusieurs domaines auxiliaires et parfois des modules associés fixant les sucres.

La nature modulaire des CBMs a été mise en évidence la première fois en 1986 (Van Tilbeurgh *et al.*, 1986), confirmée en 1988 (Gilkes *et al.*, 1988), même année que la première caractérisation publiée (Tomme *et al.*, 1988). Pouvant être séparés des domaines qui possèdent une activité cellulase par protéolyse ménagée, ils ont été définis comme des polypeptides stables d'une vingtaine de résidus ayant pour fonction la fixation de la cellulose. Seul le terme de CBD (« Cellulose Binding Domain ») était utilisé jusqu'en 1999, date à laquelle des modules associés aux GHs fixant d'autres molécules que la cellulose ont été mis en évidence (Boraston *et al.*, 1999).

Le terme de « CBM », plus générique, a progressivement été utilisé depuis. Sa nomenclature, comparable à la classification des GHs, se base sur la similarité de séquence (Tomme *et al.*, 1995).

A ce jour, 66 familles de CBMs sont répertoriées dans CAZy (de CBM1 à CBM67, la famille CBM7 ayant été retirée) parmi lesquelles 51 ont au minimum un représentant structural. Cette classification contribue à l'identification de ces modules et permet d'en prédire les spécificités de fixation, d'identifier les résidus fonctionnels ainsi que révéler les relations évolutives et prédire des repliements tridimensionnels de ces polypeptides.

Ces modules, pouvant contenir jusqu'à environ 150 résidus, sont susceptibles d'adopter différents repliements. Sont ainsi retrouvés le sandwich β, le « β-trefoil » (en forme de trèfle), le repliement en « feuillet β et bobine uniquement » (typiquement retrouvé pour la reconnaissance de la cellulose et des oligosaccharides) et, enfin, le repliement « hevein-like », majoritairement en boucle désordonnée (Boraston *et al.*, 2004).

Étant donné que le repliement ne suffit pas à la prédiction de la fonction, un autre classement est proposé selon la modalité de fixation du substrat. Le type A correspond à une fixation en surface du substrat, ce dernier étant généralement cristallin (Creagh *et al.*, 1996). Le type B englobe les enzymes ayant un site de fixation étendu et sous forme de crevasse (Figure 9). Enfin, les sites de type C sont optimaux pour des substrats de petites tailles comme les tri-saccharides.

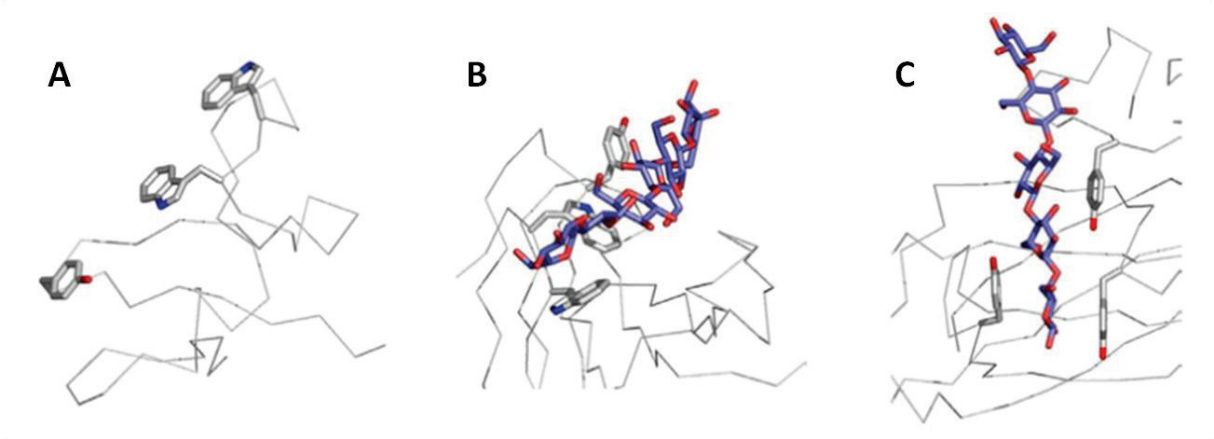


Figure 9 : Visualisation des « plateformes » formées par les résidus aromatiques au niveau du site de fixation de CBMs (issu de Boraston *et al.*, 2004). *A* : plateforme plane retrouvée chez le type A. B : plateforme tordue retrouvée chez certains CBMs de type B. *C* : plateforme sandwich retrouvée également chez le type B. Les résidus aromatiques du site de fixation sont représentés en bâtons gris et les oligosaccharides en bâtons violet.

Chapitre 3 : Les α-glucosidases

D'après la nomenclature IUPAC-IUBMB, les α-glucosidases (EC 3.2.1.20, α-D-glucoside glucohydrolase) sont définies comme des exo-glycosidases catalysant majoritairement l'hydrolyse des liaisons α-1,4-glucosidiques d'un substrat (comme chez les maltooligosaccharides ou le glycogène par exemple) pour libérer un α-D-glucose. Elles peuvent également catalyser des réactions de transglycosylation pour synthétiser des composés α-D-glucosylés (Kato *et al.*, 2002; Ferrer *et al.*, 2005; Seo *et al.*, 2011). L'activité de transglycosylation des α-glucosidases est utilisée dans l'industrie pour produire des oligosaccharides et des sucres conjugués dans le but d'améliorer leurs propriétés chimiques et leurs fonctions physiologiques (Kren & Martinkova, 2001; Murata & Usui, 2006).

Ces enzymes sont retrouvées dans six familles différentes de glycoside hydrolases répertoriées dans la base de données CAZy : GH4, 13, 31, 63, 97 et 122. Cette dernière, récemment créée, ne possède pas à ce jour de représentant structural (Comfort *et al.*, 2008).

I. Description des différentes familles

A. Famille GH4

1. Activité

Dans cette famille sont retrouvées les activités enzymatiques : maltose-6-phosphate glucosidase (EC 3.2.1.122), α-galactosidase (EC 3.2.1.22), 6-phospho-β-glucosidase (EC 3.2.1.86), α-glucuronidase (EC 3.2.1.139) et α-glucosidase (EC 3.2.1.20). Les enzymes de cette famille sont retrouvées très majoritairement chez les Eubactéries et aussi chez quelques Archées, catalysant leur réaction d'hydrolyse selon un mécanisme inhabituel impliquant du NAD^+ dont la dépendance a été mise en évidence dès 1971 (Burstein & Kepes, 1971). Cette famille a la particularité de regrouper des enzymes ayant des spécificités pour des liaisons α- ou β-glycosidiques (Pikis *et al.*, 2002).

2. Description structurale

La première structure tridimensionnelle décrivant une α-glucosidase (provenant de *Thermotoga maritima*) date de 2003 (Lodge *et al.*, 2003). Elle est accessible dans la base de données PDB sous le code 1OBB. Jusqu'en 2011, il s'agissait de la seule α-glucosidase de cette famille dont la structure cristalline avait été résolue. Depuis, la structure d'une α-

glucosidase putative issue de *Thermotoga neapolitana* a été déposée (code PDB : 3U95) (Yun *et al.*, 2011). Au total, cette famille possède sept représentants structuraux. Le site de fixation du NAD^+ est typique avec un repliement de type « Rossman » (Rao & Rossmann, 1973). Chez l'α-glucosidase de *Thermotoga maritima*, la structure résolue par cristallographie aux rayons X montre la présence du NAD^+ ainsi qu'une molécule de maltose dans le site actif avec la partie nicotinamide proche du sucre (Figure 10). Un ion Mn^{2+} décrit comme essentiel à l'activité est absent dans cette structure (Lodge *et al.*, 2003) mais retrouvé chez l'α-glucosidase de *Thermotoga neapolitana* (Yun *et al.*, 2011).

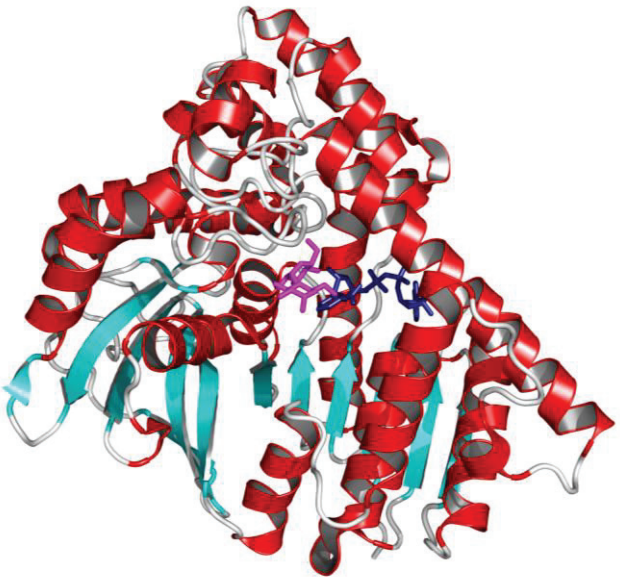


Figure 10 : Représentation en ruban de la structure tridimensionnelle résolue par cristallographie aux rayons X de l'α-glucosidase de *Thermotoga maritima* (Lodge *et al.*, 2003). Les hélices α sont représentées en rouge, les brins β en cyan. Le maltose et le NAD^+ sont représentés en bâtons colorés respectivement en magenta et bleu.

B. Famille GH13

1. Activité

Cette famille appartenant au clan GH-H est la famille majeure agissant sur les substrats possédant des liaisons α-glucosidiques. Initialement appelée famille des α-amylases, elle fut définie comme GH13 lors de la création de la classification basée sur les séquences en 1991 (Henrissat, 1991) et est présente chez les six règnes du Vivant. Différentes classes d'enzymes sont retrouvées dans cette famille comme des hydrolases (EC 3.x.x.x), des glycosyltransférases (EC 2.4.x.x) et des isomérases (EC 5.x.x.x) (MacGregor *et al.*, 2001). Il a même été rapporté que des transporteurs d'acides aminés d'origine animale, sans activité glycosidase, appartenaient également à cette famille (Fort *et al.*, 2007; Gabrisko & Janecek, 2009).

Ainsi sont retrouvées les activités : α-amylase (EC 3.2.1.1), oligo-1,6-glucosidase (EC 3.2.1.10), α-glucosidase (EC 3.2.1.20), pullulanase (EC 3.2.1.41), cyclomaltodextrinase (EC 3.2.1.54), glucane 1,4-α-maltotétraohydrolase (EC 3.2.1.60), isoamylase (EC 3.2.1.68), dextrane 1,6-α-glucosidase (EC 3.2.1.70), tréhalose-6-phosphate hydrolase (EC 3.2.1.93), glucane 1,4-α-maltohexaosidase (EC 3.2.1.98), glucane 1,4-α-maltotriohydrolase (EC 3.2.1.116), glucane 1,4-α-maltohydrolase (EC 3.2.1.133), néopullulanase (EC 3.2.1.135), malto-oligosyltréhalose tréhalohydrolase (EC 3.2.1.141), dextrinase limite (EC 3.2.1.142), glucane 1,4-α-maltopentaohydrolase (EC 3.2.1.-), amylosucrase (EC 2.4.1.4), sucrose phosphorylase (EC 2.4.1.7), enzyme 1,4-α-glucane branchante (EC 2.4.1.18), cyclomaltodextrine glucanotransférase (CGTase) (EC 2.4.1.19); 4-α-glucanotransférase (EC 2.4.1.25), isomaltulose synthase (EC 5.4.99.11) et tréhalose synthase (EC 5.4.99.16).

Étant donné cette polyspécificité, la famille GH13 a été subdivisée en 37 sous-familles (Stam *et al.*, 2006). De plus, il a été rapporté qu'un nombre important d'enzymes de cette famille possédaient des CBMs (« Carbohydrate Binding Modules ») fixant l'amidon comme les CBM20, CBM21, CBM25, CBM26, CBM34, CBM41, CBM45, CBM48, CBM53 et CBM58 (Svensson *et al.*, 1989; Janecek & Sevcik, 1999; Rodriguez-Sanoja *et al.*, 2005; Machovic & Janecek, 2006; Christiansen *et al.*, 2009).

Il a été démontré que les enzymes de cette famille agissaient selon un mécanisme de rétention de la configuration du carbone anomérique (Isoda *et al.*, 1992). Le site actif est constitué d'une triade catalytique composée de deux acides aspartiques : un ayant le rôle de nucléophile, le second participant à la stabilisation de l'état de transition (Uitdehaag *et al.*, 1999) et un acide glutamique défini comme le résidu acide/base (Sinnott, 1998; Rydberg *et al.*, 2002).

2. Description structurale

A ce jour, 82 structures tridimensionnelles d'enzymes appartenant à cette famille ont été déposées dans la PDB ; la première, une α-amylase, a été résolue en 1984 (code PDB : 2TAA) (Matsuura *et al.*, 1984). Une seule structure d'α-glucosidase de cette famille a été résolue jusqu'à présent, issue de l'organisme *Geobacillus* sp. HTA-462 (code PDB : 2ZE0) (Shirai *et al.*, 2008).

Les enzymes de la famille GH13 possèdent trois domaines caractéristiques : une partie N-terminale avec un repliement en tonneau $(\beta/\alpha)_8$ de type « TIM barrel » (Banner *et al.*, 1975) constituant le domaine catalytique (appelé domaine A ou N), un sous-domaine (appelé domaine B) inséré entre le brin β_3 et l'hélice α_3 du tonneau catalytique (Janecek *et al.*, 1997) jouant un rôle dans la spécificité de substrat (MacGregor *et al.*, 2001) et une partie

carboxy-terminale, formant un sandwich β composé de brins β antiparallèles, appelée domaine C (Figure 11).

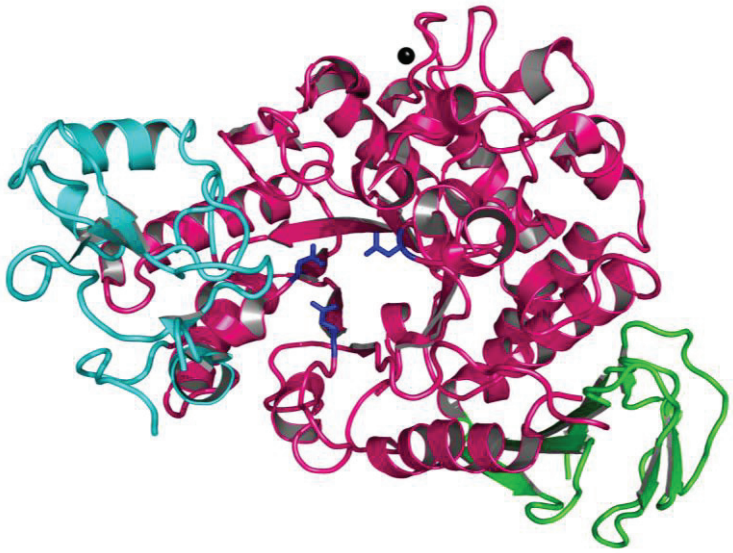


Figure 11 : Représentation en ruban de la structure tridimensionnelle résolue par cristallographie aux rayons X de l'α-glucosidase de *Geobacillus* sp. HTA-462 (Shirai *et al.*, 2008). Le domaine catalytique A est coloré en magenta, le domaine B en cyan et le domaine C en vert. Les résidus catalytiques sont représentés en bâtons et colorés en bleu ; l'ion calcium est représenté sous forme de sphère noire.

C. Famille GH31

1. Activité

Cette famille appartient au clan GH-D dans laquelle sont retrouvées les activités α-glucosidase (EC 3.2.1.20), α-1,3-glucosidase (EC 3.2.1.84), sucrase-isomaltase (EC 3.2.1.48, EC 3.2.1.10), α-xylosidase (EC 3.2.1.177), α-glucane lyase (EC 4.2.2.13) et isomaltosyltransférase (EC 2.4.1.-). Les enzymes de cette famille sont présentes chez les Archées, les Bactéries et les Eucaryotes.

Les enzymes de la famille GH31 agissent selon un mécanisme de rétention de la configuration du carbone anomérique et semblent suivre le mécanisme de Koshland classique à deux étapes (Frandsen & Svensson, 1998). Cette idée est confortée par des études sur l'inactivation d'α-glucosidases par du conduritol B epoxide (Hermans *et al.*, 1991; Iwanami *et al.*, 1995). Des études ultérieures ont permis de mettre en évidence des intermédiaires catalytiques en utilisant des réactifs de type 5-fluoroglycosyl fluoride (Lee *et al.*, 2001; Lovering *et al.*, 2005). Les résidus catalytiques acide/base et nucléophile sont toujours des acides aspartiques, plusieurs techniques le prouvant : par le conduritol B epoxide (Hermans *et al.*, 1991; Iwanami *et al.*, 1995), du 5-fluoro-α-D-glucopyransyl fluoride couplé à un séquençage par LC/MS/MS (Lee *et al.*, 2001), des études cristallographiques (Lovering *et al.*, 2005) et de mutagenèse dirigée (Kim *et al.*, 2006).

2. *Description structurale*

A ce jour, huit représentants structuraux ont été déposés dans la base de données PDB, dont trois correspondent à des α-glucosidases, chacune issue d'organismes appartenant à un règne différent : *Sulfolobus solfataricus* (code PDB : 2G3M) (Ernst *et al.*, 2006) chez les Archaea, *Ruminococcus obeum* (code PDB : 3NO4) (Tan *et al.*, 2010) chez les bactéries et *Homo sapiens* (code PDB : 2QLY) (Sim *et al.*, 2008) chez les eucaryotes. La structure résolue de l'α-glucosidase humaine correspond à la sous-unité catalytique N-terminale. La première structure tridimensionnelle de la famille ayant été résolue est une α-xylosidase (code PDB : 1XSI) (Lovering *et al.*, 2005).

Les enzymes de cette famille sont subdivisées en 4 sous-groupes dont chacun contient un motif de séquence caractéristique situé dans la région encadrant le résidu catalytique nucléophile. Ces signatures sont corrélées à la fonction et/ou la spécificité des enzymes et les variations au sein du motif consensus reflètent la distance évolutionnaire entre ces 4 sous-familles. Ainsi, toutes les α-glucosidases de la GH31 connues sont retrouvées dans le sous-groupe 1 (Ernst *et al.*, 2006).

Ces enzymes sont composées de 4 domaines : un domaine N-terminal en sandwich β (domaine N) suivi du domaine catalytique formé d'un tonneau $(\beta/\alpha)_8$ de type « TIM barrel » (domaine A) à la suite duquel se trouvent deux domaines en sandwich β (domaine C ou proximal et domaine D ou distal). Au sein du domaine catalytique sont retrouvées deux insertions : entre le brin β_3 et l'hélice α_3 (sous-domaine B) mais aussi entre brin β_4 et l'hélice α_4 (sous-domaine B') (Figure 12). Les trois domaines en sandwich β sont formés de brins β antiparallèles.

Le domaine N est composé d'une dizaine de brins β incluant de grandes boucles interagissant avec le domaine A et qui semblent particulièrement importantes. Ce domaine N-terminal parait avoir un rôle dans la fixation du substrat et/ou dans le maintien de l'architecture du domaine catalytique (Ernst *et al.*, 2006). Généralement, les feuillets β sont connus pour jouer un rôle dans la fixation des sucres (Boraston *et al.*, 2004) et aider les polysaccharides à se positionner avant d'approcher le domaine catalytique (Tan *et al.*, 2010).

Le domaine C, constitué de 8 brins β, semble jouer un rôle similaire à son équivalent chez la famille GH13 (MacGregor *et al.*, 2001), à savoir stabiliser le domaine A en le protégeant du solvant (Ernst *et al.*, 2006). Le domaine D, quant à lui formé de 11 brins β et moins conservé, parait avoir évolué en acquérant des fonctions distinctes. Mais des études plus approfondies sont nécessaires afin de confirmer ou infirmer cette hypothèse (Ernst *et al.*, 2006).

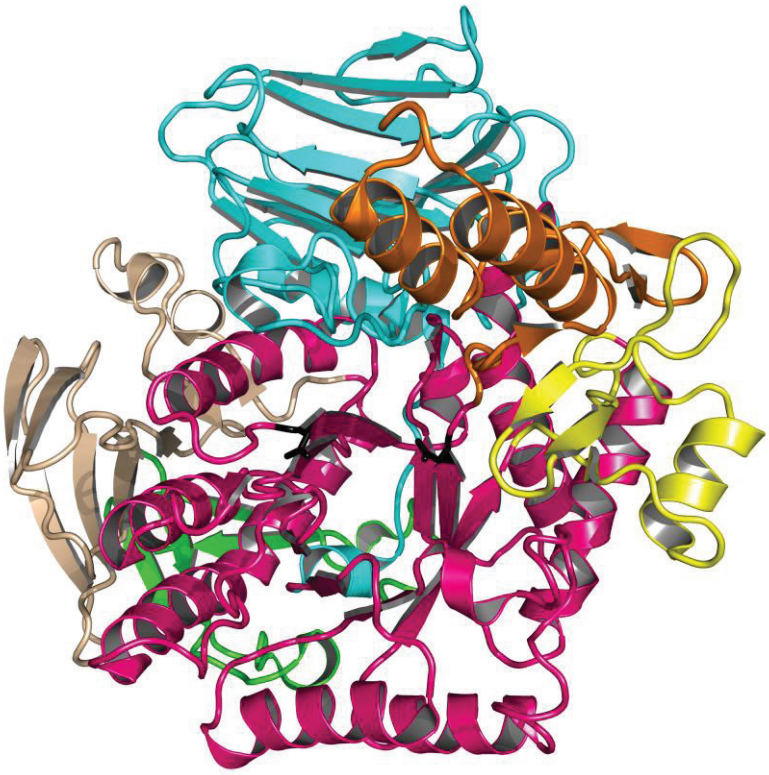


Figure 12 : Représentation en ruban de la structure tridimensionnelle résolue par cristallographie aux rayons X de l'α-glucosidase de *Ruminococcus obeum* (Tan *et al.*, 2010). Le domaine N, le domaine catalytique A et les domaines C et D sont respectivement colorés en cyan, rose, beige et vert. Les sous-domaines B et B' sont colorés en jaune et orange. Les résidus catalytiques supposés sont représentés en bâtons et colorés en noir.

D. Famille GH63

1. Activité

Les enzymes de cette famille ont des activités α-glucosidase processive (EC 3.2.1.106), α-1,3-glucosidase (EC 3.2.1.84) ou encore α-glucosidase (EC 3.2.1.20) et appartiennent au clan GH-G. Celles issues d'organismes eucaryotes hydrolysent spécifiquement les liaisons α-1,2-glucosidique du $Glc_3Man_9GlcNAc_2$ (Herscovics, 1999; Dhanawansa *et al.*, 2002), un précurseur de N-glycoprotéines (Dairaku & Spiro, 1997). Des gènes codant des enzymes de la famille GH63 ont également été retrouvés chez les Archaea et les bactéries avec des activités α-1,3-glucosidase ou α-glucosidase. Une de ces enzymes bactériennes issue d'*Escherichia coli*, YgjK, a une forte activité hydrolytique sur le nigerose (Glc-α-1,3-Glc) (Kurakata *et al.*, 2008).

Elles agissent selon un mécanisme d'inversion de la configuration du carbone anomérique, ce qui a été montré par RMN avec l'α-glucosidase processive de *Saccharomyces cerevisiae* (Palcic *et al.*, 1999).

Les résidus catalytiques ne sont pas définis expérimentalement mais ont été déduits par comparaison des domaines catalytiques des enzymes de la famille GH15 et GH37. Ainsi le résidu catalytique acide est prédit comme étant un acide aspartique tandis que le résidu catalytique base serait un acide glutamique (Kurakata *et al.*, 2008).

2. Description structurale

Seuls deux représentants structuraux ont été répertoriés jusqu'à aujourd'hui. L'un des deux est une α-glucosidase issue d'*Escherichia coli*, YgjK (code PDB : 3D3I) (Kurakata *et al.*, 2008).

La structure tridimensionnelle résolue par cristallographie aux rayons X révèle la présence de deux domaines principaux connectés par deux hélices α. Le domaine N-terminal, appelé domaine N, est constitué de 18 brins β antiparallèles divisés en deux feuillets β. Il semble avoir des homologies structurales avec le domaine N des GH31 (Tan *et al.*, 2010). Le second domaine, appelé domaine A, est composé de 16 hélices α organisées en deux couches concentriques constituant un tonneau $(\alpha/\alpha)_6$ et déterminant le domaine catalytique. Les deux hélices non modulaires imposent une courbure importante entre les deux domaines. YgjK possède un sous-domaine au sein du domaine A (appelé sous-domaine A') composé d'hélices α et de brins β (Figure 13).

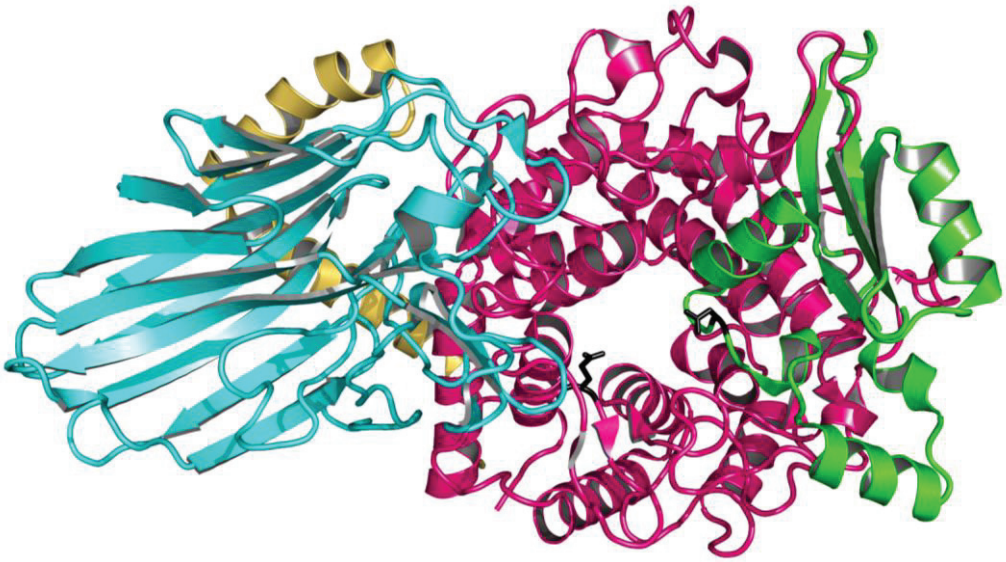


Figure 13 : Représentation en ruban de la structure tridimensionnelle résolue par cristallographie aux rayons X de l'α-glucosidase d'*Escherichia coli*, YgjK (Kurakata *et al.*, 2008). Le domaine N et le domaine catalytique A sont colorés respectivement en cyan et rose. Les deux hélices α reliant ces deux domaines et le sous domaine A' sont colorés respectivement en or et vert. Les résidus catalytiques supposés sont représentés en bâtons et colorés en noir.

E. Famille GH97

1. Activité

Les enzymes de cette famille sont toutes issues d'organismes archées ou bactériens. Jusqu'à présent, il a été rapporté qu'elles ont une activité α-glucosidase (3.2.1.20) ou α-galactosidase (EC 3.2.1.22) (Gloster *et al.*, 2008; Kitamura *et al.*, 2008).

La particularité de cette famille est de comprendre les deux mécanismes d'hydrolyse principaux. En effet, l'α-glucosidase étudiée agit selon un mécanisme d'inversion de la configuration du carbone anomérique, comme l'ont démontré les analyses par RMN et HPLC (Gloster *et al.*, 2008; Kitamura *et al.*, 2008). A l'inverse, l'α-galactosidase agit selon un mécanisme de rétention de la configuration du carbone anomérique également confirmé par RMN et HPLC (Gloster *et al.*, 2008; Kitamura *et al.*, 2008; Okuyama *et al.*, 2009).

Étant donné qu'une multitude de glutamates est retrouvée au niveau du site actif, il n'est pas aisé de définir les résidus catalytiques. Toutefois, des alignements structuraux avec la famille GH27 et des alignements de séquences tendent à prouver que le résidu catalytique donneur de proton serait dans tous les cas un acide glutamique. Le résidu nucléophile, quant à lui, serait un acide glutamique dans le cas d'un mécanisme par inversion, ou un acide aspartique pour un mécanisme par rétention (Gloster *et al.*, 2008; Kitamura *et al.*, 2008; Okuyama *et al.*, 2009).

2. Description structurale

Seuls deux représentants structuraux ont été déposés dans la base données PDB pour cette famille. Deux équipes ont résolu quasi simultanément la structure tridimensionnelle de l'α-glucosidase issue de *Bacteroides thetaiotaomicron* (SusB) par cristallographie aux rayons X (code PDB : 2JKA (Gloster *et al.*, 2008) ; code PDB : 2D73 (Kitamura *et al.*, 2008)). Le deuxième représentant structural est une α-galactosidase également issue de *B. thetaiotaomicron* (code PDB : 3A24) (Okuyama *et al.*, 2009).

Ainsi, les structures tertiaires révèlent la présence de trois domaines : un domaine N-terminal replié en super sandwich β distordu suivi du domaine catalytique A constitué d'un tonneau $(\beta/\alpha)_8$ pour finir par un domaine C-terminal formant un feuillet β (Figure 14). Les études structurales montrent que le domaine N contribue à la formation du site actif par l'intermédiaire de boucles qui interagissent avec le domaine A (Kitamura *et al.*, 2008). Une partie du domaine C interagit avec le domaine catalytique en protégeant du solvant certaines hélices du domaine A. Aussi, ceci suggère une stabilisation de ce dernier (Kitamura *et al.*, 2008). Il apparait également que le domaine A n'est pas composé d'un tonneau $(\beta/\alpha)_8$ strict. En effet, le brin 7 du tonneau canonique est ici présent sous forme d'une boucle.

La résolution des structures d'enzymes appartenant à cette famille montre la présence d'un ion calcium au niveau du site actif coordonné à plusieurs acides glutamiques ainsi que des molécules d'eau. Des études enzymatiques montrent que cet ion divalent est nécessaire à l'hydrolyse du substrat et qu'il jouerait un rôle dans l'assistance acide au cours de la réaction (Gloster *et al.*, 2008; Kitamura *et al.*, 2008; Okuyama *et al.*, 2009). Il apparait dans la structure de Kitamura qu'une molécule d'eau proche du carbone anomérique semble être la meilleure candidate pour être la molécule d'eau catalytique. Deux acides glutamiques liant cette dernière joueraient le rôle de bases catalytiques et faciliteraient l'attaque nucléophile de celle-ci. Dans le cas de l'α-galactosidase (Okuyama *et al.*, 2009), les études structurales montrent que deux acides glutamiques coopéreraient pour agir comme résidus « acide/base ». Dans cette structure, l'acide aspartique qui jouerait le rôle de résidu nucléophile se trouve à la place de la molécule d'eau supposée catalytique chez SusB (Okuyama *et al.*, 2009).

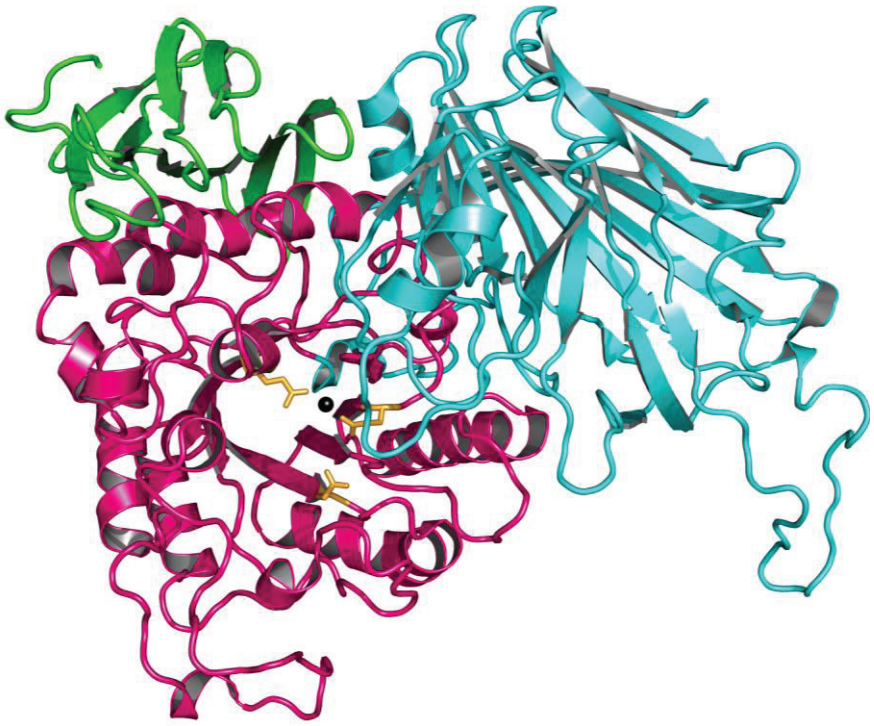


Figure 14: Représentation en ruban de la structure tridimensionnelle résolue par cristallographie aux rayons X de l'α-glucosidase de *Bacteroides thetaiotaomicron* (Kitamura *et al.*, 2008). Le domaine N, le domaine catalytique A et le domaine C sont respectivement colorés en cyan, rose et vert. Les résidus catalytiques supposés sont représentés en bâtons et colorés en jaune. L'ion calcium sous forme de sphère est coloré en noir.

F. Famille GH122

Très peu d'informations sont rapportées pour cette famille. Elle a été créée après des études bioinformatiques suggérant la découverte d'une nouvelle famille de GH (Comfort *et al.*, 2008). Sont retrouvées dans cette famille des enzymes issues d'organismes appartenant au règne des Archaea et considérés comme extrémophiles. Une seule enzyme de cette famille a été caractérisée (Costantino *et al.*, 1990) révélant ainsi son caractère hyperthermophile.

Jusqu'à présent, il n'existe aucun représentant structural pour cette famille et le mécanisme d'hydrolyse n'est pas encore connu.

G. Récapitulatif

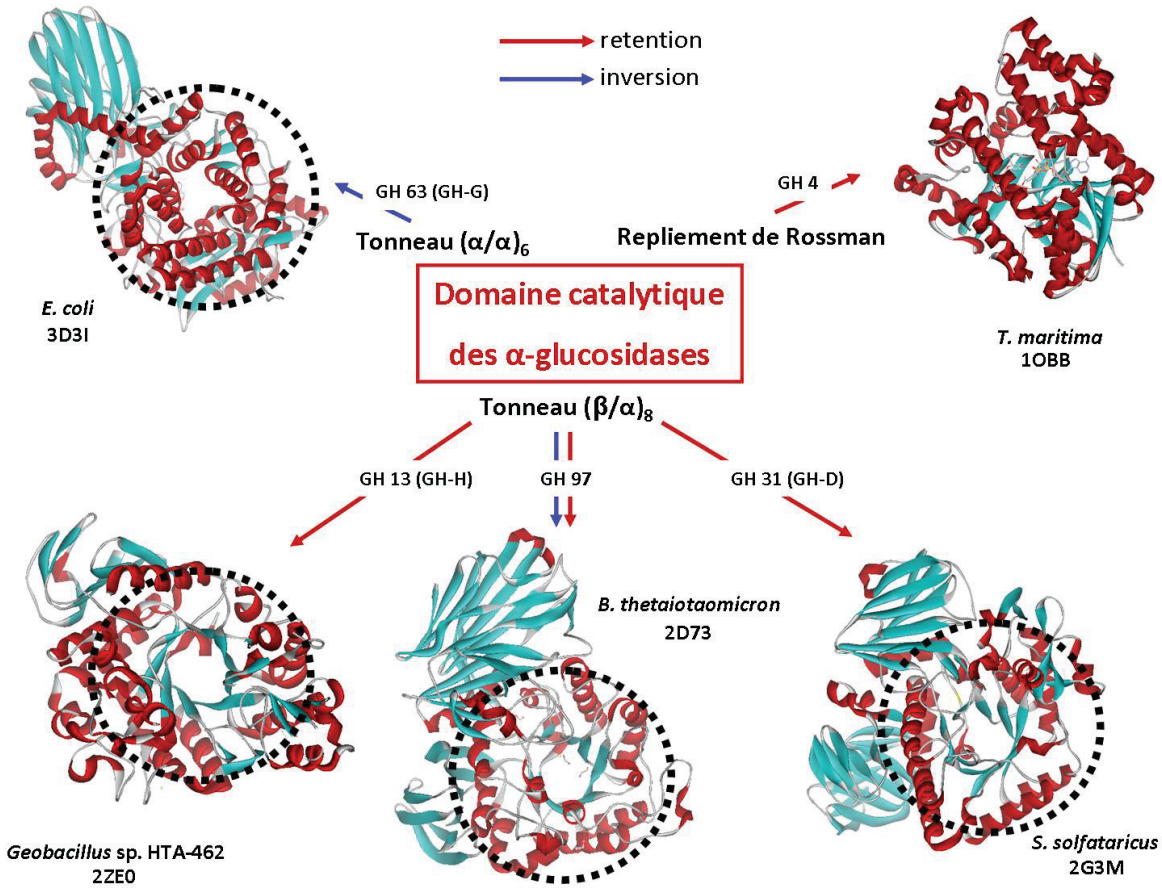
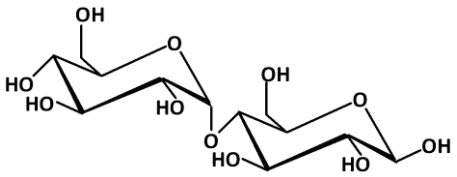
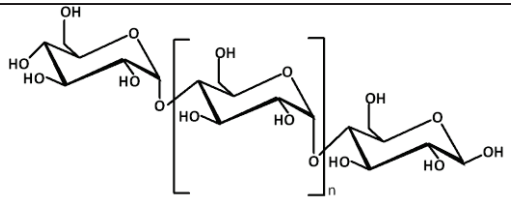
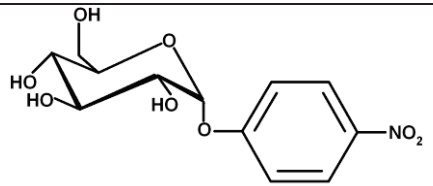
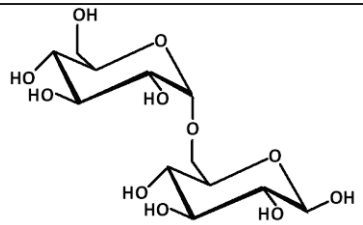
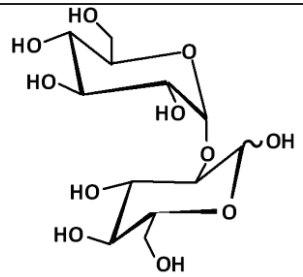


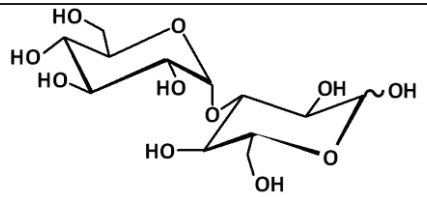
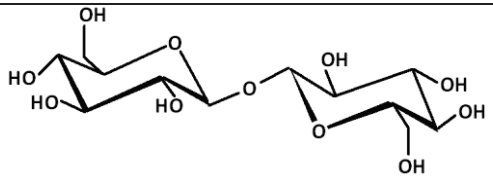
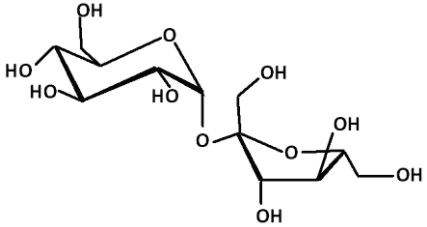
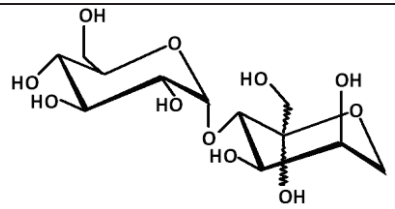
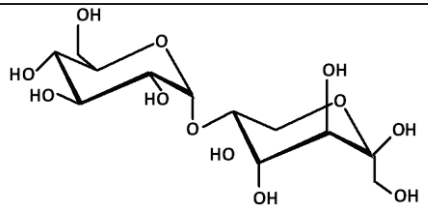
Figure 15: Représentations en ruban des repliements du domaine catalytique chez les différentes familles de GH où sont retrouvées les α-glucosidases (EC 3.2.1.20). Les flèches rouges et bleues indiquent selon quel mécanisme les différentes familles agissent.

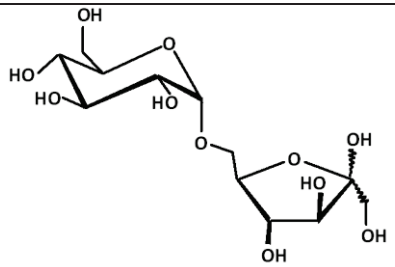
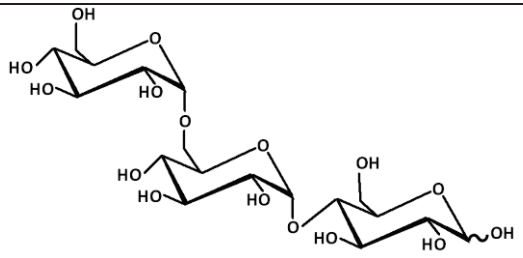
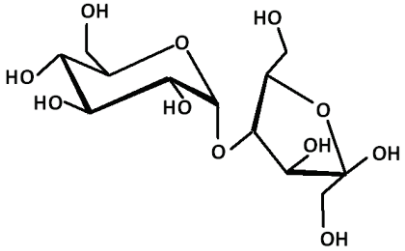
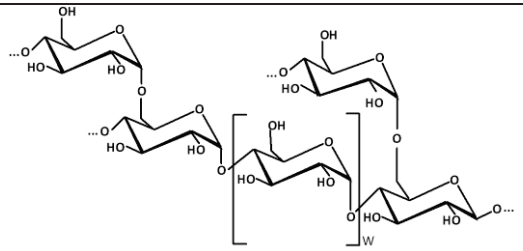
II. Caractéristiques biochimiques

A. Spécificité de substrat

Bien que peu de structures tridimensionnelles d'α-glucosidases aient été résolues jusqu'à présent, il s'avère qu'un certain nombre de ces enzymes a été étudié afin de déterminer leurs caractéristiques biochimiques et leur spécificité. Ainsi, la littérature révèle que ces enzymes ont une grande spécificité de substrat, contrairement à ce qui était annoncé selon la nomenclature IUPAC-IUBMB. En effet, ces études montrent que différents types de liaison glucosidique peuvent être hydrolysés : α-(1,1), α-(1,2), α-(1,3), α-(1,4) et α-(1,6). Différentes formes d'hétérodiholosides composés de glucose et de fructose sont également susceptibles d'être hydrolysées par les α-glucosidases. Le Tableau 1 dresse une liste non exhaustive des principaux sucres hydrolysés par les α-glucosidases.

Sucre	Structure	Composition	Références
Maltose		α-D-glucopyranosyl-(1,4)-D-glucopyranose	Nakao *et al.*, 1994a; Kato *et al.*, 2002; Tanaka *et al.*, 2002; Giuliano *et al.*, 2004; Ferrer *et al.*, 2005; Hung *et al.*, 2005; Angelov *et al.*, 2006; Giannesi *et al.*, 2006; Alarico *et al.*, 2008; Kitamura *et al.*, 2008; Kang *et al.*, 2009
MOS		α-D-glucopyranosyl-(1,4)-[α-D-glucopyranosyl-(1,4)]$_n$-D-glucopyranoside	Suzuki *et al.*, 1997; Giuliano *et al.*, 2004; Ferrer *et al.*, 2005; Hung *et al.*, 2005; Kitamura *et al.*, 2008; Kang *et al.*, 2009
pNPG		*p*-nitrophenyl-α-D-glucopyranoside	Tanaka *et al.*, 2002; van den Broek *et al.*, 2003; Hung *et al.*, 2005; Angelov *et al.*, 2006; Kitamura *et al.*, 2008; Kang *et al.*, 2009
Isomaltose		α-D-glucopyranosyl-(1,6)-D-glucopyranose	Nakao *et al.*, 1994a; Suzuki *et al.*, 1997; van den Broek *et al.*, 2003; Alarico *et al.*, 2008; Kitamura *et al.*, 2008; Kang *et al.*, 2009
Kojibiose		α-D-glucopyranosyl-(1,2)-D-glucopyranose	Tanaka *et al.*, 2002; Alarico *et al.*, 2008; Kitamura *et al.*, 2008; Kang *et al.*, 2009

Nigérose		α-D-glucopyranosyl-(1,3)-D-glucopyranose	Nakao *et al.*, 1994a; Suzuki *et al.*, 1997; Tanaka *et al.*, 2002; Alarico *et al.*, 2008; Kitamura *et al.*, 2008; Kang *et al.*, 2009
Tréhalose		α-D-glucopyranosyl-(1,1)-α-D-glucopyranoside	Nakao *et al.*, 1994a; van den Broek *et al.*, 2003; Alarico *et al.*, 2008
Sucrose		α-D-glucopyranosyl-(1,2)-β-D-fructofuranoside	Nakao *et al.*, 1994b; Ferrer *et al.*, 2005; Alarico *et al.*, 2008
Turanose		α-D-glucopyranosyl-(1,3)-D-fructopyranose	Suzuki *et al.*, 1997; Alarico *et al.*, 2008; Kang *et al.*, 2009
Leucrose		α-D-glucopyranosyl-(1,5)-D-fructopyranose	Alarico *et al.*, 2008; Kang *et al.*, 2009

Palatinose		α-D-glucopyranosyl-(1,6)-D-fructofuranose	Suzuki *et al.*, 1997; Alarico *et al.*, 2008
Panose		α-D-glucopyranosyl-(1,6)-α-D-glucopyranosyl-(1,4)-D-glucopyranoside	Suzuki *et al.*, 1997; Alarico *et al.*, 2008; Kitamura *et al.*, 2008
Maltulose		α-D-glucopyranosyl-(1,4)-D-fructofuranose	Kang *et al.*, 2009
Amidon		Mélange d'amylose et d'amylopectine	Tanaka *et al.*, 2002; Giannesi *et al.*, 2006; Kitamura *et al.*, 2008; Kang *et al.*, 2009

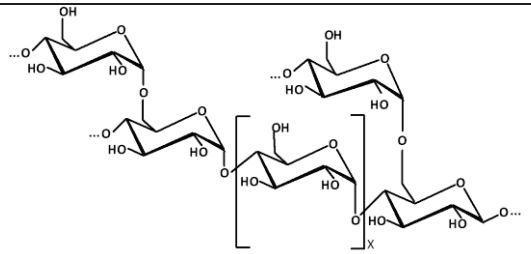
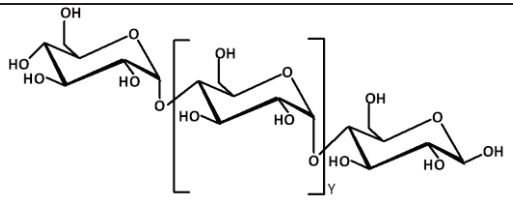
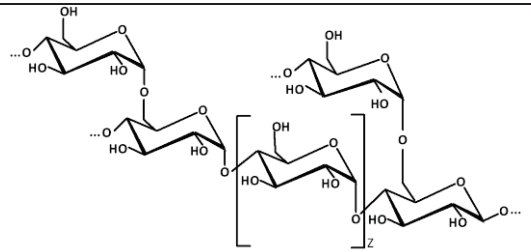
Amylopectine		α-D-glucopyranosyl-(1,4)-[α-D-glucopyranosyl-(1,4)]$_X$-D-glucopyranoside ramifié en α-(1,6)-D-glucopyranosyl (où X est compris entre 23 et 29)	Tanaka *et al.*, 2002
Amylose		α-D-glucopyranosyl-(1,4)-[α-D-glucopyranosyl-(1,4)]$_Y$-D-glucopyranoside (où Y est compris entre 250 et 6000)	Tanaka *et al.*, 2002; Giannesi *et al.*, 2006
Glycogène		α-D-glucopyranosyl-(1,4)-[α-D-glucopyranosyl-(1,4)]$_Z$-D-glucopyranoside ramifié en α-(1,6)-D-glucopyranosyl (où Z est compris entre 7 et 11)	Rolfsmeier *et al.*, 1998; Tanaka *et al.*, 2002; McVie-Wylie *et al.*, 2008

Tableau 1: Liste non exhaustive des principaux sucres hydrolysés par les α-glucosidases selon la littérature. MOS signifiant maltooligosaccharides.

B. Conditions optimales de réaction

Selon la littérature, les différentes α-glucosidases étudiées montrent qu'elles tendent à avoir un pH optimum acide, avec un indice atteignant même 3 pour *Oryza sativa, Penicillium purpurogenu* et *Ferroplasma acidophilum* (Yamasaki *et al.*, 1976; Yamasaki & Konno, 1987; Ferrer *et al.*, 2005). Selon la base de données BRENDA, compilant différents types d'informations sur les enzymes (www.brenda-enzymes.org) (Schomburg *et al.*, 2002), on peut remarquer que seulement quelques-unes (10% des enzymes répertoriées) ont un pH optimum de 7 ou au-delà (Figure 16) dont une seule à pH 8 (Zibaee *et al.*, 2009).

Seulement environ 20% d'entre elles supportent des variations de pH avec des écarts plus ou moins importants pouvant aller jusqu'à 6 unités pH pour les plus extrêmes (Yamasaki *et al.*, 1976). En comparaison, la plage moyenne des valeurs dans laquelle les enzymes conservent leur activité est de 3 à 4 unités pH (Brizova *et al.*, 1992; Rolfsmeier & Blum, 1995; Raasch *et al.*, 2000).

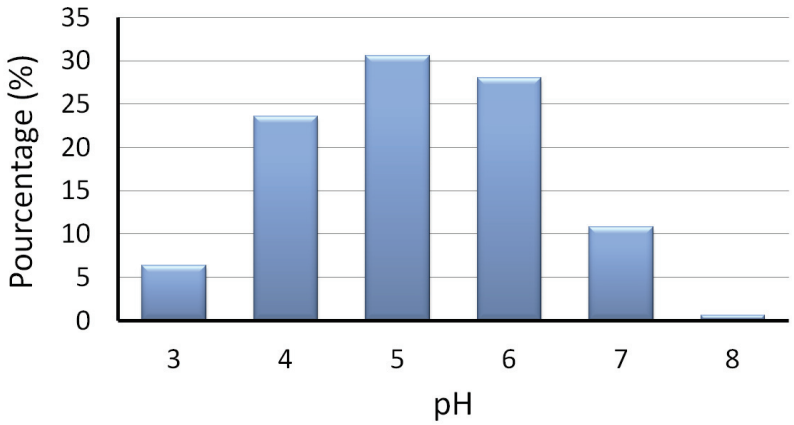


Figure 16 : Répartition des pH optimaux chez les différentes α-glucosidases caractérisées selon la base de données BRENDA.

Les températures optimales répertoriées dans cette même base de données sont retrouvées sur une large gamme (Figure 17) allant de 21°C à 105°C selon l'origine de l'enzyme (Rolfsmeier & Blum, 1995; Rowe & Margaritis, 2004). A partir de 80°C, les enzymes proviennent d'organismes thermo- ou hyperthermophiles (Costantino *et al.*, 1990; Rolfsmeier & Blum, 1995; Kashiwabara *et al.*, 2000).

Toujours selon cette base de données, une seule α-glucosidase (issue d'*Aspergillus niger*) est rapportée comme étant encore active à hauteur de 50% à 5°C (Brizova *et al.*, 1992). Pour les autres enzymes, les températures minimales supportées varient de 20°C à 95°C (Li & Chan, 1983; Costantino *et al.*, 1990) avec une activité se réduisant rapidement en dessous de ces températures.

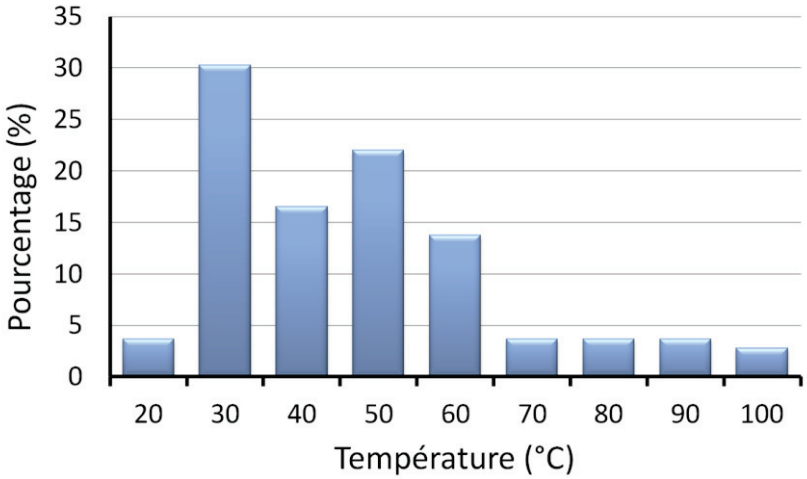


Figure 17 : Répartition des températures optimales chez les différentes α-glucosidases caractérisées selon la base de données BRENDA.

C. Récapitulatif des enzymes citées

Origine	GH	pH_{opt}	$T°C_{opt}$	Gamme pH- T°C	Référence
Aspergilus nidulans	31	5,5	45	-	Kato *et al.*, 2002)
Aspergillus niger	31	5	40	de 5°C à 50°C	Brizova *et al.*, 1992
Bacillus sp.	13	5,5	75	-	Nakao *et al.*, 1994a
Bacillus thermoamyloliquefaciens	31	6,8	80	-	Kashiwabara *et al.*, 2000
Bacillus thuringiensis	13	7	30	-	Rowe & Margaritis, 2004
Bacteroides thetaiotaomicron	97	6,5	45	-	Kitamura *et al.*, 2008
Bifidobacterium adolescentis	13	6,6	37	-	van den Broek *et al.*, 2003
Chaetomium thermophilum		7	60	de pH 4 à 7,5 de 35°C à 70°C	Giannesi *et al.*, 2006
Chilo suppressalis		8	45	-	Zibaee *et al.*, 2009
Ferroplasma acidophilum		3	50	de pH 2,4 à 3,5	Ferrer *et al.*, 2005
Geobacillus HTA-462	13	6,6	37	-	Hung *et al.*, 2005
Homo sapiens	31	-	37	-	McVie-Wylie *et al.*, 2008
Mortierella alliacea		5	55	-	Tanaka *et al.*, 2002
Oryza sativa		3	50	de pH 3 à 5	Yamasaki & Konno, 1987
Penicillium purpurogenum		3	50	de pH 1 à 7 de 20°C à 60°C	Yamasaki *et al.*, 1976
Picrophilus torridus	31	5	87	-	Angelov *et al.*, 2006
Pyrococcus furiosus		5	105	de pH 4,5 à 7,3 de 65°C à 115°C	Costantino *et al.*, 1990
Sulfolobus solfataricus	31	4,5	105	de pH 3,5 à 6,5 de 65°C à 115°C	Rolfsmeier & Blum, 1995
Thermotoga maritima	4	7,5	90	-	Raasch *et al.*, 2000
Thermus thermophilus	13	6,2	90	-	Alarico *et al.*, 2008

Tableau 2 : Caractéristiques des α-glucosidases citées. GH réfère à quelle famille CAZy de glycoside hydrolase appartient l'enzyme, pH_{opt} est le pH optimum et $T°C_{opt}$ est la température optimale en degré Celsius.

III. Les inhibiteurs

De manière générale, de nombreux inhibiteurs de glucosidases sont développés du fait des implications fondamentales de ces enzymes dans plusieurs processus biochimiques comme (1) l'hydrolyse de polysaccharides alimentaires en unités monosaccharidiques absorbées et utilisées par l'organisme, (2) le catabolisme lysosomal de glycoconjugués et la transformation des glycoprotéines et (3) la biosynthèse d'unités oligosaccharidiques dans les glycoprotéines et glycolipides. Ainsi, ces multiples fonctions justifient les recherches sur les inhibiteurs, notamment en vue d'être utilisés dans le traitement du diabète, de l'obésité, des maladies de surcharge lysosomale, les infections V.I.H. ou encore les tumeurs en général (Borges de Melo *et al.*, 2006). En outre, sur un plan plus fondamental, ces ligands sont des outils précieux pour l'étude des mécanismes de ces enzymes.

Lors de l'hydrolyse des glycosides, l'état de transition (de très courte durée) implique généralement un intermédiaire de type ion oxocarbénium. Ceci consiste en des liaisons partiellement formées ou rompues entre le nucléophile qui attaque et l'oxygène glycosidique du groupe partant. A ce stade, il y a accumulation de charges positives au niveau du cycle pyranose provoquant une distorsion du substrat proche de l'état de transition (Gloster & Davies, 2010). Certains inhibiteurs mimeraient cet état de transition formé furtivement durant le processus d'hydrolyse du glycoside. Ceci n'a été démontré que dans quelques cas où le ligand mimant l'état de transition se fixe à l'enzyme avec une très grande affinité (Wolfenden & Snider, 2001). Dans les années 1940, Pauling a fait le postulat que les inhibiteurs ont de meilleures affinités pour une enzyme lorsqu'ils miment la structure d'un « complexe activé contraint » (c'est-à-dire, l'état de transition) (Pauling, 1946; Pauling, 1948).

Il existe différents types chimiques d'inhibiteurs pour les α-glucosidases comme les thiosucres, les iminosucres et les carbasucres, par exemple.

A. Les thiosucres

Ces composés consistent en le remplacement de l'atome d'oxygène cyclique par un atome de soufre. Le salacinol (Figure 18), un composé naturel de cette famille, a été isolé de la plante *Salacia reticulata* par Yoshikawa et collaborateurs (Yoshikawa *et al.*, 1997) et est traditionnellement utilisé en Inde et au Sri Lanka pour traiter le diabète.

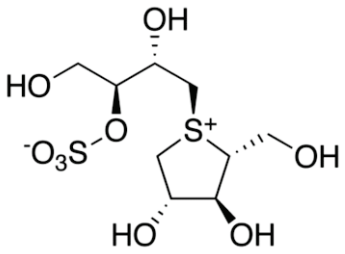


Figure 18 : Représentation du salacinol.

B. Les iminosucres

Les iminosucres, isolés de plantes ou de microorganismes, sont capables de mimer l'état de transition des substrats naturels des glycosidases. En effet, ils peuvent imiter la conformation et la charge de l'intermédaire oxocarbénium.

1. Déoxynojirimycine

Historiquement, le premier analogue du glucose découvert fut la nojirimycine en 1966. Cette molécule issue d'une souche *Streptomyces* comprend un atome d'azote endocyclique à la place de l'atome d'oxygène pyranosique (Inouye *et al.*, 1966). Toutefois, il s'agit d'un composé très instable du fait de la présence d'un groupement hydroxyle au niveau du carbone C-1 (Inouye *et al.*, 1968). La synthèse d'un composé plus stable, dérivé de celui-ci, auquel manque cet hydroxyl a donc été réalisée par Paulsen et collaborateurs (Paulsen *et al.*, 1967). Il s'avère que cette molécule appelée déoxynojirimycine (Figure 19) est également naturellement présente dans les racines de mûrier (Daigo *et al.*, 1986). Un dérivé de cet inhibiteur, le miglitol (Diastabol® en France et Glyset® aux Etats-Unis) est commercialisé comme antidiabétique.

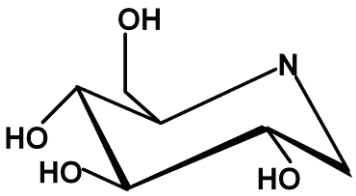


Figure 19 : Représentation de la déoxynojirimycine.

2. *Castanospermine*

Cette molécule trouvée dans l'arbre *Castanospermum australe* consiste en la fusion de cycles pipéridine et pyrrolidine. Plus schématiquement, il s'agit d'un dérivé de la déoxynojirimycine associé à un cycle à 5 atomes avec l'atome d'azote en commun à la jonction, et présentant des propriétés structurales et biologiques similaires (Figure 20). Cette structure engendre une certaine rigidité expliquant son activité anti-glucosidase (Borges de Melo *et al.*, 2006).

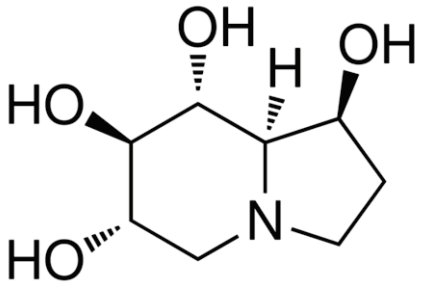


Figure 20 : Représentation de la castanospermine.

C. Les carbasucres

Il s'agit d'analogues de sucres dont l'oxygène intracyclique est remplacé par un atome de carbone ; l'exemple le plus connu étant l'acarbose (commercialisé sous le nom de Glucobay® en Europe, Precose® aux Etats-Unis et Glucor®/Acarbose en France). Il est possible de traiter des patients diabétiques avec ce composé qui peut être industriellement produit chez les actinomycètes (Wehmeier & Piepersberg, 2004).

L'acarbose est un pseudotétrasaccharide (Figure 21) dont l'acarviosine est lié à un maltose (Junge *et al.*, 1984). L'acarviosine est composée d'une unité de valiénamine (le carbasucre) lié à un dideoxyglucose par une liaison *N*-glycosidique qui mime la liaison *O*-glycosidique des substrats naturels. Il s'agit d'une molécule capable d'inhiber des α-glucosidases d'origine diverses : bactériennes (Kitamura *et al.*, 2008) ou mammifères (Sim *et al.*, 2008) par exemple.

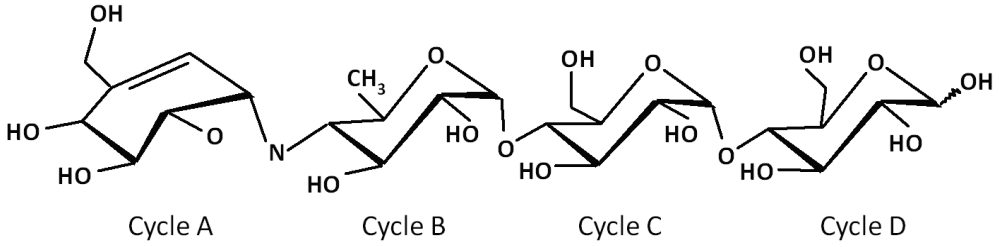


Figure 21 : Représentation de l'acarbose. Le cycle A seul correspond à l'unité valiénamine. Le dihétéroside A+B correspond à l'acarviosine.

D. Le Tris

Les molécules présentées jusqu'alors possédaient une structure relativement proche des substrats naturels des α-glucosidases avec certaines modifications chimiques leur conférant leurs propriétés inhibitrices.

Le Tris (Tris(hydroxyméthyl)aminométhane) est un exemple radicalement différent d'inhibiteur (Figure 22). Il a été rapporté à plusieurs reprises que ce composé inhibait de façon compétitive des α-glucosidases (Bravo-Torres *et al.*, 2004; Kitamura *et al.*, 2008; Zhou *et al.*, 2009).

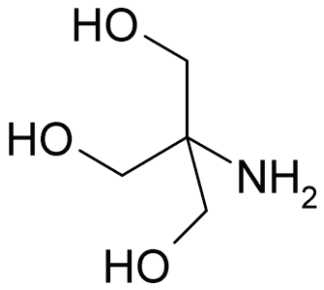


Figure 22 : Représentation d'une molécule de Tris.

Chapitre 4 : Les applications et la bio-ingénierie

Comme définis précédemment, les GH et les GT se trouvent être extrêmement utiles à la manipulation des saccharides *in vitro*. L'utilisation des glycosidases dans les procédés industriels devient de plus en plus répandue de par leur capacité à agir selon des conditions dites « vertes » et la facilité avec laquelle il est possible d'en produire en grandes quantités (Panke *et al.*, 2004; Panke & Wubbolts, 2005).

Ainsi, les glycosidases ont été soumises à de nombreuses études d'ingénierie protéique dans le but d'améliorer ou d'altérer par des stratégies de mutagenèse dirigée leurs activités catalytiques, leur spécificité de substrat ou encore leurs conditions optimales de réaction (Hancock *et al.*, 2006). Ces enzymes se sont avérées être un outil de choix pour la synthèse de glycosides par des réactions de transglycosylation avec de nombreuses applications fructueuses (Feng *et al.*, 2005).

L'utilisation de ces enzymes dans le cadre de la synthèse de sucres est limitée par deux paramètres : 1) conduire quand bien même la réaction qui est thermodynamiquement défavorable et 2) limiter ou abolir la dégradation enzymatique du produit de la réaction (Hancock *et al.*, 2006).

Une nouvelle classe d'enzymes mutées, appelées glycosynthases, a été introduite afin d'autoriser la synthèse de nouvelles liaisons glycosidiques tout en empêchant l'hydrolyse des liaisons nouvellement formées (Mayer *et al.*, 2000), ce, en mutant le résidu nucléophile en un autre non nucléophile (Gly, Ala ou Ser) (Sugimura *et al.*, 2006; Blanchard *et al.*, 2007) et en utilisant des donneurs glycosyl fluorides (Figure 23).

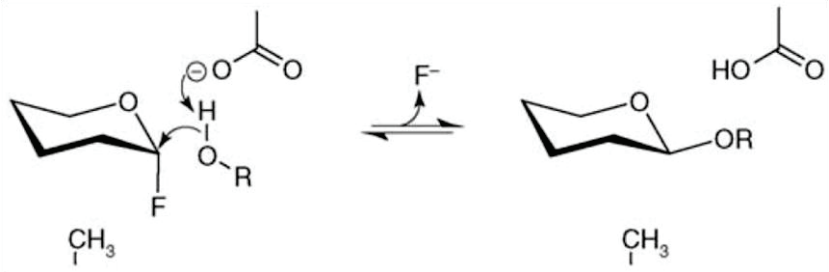


Figure 23 : Schéma de réaction d'une glycosynthase (issu de Hancock *et al.*, 2006).

Jusqu'en 2006, les glycosynthases obtenues agissaient selon un mécanisme de rétention de la configuration du carbone anomérique. Des études sur une exo-oligoxylanase agissant selon un mécanisme d'inversion et des mutations de cette enzyme sur un unique résidu (celui servant de base catalytique) ont mené à la découverte de neufs substitutions conférant une activité glycosynthase (Honda & Kitaoka, 2006). Une glycosynthase dérivée

d’une autre « *inverting* » a été étudiée en 2012 (Ohnuma *et al.*, 2012). Ainsi, ces perspectives semblent prometteuses et pourraient s’étendre au cadre de la synthèse de sucres prébiotiques en mutant en glycosynthases des α-glucosidases possédant des activités et des spécificités particulières et opérant selon un mécanisme de rétention ou d’inversion.

Chapitre 5 : Les α-glucosidases du projet

I. Issue de *Shewanella* sp. ANA-3

A. Informations sur la souche

Les *Shewanellae* font partie de communautés complexes dans les écosystèmes aquatiques et sédimentaires (Hofle & Brettar, 1996). Les membres de ce genre sont largement distribués partout dans le monde, dans des environnements marins et d'eau douce (Hau & Gralnick, 2007) et sont capables, entre autres, de supporter les basses températures et les fortes concentrations en sel (Karpinets *et al.*, 2010). Ces micro-organismes sont des γ-protéobactéries Gram-négatives possédant un flagelle unique. La majorité des espèces de *Shewanella* testées semble avoir une température optimale de croissance proche de 25°C en moyenne. Elles sont généralement classées comme des organismes mésophiles (Vieille *et al.*, 2001). Toutefois, elles montrent des différences dans leur capacité à maintenir leur croissance à plus basse température (Karpinets *et al.*, 2010) (Figure 24) et paraissent très souvent être psychrotolérantes (Hau & Gralnick, 2007). Ceci signifie qu'elles sont capables de se développer à basses températures (<5°C) avec, toutefois, une température optimale de croissance supérieure à 16°C. Cette capacité donne un avantage à ces organismes dans des environnements froids en permanence ou connaissant de grandes fluctuations thermiques.

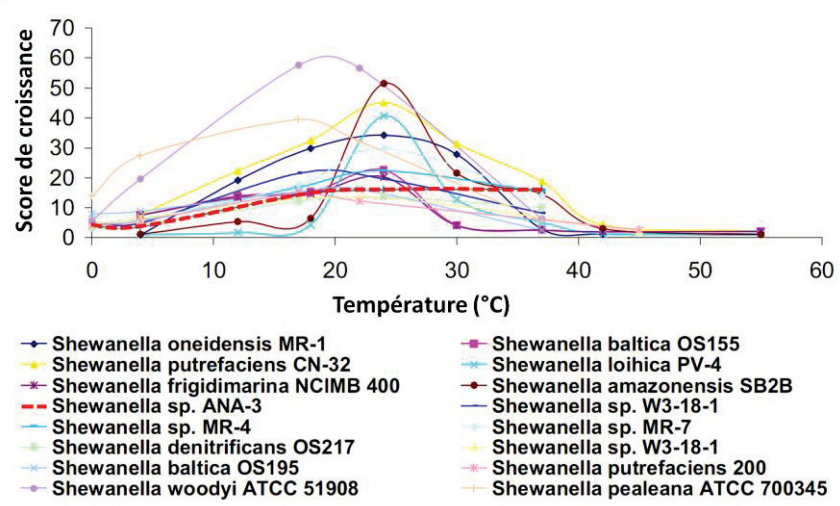


Figure 24 : Scores de croissance mesurés après 100h de culture en fonction de la température chez différentes souches de *Shewanella* (issu de Karpinets *et al.*, 2010).

Shewanella sp. ANA-3 est une souche qui a été isolée aux États-Unis (Woods Hole, Massachussetts) dans une eau saumâtre (mélange d'eau douce et d'eau salée), au niveau d'un ponton en bois traité à l'arsenic (Saltikov *et al.*, 2003). Sur un total de 4111 gènes codant des protéines, cette souche en possède 87 impliqués dans des voies de métabolisme des sucres (Rodionov *et al.*, 2010). Un mécanisme typique d'utilisation de sucre comprend un système de transport et des enzymes intracellulaires nécessaires aux transformations biochimiques de ces composés, dont les gènes sont souvent organisés en opéron. Aussi, il s'avère que *Shewanella* sp. ANA-3 possède une voie métabolique associée à l'utilisation de l'amidon comprenant une α-glucosidase (Figure 25, appelée ici SusB). Il semblerait que le glucose soit produit dans le périplasme de la bactérie pour ensuite être transporté vers le cytoplasme (Rodionov *et al.*, 2010), d'autant que la souche est capable de se développer dans un milieu contenant des maltodextrines comme source unique de carbone (Rodionov *et al.*, 2010).

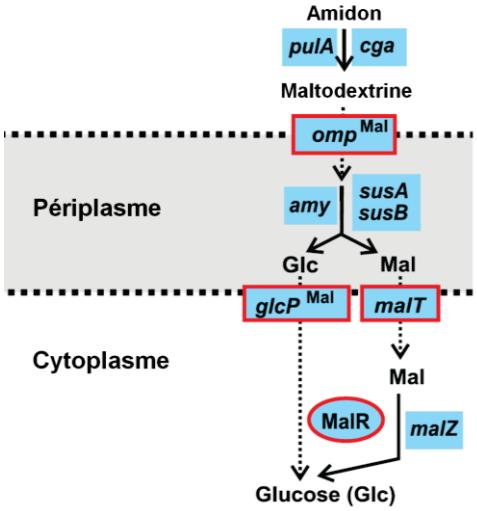


Figure 25 : Mécanisme d'utilisation supposé des maltodextrines chez *Shewanella* sp. ANA-3 (adapté de Rodionov *et al.*, 2010). Les enzymes sont surlignés en bleu, les transporteurs sont colorés en bleu et encadrés en rouge, le facteur de transcription est coloré en bleu et encerclé en rouge. Les différents sucres présents et/ou libérés dans cette voie sont l'amidon, les maltodextrines, le maltose et le glucose.

B. Informations sur l'enzyme

La séquence protéique de l'α-glucosidase de *Shewanella* sp. ANA-3 (« SHWaglu », code Uniprot : A0KWJ3) dont le génome complet a été séquencé (référence de séquence NCBI : NC_008573.1) est classée dans la base de données KEGG (www.genome.jp/kegg). Le gène correspondant (*SHWaglu*) est répertorié dans cette dernière sous le code Shewana3_1931 et possède 2012 pb. Ce gène se trouve à proximité d'autres gènes impliqués dans la voie d'utilisation du maltose codant une glucoamylase et un transporteur de maltooligosachharides (Figure 26).

Figure 26 : Localisation schématique du gène *SHWaglu* sur le génome de *Shewanella* sp. ANA-3. *RT* et *TM/M* représentent respectivement les gènes de régulateurs de transcription de la famille AraC et le gène d'un transporteur de maltose/maltooligosaccharides. « *3.2.1.3* » et « *3.2.1.20* » correspondent aux gènes d'une glucoamylase et de SHWaglu.

Les caractéristiques générales et théoriques de cette protéine ont pu être déterminées grâce à l'outil ProtParam Tool sur ExPASy (Gasteiger *et al.*, 2005) : avec 703 résidus, cette enzyme a un poids moléculaire de 79 kDa et un pI théorique de 6.

Dans la base de données CAZy, cette protéine est référencée comme étant une GH97. Un alignement de séquences a été réalisé avec des séquences des α-glucosidases des différentes familles de GH dont la structure tridimensionnelle a été résolue (Tableau 3). Le meilleur pourcentage d'identité (36%) pour l'α-glucosidase *Shewanella* sp. ANA-3 est obtenu avec l'α-glucosidase de *B. thetaiotaomicron* appartenant à la famille GH97. Un alignement de séquence a été réalisé avec ces deux dernières (Figure 27).

Origine	Code Uniprot	Famille	Code PDB	Référence	Identité de séquence
Thermotoga maritima	O33830	GH4	1OBB	Lodge *et al.*, 2003	**12%**
Geobacillus sp HTA-462	Q33E90	GH13	2ZE0	Shirai *et al.*, 2008	**12%**
Sulfolobus solfataricus	P0CD66	GH31	2G3M	Ernst *et al.*, 2006	**10%**
Escherichia coli K-12	MG1655	GH63	3D3I	Kurakata *et al.*, 2008	**11%**
Bacteroides thetaiotaomicron	P71094	GH97	2JKA 2D73	Gloster *et al.*, 2008 Kitamura *et al.*, 2008	**36%**

Tableau 3 : Récapitulatif des pourcentages d'identité de séquence obtenus lors de l'alignement de la séquence de SHWaglu avec celles d'α-glucosidases de structure connue.

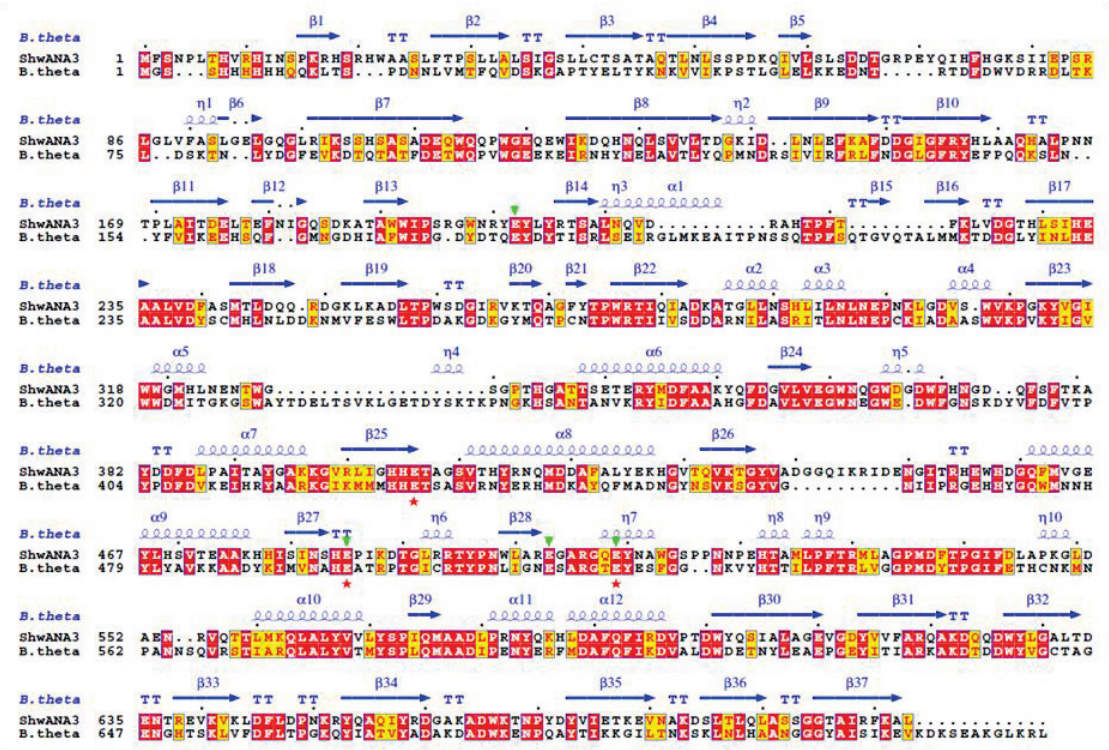


Figure 27 : Alignement de séquences des α-glucosidases de *Shewanella* sp. ANA-3 et de *B. thetaiotaomicron* (SusB). Les résidus complexant l'ion calcium et ceux potentiellement catalytiques sont signalés respectivement par un triangle vert et une étoile rouge. Les structures secondaires issues de la structure tridimensionnelle de SusB sont colorées en bleu.

II. Issue de *Lactococcus lactis* IL1403

A. Informations sur la souche

Anciennement nommée *Streptococcus lactis*, *Lactococcus lactis* est retrouvée à la surface ou dans le tractus gastro-intestinal d'animaux mais aussi chez des plantes. Cette espèce fait partie des bactéries lactiques (LAB), groupe de micro-organismes convertissant les sucres en acide lactique (Bolotin *et al.*, 2001). Les LAB regroupent des bactéries pathogènes (comme les *Streptococcus pneumoniae* ou *pyogenes*), mais aussi des micro-organismes utiles. Ainsi, *L. lactis* est utilisée dans l'industrie laitière comme déclencheur de fermentation du fromage mais aussi pour préparer la bière, le vin et les légumes marinés.

Lactococcus lactis IL1403 est une bactérie Gram-positive couramment utilisée dans l'industrie agro-alimentaire (de Vos, 1996) mais est également un organisme modèle pour des études moléculaires grâce au fait que son génome a entièrement été séquencé (Bolotin *et al.*, 2001). Cette bactérie mésophile, avec une température optimale de croissance de 30°C (Panoff *et al.*, 1994), est incapable de se développer en dessous de 8°C et a pour température létale 52°C (Panoff *et al.*, 1995). Son génome code 2310 protéines dont 32 enzymes impliquées dans la dégradation de oligo- et polysaccharides (Bolotin *et al.*, 2001; Cantarel *et al.*, 2009).

B. Informations sur l'enzyme

La séquence protéique de l'α-glucosidase de *Lactococcus lactis* IL1403 (code Uniprot : Q9CF03) dont le génome complet a été séquencé (NCBI complete genome : NC_002662.1) (Bolotin *et al.*, 2001), est classée dans la base de données KEGG. Le gène correspondant (*IL1403*) est répertorié par cette dernière sous le code L128691 et possède 1638 pb. Ce gène se trouve à proximité d'autres gènes impliqués dans la voie d'utilisation du maltose codant pour des enzymes et des systèmes de transport de maltooligosaccharides (Figure 28).

Figure 28 : Localisation schématique du gène *IL1403* dans le génome de *Lactococcus lactis* IL1403. De gauche à droite, sont représentés les gènes d'une 6-phospho-β-glucosidase (*3.2.1.86*), maltose phosphorylase (*2.4.1.8*), α-glucosidase (*3.2.1.20*), α-amylase (*3.2.1.1*), maltose *O*-acétyltransférase (*2.3.1.79*), oligo-1,6-glucosidase (*3.2.1.10*), néopullulanase (*3.2.1.135*). S'ensuit *MalE*, le gène d'une protéine périplasmique soluble fixant le maltose et les maltodextrines et les délivrant à un transporteur. Pour finir, *MalF* et *MalG* codent des perméases de type transporteur ABC pour le maltose.

En utilisant ProtParam Tool sur ExPASy, les caractéristiques générales et théoriques de cette protéine ont pu être déterminées : avec 545 résidus, la protéine a un poids moléculaire de 63 kDa et un pI théorique de 5,6.

Dans la base de données CAZy, cette protéine est référencée comme étant une GH13. Un alignement de séquences a été réalisé avec des séquences des α-glucosidases des différentes familles de GH dont la structure tridimensionnelle a été résolue (Tableau 4). Ainsi le meilleur pourcentage d'identité pour l'α-glucosidase *Lactococcus lactis* IL1403 est obtenu avec l'α-glucosidase de *Geobacillus* sp. HTA-462 appartenant à la famille GH13 avec 44% d'identité. Un alignement de séquence a été réalisé entre ces deux dernières (Figure 29).

Origine	Code Uniprot	Famille	Code PDB	Référence	Identité de séquence
Thermotoga maritima	O33830	GH4	1OBB	Lodge *et al.*, 2003	**11%**
Geobacillus sp HTA-462	Q33E90	GH13	2ZE0	Shirai *et al.*, 2008	**44%**
Sulfolobus solfataricus	P0CD66	GH31	2G3M	Ernst *et al.*, 2006	**11%**
Escherichia coli K-12	MG1655	GH63	3D3I	Kurakata *et al.*, 2008	**12%**
Bacteroides thetaiotaomicron	P71094	GH97	2JKA 2D73	Gloster *et al.*, 2008 Kitamura *et al.*, 2008	**11%**

Tableau 4 : Récapitulatif des pourcentages d'identité de séquence obtenus lors de l'alignement de la séquence d'IL1403 avec celles d'α-glucosidases de structure connue.

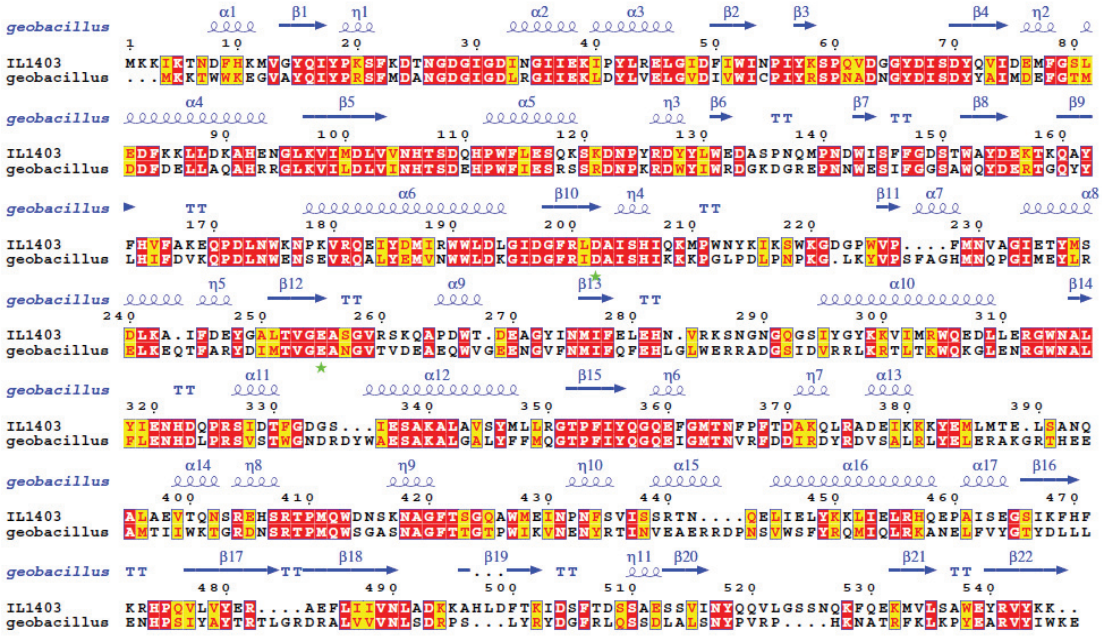


Figure 29 : Alignement de séquences des α-glucosidases de *Lactococcus lactis* IL1403 et de *Geobacillus* sp. HTA-462. Les résidus catalytiques sont indiqués par une étoile rouge. Les structures secondaires de l'α-glucosidase de *Geobacillus* sp. HTA-462 sont colorées en bleu.

III. Issue de *Lactobacillus bulgaricus* ATCC11842

A. Informations sur la souche

Lactobacillus delbrueckii ssp. *bulgaricus* ou, plus simplement, *Lactobacillus bulgaricus* est un des représentants économiquement les plus importants du groupe hétérogène des bactéries lactiques, avec une application mondiale dans la production de yaourt (van de Guchte *et al.*, 2006). Cet aliment est à la base du concept de « probiotique » (Metchnikoff, 1907; Heller, 2001) et il a été démontré que sa teneur en *Lactobacillus bulgaricus* vivant jouait un rôle dans l'atténuation de l'intolérance au lactose (Mercenier *et al.*, 2003).

A l'origine, la souche ATCC11842 a été isolée d'un yogourt bulgare par S. Orla-Jensen en 1919 (travaux non publiés, rapporté par van de Guchte *et al.*, 2006). L'étude de son génome associé à une comparaison avec d'autres membres du groupe des acidophiles et de *Streptococcus thermophilus* contribue à mieux comprendre son adaptation spécialisée au lait et sa protocoopération avec *S. thermophilus* (van de Guchte *et al.*, 2006). Il apparait que le génome de cette souche semble être en phase rapide d'évolution avec la présence significative de pseudogènes et de voies métaboliques incomplètes. Ceci indiquerait que la bactérie s'adapte à l'environnement lacté riche en lactose et, ce, par la perte de fonctions accessoires et par sa protocoopération avec *Streptococcus thermophilus* (van de Guchte *et al.*, 2006). Sa température optimale de croissance est 45°C (minimale 22°C ; maximale 52°C) (Heller, 2001).

B. Informations sur l'enzyme

La séquence protéique de l'enzyme de *Lactobacillus bulgaricus* ATCC11842 (code Uniprot : Q1GC15) dont le génome complet a été séquencé (NCBI complete genome : NC_008054.1) (van de Guchte *et al.*, 2006) est classée dans la base de données KEGG. Le gène correspondant (*11842aglu*) est répertorié par cette dernière sous le code Ldb0196 et possède 3006 pb. Il s'avère que ce gène est annoté comme protéine hypothétique et est localisé à proximité d'un gène codant une perméase MalG, une protéine permettant le transport de maltooligosaccharides. On peut également constater la présence de pseudogènes voisins, annotés comme étant *MalF* (transporteur de maltooligosaccharides) et *dexB* (glucane-1,6-α-glucosidase).

Figure 30 : Localisation schématique du gène *11842* dans le génome de *Lactobacillus bulgaricus* ATCC 11842. Les gènes représentés en cyan correspondent à des pseudogènes, avec *dexB* et *MalF* annotés comme glucane-1,6-a-glucosidase (3.2.1.70) et transporteur de maltooligosaccharides, respectivement. *MalG* et *3.2.1.20* représentent les gènes d'un transporteur ABC pour les maltose/maltodextrines et *11842glu*.

En utilisant ProtParam Tool sur ExPASy, les caractéristiques générales et théoriques de cette protéine ont pu être déterminées : avec 1001 résidus, la protéine a un poids moléculaire de 113 kDa et un pI théorique de 6,7.

Dans la base de données CAZy, cette protéine est référencée comme étant une GH31. Un alignement de séquences a été réalisé avec des séquences des α-glucosidases membres des différentes familles de GH dont la structure tridimensionnelle a été résolue (Tableau 5). Au final, par ce protocole, il n'a pas été possible d'identifier à partir de la séquence complète un homologue structural au sein d'une quelconque famille.

Origine	Code Uniprot	Famille	Code PDB	Référence	Identité de séquence
Thermotoga maritima	O33830	GH4	1OBB	Lodge *et al.*, 2003	**9%**
Geobacillus sp HTA-462	Q33E90	GH13	2ZE0	Shirai *et al.*, 2008	**10%**
Sulfolobus solfataricus	P0CD66	GH31	2G3M	Ernst *et al.*, 2006	**12%**
Ruminococcus obeum	A5ZY13	GH31	3FFJ	Tan *et al.*, 2010	**11%**
Escherichia coli K-12	MG1655	GH63	3D3I	Kurakata *et al.*, 2008	**11%**
Bacteroides thetaiotaomicron	P71094	GH97	2JKA 2D73	Gloster *et al.*, 2008 Kitamura *et al.*, 2008	**10%**

Tableau 5 : Récapitulatif des pourcentages d'identité de séquence obtenus lors de l'alignement de la séquence de 11842 avec celles d'a-glucosidases de structure connue.

Matériel, Méthodes et Principes

Chapitre 1 : Techniques de biologie moléculaire

I. Plasmides

Les plasmides du système pET (Novagen®), initialement développés par Studier en 1990 (Studier *et al.*, 1990), sont dérivés du plasmide pBR322 et confèrent une résistance à l'ampicilline. Cette série de vecteurs d'expression permet la transcription régulée de gènes d'intérêt par l'ARN polymérase du bactériophage T7. Cette dernière doit être produite par la cellule hôte *E. coli* grâce à l'insertion d'une copie de gène dans le chromosome bactérien (bactéries BL21(DE3)). L'expression de la protéine d'intérêt est sous le contrôle du promoteur constitutif fort du bactériophage T7 et de la séquence opératrice lac. Les plasmides pET portent le gène *lacI* codant un répresseur de l'opérateur qui contrôle l'induction de la surproduction de la protéine (Figure 31).

L'IPTG (IsoPropyl-β-D-1-ThioGalactopyranoside) est fréquemment utilisé comme analogue de l'allolactose qui est un métabolite du lactose induisant l'opéron *lac* chez *E. coli*. Dans le cas de cellules transformées avec un vecteur pET, l'IPTG se lie au répresseur issu de la transcription du gène *lacI*, l'empêchant de s'unir à l'opérateur en amont du gène d'intérêt et lève ainsi l'inhibition de la transcription du gène sous contrôle du promoteur T7. L'IPTG n'étant pas métabolisé par *E. coli*, sa concentration reste donc constante tout au long de la phase d'induction.

Les bactéries BL21(DE3) pLysS permettent de prévenir une expression basale de la protéine en empêchant une fuite du promoteur en absence d'IPTG, et ce par la production du lysozyme T7 qui est un inhibiteur naturel de l'ARN polymérase T7.

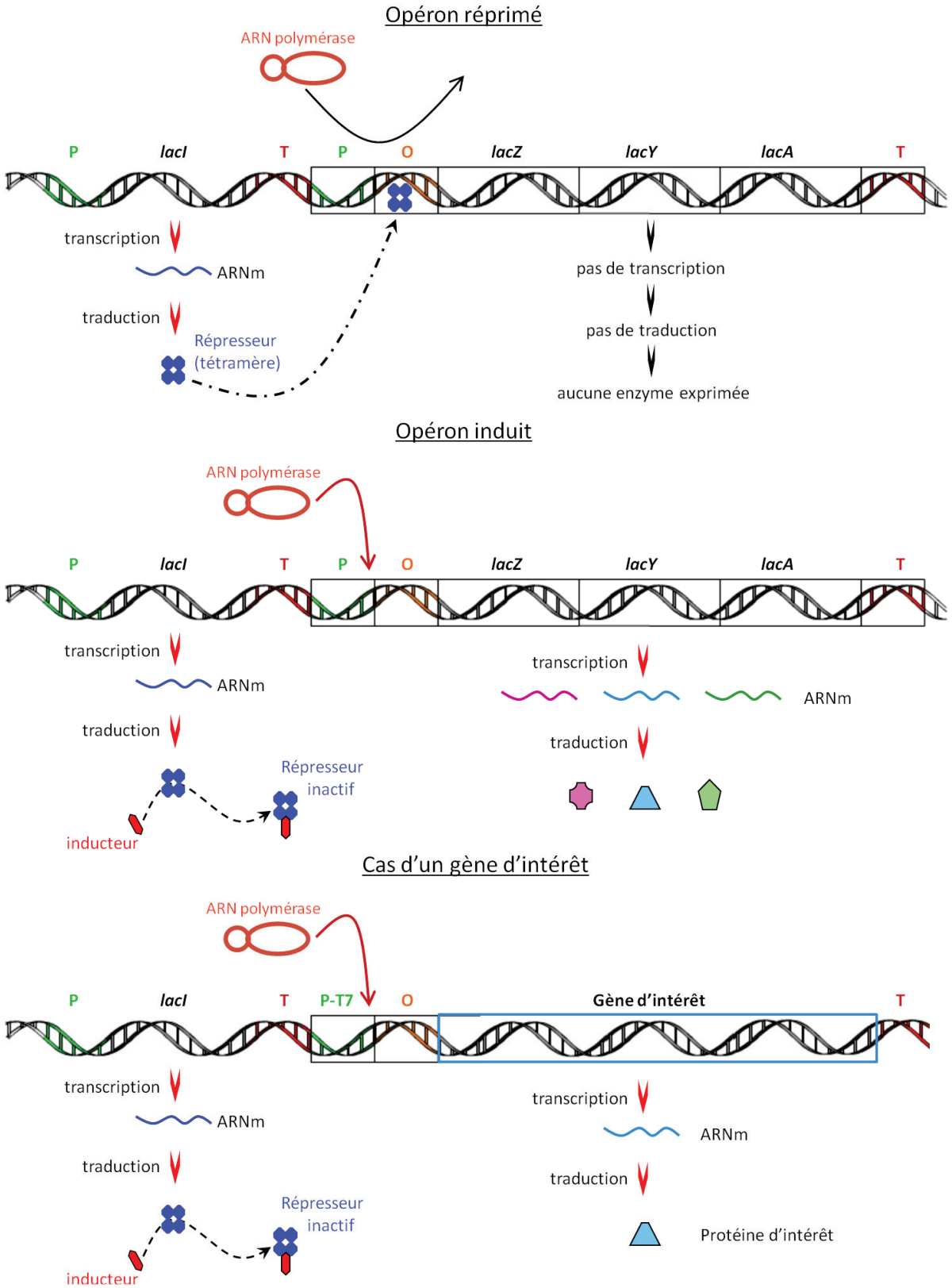
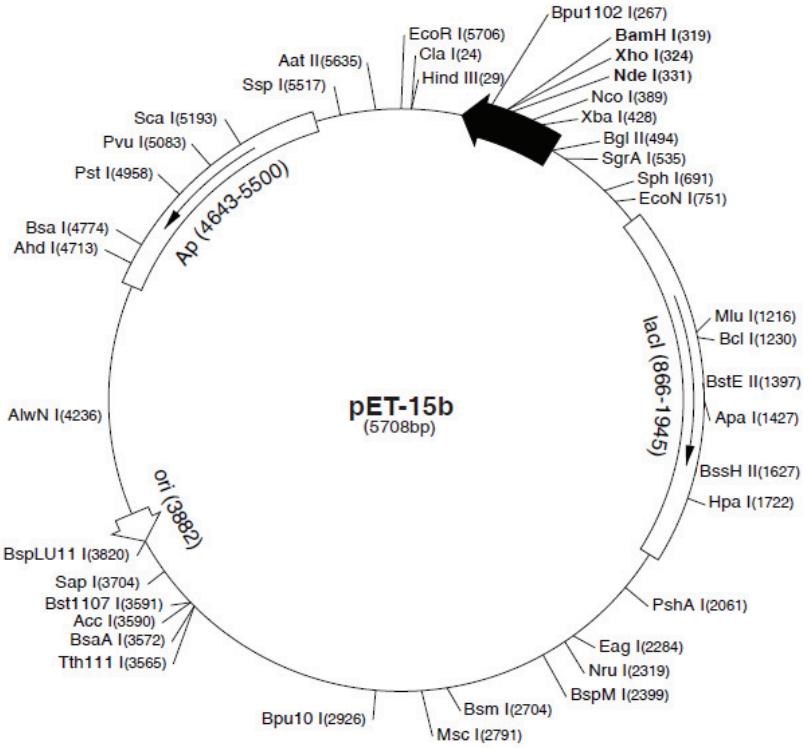


Figure 31 : Fonctionnement de l'opéron *lac* chez les bactéries *E. coli*. En l'absence de l'inducteur, le répresseur est produit et se fixe sur le site opérateur (« O » sur le schéma) et empêche ainsi la transcription des gènes placés en aval. En présence d'inducteur (lactose ou IPTG), ce répresseur ne se fixe plus sur le site opérateur, permettant la transcription des gènes et donc la production des protéines.

Le vecteur pET15b porte une séquence N-terminale His-Tag $(His)_6$ suivie d'un site de coupure à la thrombine puis de trois sites de clonage (Figure 32). Ceci permet la fusion d'une étiquette poly-histidine suivie d'un site de coupure à la thrombine à l'extrémité N-terminale de la protéine d'intérêt.

pET-15b sequence landmarks	
T7 promoter	453-469
T7 transcription start	452
His•Tag coding sequence	362-380
Multiple cloning sites	
(*Nde* I - *Bam*H I)	319-335
T7 terminator	213-259
lacI coding sequence	(866-1945)
pBR322 origin	3882
bla coding sequence	4643-5500



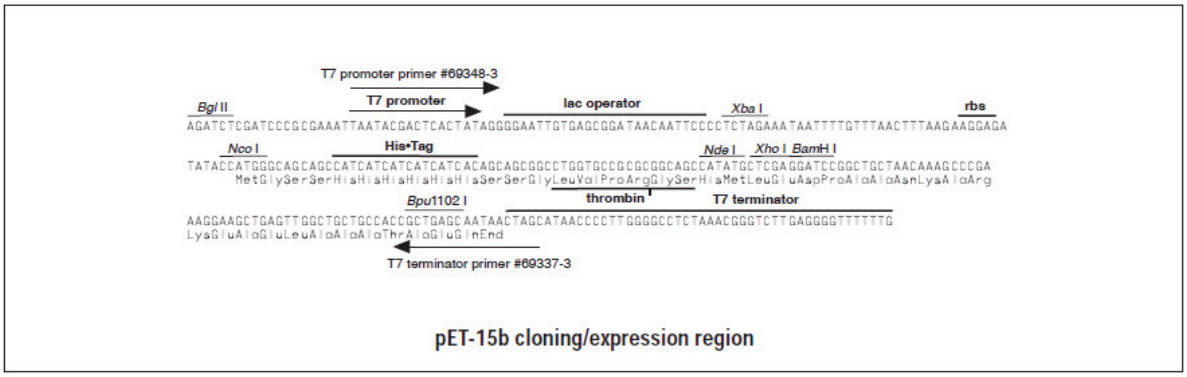


Figure 32 : Carte du plasmide pET15b et représentation de son site multiple de clonage (fourni par Novagen).

II. Techniques classiques de manipulation de l'ADN

A. Transformation de bactéries

La transformation est un transfert génétique au cours duquel de l'ADN bicaténaire, nu et en solution est introduit dans une bactérie dite compétente. Cette dernière est alors utilisée pour amplifier l'ADN exogène plasmidique.
La transformation de bactéries compétentes par le plasmide d'intérêt a été réalisée par la technique du choc thermique.

Environ 10 ng de préparation de plasmide circulaire ou d'ADN issu de réaction de ligature sont incubés avec 50 µl de bactéries compétentes de clonage (DH5α™, Invitrogen™ ou *E. cloni*® 10G, Lucigen®) ou d'expression (*E. cloni*® BL21 (DE3), Lucigen®) pendant 30 minutes dans la glace. Le volume d'ADN ne doit pas dépasser 10% du volume total. Les cellules sont ensuite soumises à un choc thermique à 42°C pendant 30 secondes puis plongées dans la glace pendant 2 minutes. Puis, les bactéries sont resuspendues dans 300 µl de milieu enrichi sans antibiotique afin de leur permettre l'expression de la résistance à l'antibiotique portée par le vecteur recombiné. Les clones recombinants sont sélectionnés grâce à un étalement de cette resuspension sur des boites de milieu sélectif contenant l'antibiotique adéquat. Les boites sont incubées toute une nuit à 37°C.

Les clones recombinants sont identifiés par digestion par des endonucléases de restriction ou par PCR après avoir extrait et purifié les plasmides. Un séquençage automatisé permet de s'assurer que l'étape de PCR n'a pas introduit de mutation.

B. Extraction et purification de l'ADN plasmidique

1. Miniprep

A partir de 7 à 10 ml de culture bactérienne - contenant les antibiotiques requis - incubés toute une nuit à 37°C, les ADNs plasmidiques sont extraits et purifiés à partir du kit GeneJET™ Plasmid Miniprep (Thermo Scientific). Après fixation sur la membrane de silice fournie, les ADNs plasmidiques sont élués avec de l'eau Milli-Q.

2. MidiPrep

A partir de 100 à 150 ml de culture - contenant le(s) antibiotique(s) requis - incubés toute une nuit à 37°C, les ADNs plasmidiques sont extraits et purifiés à l'aide du kit NucleoBond® Xtra Midi (Macherey-Nagel).

Suite à leur extraction, les ADNs sont précipités avec 3,5 ml d'isopropanol pendant 5 minutes à température ambiante, centrifugés à 13000 g pendant 20 minutes à 4°C, puis rincés à l'éthanol 70%, séchés et, enfin, repris dans 200 µl d'eau Milli-Q.

C. Mesure de la concentration en ADN

La quantité d'ADN est estimée par mesure de son absorbance dans l'UV à 260 nm *via* un spectrophotomètre (1 unité DO correspondant à 50 µg d'ADN). La pureté de la préparation est vérifiée grâce au rapport A260/A280 ≥ 1,8 - valeur généralement acceptée pour considérer un ADN comme « pur ».

D. Digestion de l'ADN par des endonucléases de restriction

La réaction a généralement lieu dans un volume de 40 µl comprenant 4 µl de tampon de restriction 10X, 1 à 5 µg d'ADN et 0,5 à 1U d'endonucléase(s) par µg d'ADN. L'incubation à 37°C dure 1 à 2 heures. La réaction est arrêtée soit en transférant les tubes dans la glace, soit en inactivant de façon irréversible les endonucléases par un traitement à la chaleur à 65°C pendant 5 minutes.

E. Électrophorèse sur gel d'agarose

Les fragments d'ADN sont analysés par électrophorèse en gel d'agarose. Ces gels sont généralement composés de 1% d'agarose - 1 g pour 100 ml de tampon TAE (Tris 40 mM, pH 8 - acétate (acide acétique) 20 mM - EDTA 1 mM) - supplémenté de bromure d'éthidium (BET) à une concentration finale de 0,5 µg/ml. Le BET est un agent intercalant se liant aux acides nucléiques qui devient fluorescent lorsqu'il est exposé à des rayonnements UV.

Cette technique est utilisée aussi bien pour valider une étape (ADN purifié correspondant à la taille attendue, digestion totale/partielle, etc.) que pour isoler un ADN d'intérêt. Dans ce dernier cas, il faudra extraire le fragment d'ADN du gel (voir paragraphe suivant).

F. Extraction sur gel d'agarose

Après analyse électrophorétique, le fragment d'ADN d'intérêt visualisé peut être extrait grâce au kit QIAquick® (QIAGEN). Le principe consiste en la solubilisation de la portion de gel d'agarose contenant l'ADN d'intérêt suivi de son adsorption sur une membrane de silice puis de son élution grâce à de l'eau Milli-Q. Ceci permet d'éliminer une grande majorité des impuretés notamment grâce à une étape de lavage avant élution.

G. Ligature

La ligature d'un fragment d'ADN dans un vecteur est assurée par l'ADN ligase du phage T4. Cette dernière catalyse la formation d'une liaison covalente phosphodiester entre une extrémité hydroxyle 3' d'un brin et une extrémité phosphate 5' d'un autre brin. Il est possible de faire une ligature orientée - dans pET15b, par exemple - en utilisant des extrémités « cohésives » obtenues par les endonucléases d'intérêt ou une ligature aboutissant à deux orientations - dans le vecteur pGEMT-Easy - avec des extrémités « franches ».

La réaction est réalisée dans un volume de 10 à 20 µL contenant 1 unité de ligase et un rapport molaire vecteur:insert allant de 1:1 à 1:5 (mol/mol). La réaction est effectuée à 16°C pendant 18 heures.

H. Technique de PCR

L'amplification par PCR (Réaction de Polymérisation en Chaine) est une technique enzymatique d'amplification spécifique d'une séquence d'ADN. Elle emploie une ADN polymérase et deux oligonucléotides de synthèse qui s'hybrident avec des séquences qui leur sont complémentaires et bordant la séquence à amplifier. Elle est basée sur une répétition de cycles (le plus souvent 30 à 35) de variations de température après une première dénaturation à 95°C pendant 5 minutes. Chaque cycle nécessite plusieurs composés dans le milieu réactionnel (Tableau 6) et se décompose en trois étapes :

1. La dénaturation de l'ADN matriciel pendant 1 à 2 minutes.
2. L'hybridation spécifique des amorces oligonucléotidiques pendant 30 à 60 secondes à une température propre aux amorces utilisées (5°C inférieure à la température de fusion).
3. L'élongation par l'ADN polymérase à 72°C (température optimale de réaction de l'enzyme). Le temps d'élongation dépend de la taille du fragment à amplifier et de la vitesse de polymérisation de l'enzyme utilisée. La Taq ADN polymérase a une processivité d'environ 1000 bases/minute alors que celle de la Phusion® (Thermo Scientific) est d'environ 2000 bases/minute.

MILIEU REACTIONNEL	QUANTITE (CONCENTRATION) FINAL(E)
Tampon de PCR Phusion® 5X	10 µl (1X)
ADN matriciel	10 ng
Amorce oligonucléotidique 1 (10 µM)	4 µl (0,8 µM)
Amorce oligonucléotidique 2 (10 µM)	4 µl (0,8 µM)
dNTPs (10 mM)	1 µl (1mM, 200 µM de chaque dNTP)
DMSO	1,5 µl (3%)
ADN polymérase Phusion®	0,5 µl (0,2 U/µl)
Eau	q.s.p. 50 µl

Tableau 6 : Milieu réactionnel utilisé pour les réactions d'amplification d'ADN par PCR.

Après le dernier cycle, une dernière élongation est réalisée à 72°C pendant 10 minutes, ceci permettant à l'enzyme de terminer l'amplification des amplicons incomplets. Les PCR ont été réalisées dans un thermocycleur GeneAmp® PCR System 9700 (Applied Biosystems).

A la fin de la réaction, l'ADN néo-synthétisé est purifié selon le protocole utilisé pour l'extraction des fragments d'ADN d'un gel d'agarose, ceci permettant d'éliminer les oligonucléotides et les fragments qui ne sont pas à la taille attendue.

I. Mutagenèse dirigée

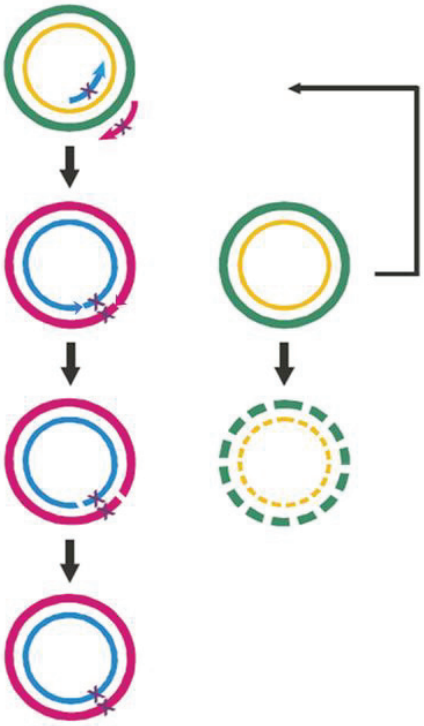
La mutagenèse dirigée est une technique de biologie moléculaire permettant d'apporter une (des) mutation(s) ponctuelle(s) ciblée(s) et voulue(s) dans la molécule d'ADN d'intérêt. Pour ce faire, il est nécessaire de modifier, dans la séquence du gène, le(s) codon(s) correspondant à (aux) l'acide(s) aminé(s) ciblé(s).

Le protocole employé est celui du kit commercial « Quikchange® II Site-Directed Mutagenesis Kit » (Agilent Technologies). Le principe est similaire à celui d'une amplification par PCR classique. L'ADN matriciel utilisé sera le vecteur de clonage ADN double brin contenant l'insert d'intérêt auquel la mutation doit être apportée. Chaque amorce doit contenir la mutation désirée et être complémentaire de la même séquence sur les brins opposés des brins du plasmide et seront prolongées pendant la réaction de PCR par la Phusion® High-Fidelity DNA polymérase (Finnzymes). La mutation apportée doit se situer au centre de l'oligonucléotide et être entourée de part et d'autre de 10 à 15 bases correspondant à la séquence correcte. Cette élongation génère un plasmide contenant la mutation désirée au niveau du gène d'intérêt.

Après la réaction PCR, avant la future étape de transformation de bactéries, il est nécessaire de réaliser une digestion de l'échantillon par l'endonucléase *Dpn*I qui ne reconnait que les sites de restriction méthylés. Il est à noter que seul l'ADN amplifié par une bactérie est méthylé. Ainsi, cette étape permet d'hydrolyser l'ADN parental matriciel et de sélectionner l'ADN néo-synthétisé contenant la mutation.

Ces différentes étapes sont résumées dans la

Figure 33.



<u>Synthèse des brins mutants</u> :

1) Dénaturation de l'ADN matriciel
2) Hybridation des amorces qui contiennent la mutation souhaitée.
3) Élongation des amorces avec l'ADN polymérase haute fidélité Phusion®

<u>Digestion de l'ADN matriciel avec *Dpn*I</u> :

Les ADN méthylés et hémiméthylés seront digérés par *Dpn*I

<u>Transformation</u> :

Transformation des molécules mutées dans les cellules compétentes pour réparer le brin et amplifier le plasmide

Figure 33 : Schéma du principe de la mutagénèse dirigée (fourni par Agilent Technologies).

III. Séquençage d'ADN

Le séquençage de fragments d'ADN a été réalisé selon un système automatisé proposé par la société GATC Biotech. Le principe de la réaction utilisée par les séquenceurs dérive de la méthode de Sanger (Sanger *et al.*, 1977).

Cette technique consiste à initier la polymérisation de l'ADN à partir d'un oligonucléotide complémentaire du fragment d'ADN à séquencer. Parmi les réactifs, il est nécessaire d'apporter les 4 désoxyribonucléotides (dATP, dTTP, dGTP et dCTP) indispensables à l'amplification mais aussi une quantité infime de didésoxyribonucléotides (ddNTPs). Les didésoxyribonucléotides agissent comme des terminateurs de chaines. Une fois intégrés, ils empêchent la poursuite de l'élongation. Comme ces ddNTPs sont en

quantité faible, la réaction produira de nouvelles molécules de tailles variables en fonction de quand a été intégré le ddNTP (Figure 34).

Avec l'automatisation, le marquage et la détection sont différents. Désormais, les sondes portent un fluorochrome distinct pour chaque ddNTP. Elles sont détectées par quatre capteurs (un par longueur d'onde) après séparation par électrophorèse capillaire. Les quatre types de signaux sont ensuite analysés par un logiciel qui traduit le pictogramme obtenu en séquence. Il est ainsi possible d'obtenir la séquence d'un fragment comprenant jusqu'à 750 nucléotides.

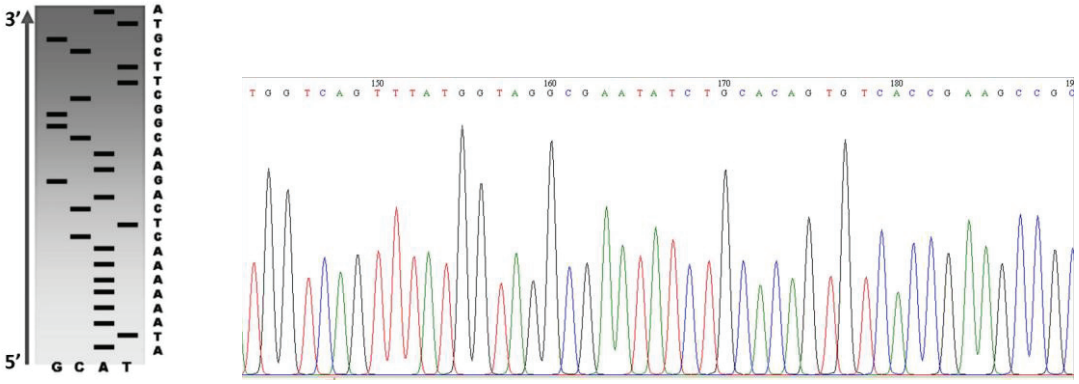


Figure 34: *À gauche, autoradiogramme d'un gel obtenu par séquençages manuel. A droite,* visualisation du pictogramme obtenu après séquençage automatisé d'un brin d'ADN.

Chapitre 2 : Techniques de biochimie

I. Surproduction des protéines de fusion

A. Culture bactérienne et induction de la surexpression

1. Milieux de culture

Les bactéries *E. coli* sont cultivées en milieu liquide LB Broth (Luria Bertani Broth, Sigma-Aldrich®), commercialisé sous forme de poudre et s'utilisant après dissolution dans de l'eau et autoclavage - composition finale : 10 g/l de tryptone (digestion pancréatique de caséine) ; 5 g/l d'extrait de levure ; 5 g/l de chlorure de sodium.

Les bactéries *E. coli* peuvent également se développer sur milieu solide (en boite de Pétri) constitué de LB Agar dans lequel le milieu est supplémenté de 15 g/l d'agar.

La croissance bactérienne en milieu liquide est suivie en mesurant l'absorbance de la culture à la longueur d'onde de 600 nm (DO_{600}).

2. Mise en culture et induction de la surexpression

À partir d'un étalement sur boite issu de la transformation de bactéries d'expression *E. coli* BL21(DE3) pLysS avec le vecteur d'intérêt, un clone est ensemencé dans une préculture contenant 5 ml de milieu LB en présence de l'antibiotique adéquat nécessaire à la sélection des cellules contenant le plasmide recombinant.

Lorsque le milieu de préculture est saturé en bactéries, il est dilué dans 300 ml de milieu de culture LB Broth supplémenté d'antibiotique adéquat pour maintenir la pression de sélection.

Selon la protéine d'intérêt, l'expression est induite au cours de la phase exponentielle de croissance bactérienne - l'ajout de l'IPTG se faisant à une concentration finale de 0,5 ou 1 mM à une DO_{600} = 0,6 ou 0,9. La température d'induction dépend également de la protéine d'intérêt et peut varier de 16°C à 25°C. Dans tous les cas, les cultures sont incubées sous agitation pendant au minimum 18 heures.

B. Lyse cellulaire

Cette étape varie selon la protéine étudiée. Dans tous les cas, les cultures bactériennes sont centrifugées à 6000 g pendant 10 minutes après quoi le surnageant est éliminé.

Chaque culot obtenu est resuspendu et homogénéisé dans son tampon de lyse adéquat. La lyse peut être aussitôt réalisée ou après une étape de congélation à -80°C pendant 24 heures.

La lyse cellulaire peut être conduite par des actions mécanique (sonication), enzymatique ou osmotique.

La lyse par sonication consiste à utiliser des ultrasons de forte puissance. Trois cycles d'une minute chacun sont réalisés en mode impulsionnel, entrecoupés par des temps d'incubation sur glace pour limiter l'échauffement et donc la dégradation des protéines libérées au cours de la lyse.

La lyse enzymatique consiste à utiliser un tampon adéquat dans lequel a été ajouté du lysozyme de blanc d'œuf (Sigma-Aldrich®) à hauteur de 10 mg/ml. Cette enzyme détruit la paroi bactérienne en catalysant l'hydrolyse du peptidoglycane la constituant. Le culot ainsi resuspendu est incubé sur glace au minimum pendant 2 heures.

La lyse osmotique nécessite, dans un premier temps, l'utilisation de saccharose dans un tampon adéquat à hauteur de 20% (w/v) pour resuspendre le culot (50 ml de tampon pour un litre de culture) avec un temps d'incubation de 10 minutes sur glace. Ceci provoque un choc osmotique et libère la fraction périplasmique bactérienne. Après centrifugation pendant 10 minutes à 7000 g à 15°C et récupération de cette fraction, le culot obtenu est resuspendu doucement avec de l'eau glacée et incubé sur glace pendant au minimum 10 minutes. Une fois centrifugé à 9000 g pendant 10 minutes, le surnageant est transféré dans un nouveau tube auquel est ajouté une solution de tampon concentrée 10 fois.

Une fois la lyse réalisée, une centrifugation à 10000 g pendant 30 minutes et à 4°C (permettant d'éliminer les débris membranaires dans le culot) suivie d'une filtration à 0,45 µm sont réalisées. Le surnageant contient, entre autres, la protéine d'intérêt surexprimée ; il est alors nécessaire de la purifier.

II. Purification des protéines d'intérêt par chromatographie

Un système chromatographique est constitué d'une phase stationnaire immobilisée (qui peut être solide, gélifiée, liquide ou un mélange solide/liquide) et d'une phase mobile (liquide ou gazeuse). Après avoir déposé le mélange de molécules à séparer, la phase mobile passe au travers ou à la surface de la phase stationnaire.

A. La chromatographie d'affinité avec métaux immobilisés (IMAC)

1. *Principe*

L'étiquette $(His)_6$ apportée lors du clonage confère à la protéine d'intérêt une affinité pour les cations divalents tels que le nickel. Cette propriété est utilisée pour extraire la protéine d'intérêt du reste du lysat bactérien obtenu après surexpression. En effet, les protéines endogènes des bactéries d'expression possèdent rarement une succession d'histidines dans leur séquence et auront donc une affinité réduite pour le nickel.

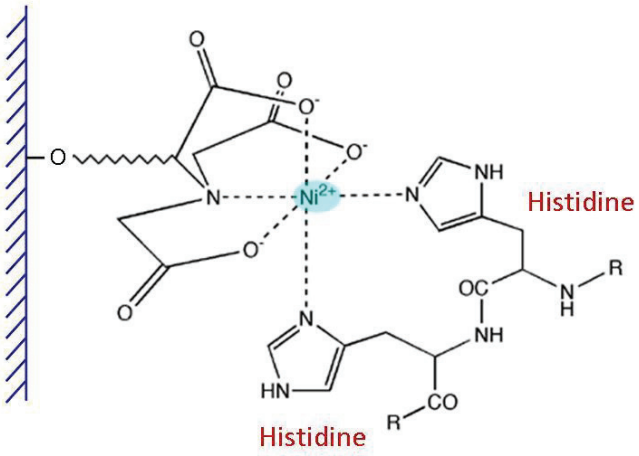


Figure 35 : Rétention des protéines ayant une étiquette poly-histidine sur le bras Ni-NTA fixé covalemment à la résine.

2. *Protocole*

Nous avons utilisé des colonnes composées de billes de sépharose ou d'agarose chélatées avec des ions Ni^{2+} par l'intermédiaire de bras espaceurs constitués d'acide nitriloacétique (HiTrap IMAC HP 1 ml - GE Healthcare ou Protino Ni-NTA 1 ml - Macherey-Nagel) montées sur un système FPLC Äkta *purifier* (GE Healthcare).

Lors de l'injection du lysat clarifié sur la colonne préalablement équilibrée avec le tampon A (Tableau 7) à un débit de 0,8 ml/min, la protéine surexprimée se lie aux ions Ni^{2+} grâce à son étiquette poly-histidine $(His)_6$. La colonne est ensuite lavée avec 10 ml de ce même tampon afin d'éliminer les protéines endogènes qui se sont fixées de façon non-spécifique.

L'élution est effectuée à 1 ml/min grâce à un gradient croissant d'imidazole - qui entre en compétition pour se lier aux ions Ni^{2+} - par mélange progressif des tampons A et B (Tableau 7) jusqu'à atteindre 100% de tampon B.

Protéine étudiée	TAMPON A	TAMPON B	TAMPON C
SHW	MES 100 mM pH 6,5	MES 100 mM pH 6,5 Imidazole 500 mM	MES 100 mM pH 6,5 NaCl 500 mM
1403	MOPS 100 mM pH 7	MOPS 100 mM pH 7 Imidazole 500 mM	-
11842	HEPES 100 mM pH 7,5 NaCl 50 mM	HEPES 100 mM pH 7,5 NaCl 50 mM Imidazole 500 mM	HEPES 100 mM pH 7,5 NaCl 500 mM

Tableau 7 : Composition des tampons utilisés lors des purifications des trois protéines étudiées.

Toutes les fractions sont ensuite analysées par SDS-PAGE afin de déterminer les fractions les plus riches en protéines d'intérêt et le degré de pureté de celles-ci.

B. La chromatographie échangeuse d'ions

1. Principe

Ce type de chromatographie se base sur la propriété qu'ont les protéines à posséder des charges globales différentes, en fonction du pH dans lequel elles se trouvent en solution. Une fois déposées sur une résine pouvant être cationique (appelée échangeuse d'anions) ou anionique (appelée échangeuse de cations), les protéines sont séparées en fonction de leur charge. Un gradient de chlorure de sodium permet alors d'éluer les différentes protéines : ces ions vont progressivement faire écran aux charges de la résine et ainsi entrainer un décrochage des protéines (Figure 36).

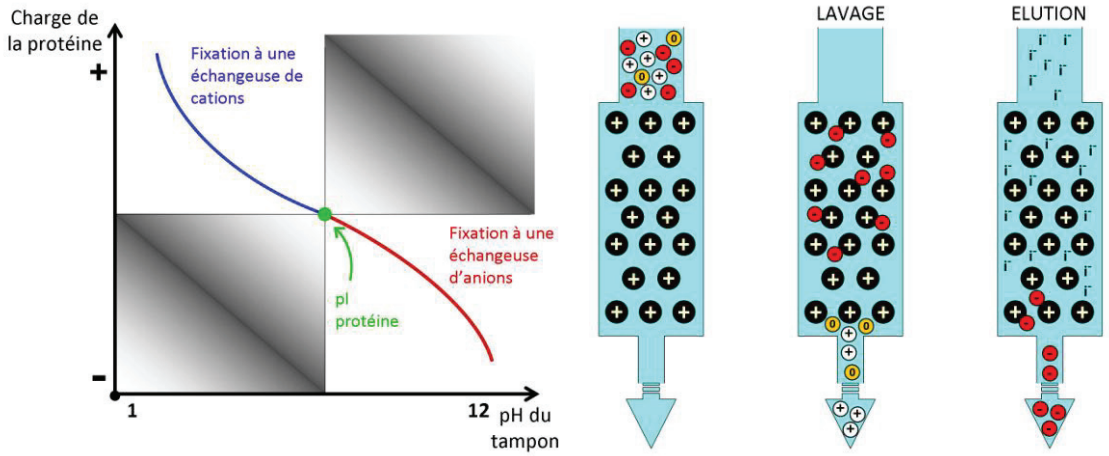


Figure 36 : Principe de la chromatographie échangeuse d'ions. *À gauche, schéma du choix de la résine en fonction de la charge de la protéine (pI : point isoélectrique de la protéine). À droite, représentation des étapes de ce type de chromatographie (exemple d'une échangeuse d'anions).*

2. *Protocole*

Cette technique a été utilisée en deuxième étape de purification lorsqu'il a été constaté que la protéine d'intérêt n'était pas suffisamment pure. Les fractions les plus riches en protéines d'intérêt issues de la chromatographie d'affinité ont été réunies avant d'être injectées. Nous avons utilisé une colonne Mono-Q™ 5/50 (GE Healthcare) montée sur un système FPLC Äkta *purifier* (GE Healthcare) puis équilibrée avec du tampon A (Tableau 7). Après l'injection à un débit constant de 1 ml/min, la colonne est lavée avec 3 volumes de tampon A avant de démarrer le gradient croissant de chlorure de sodium par un mélange progressif des tampons A et C (Tableau 7), jusqu'à atteindre 100% de ce dernier.

Toutes les fractions sont analysées par SDS-PAGE dans le but de déterminer les fractions les plus riches en protéines d'intérêt et le degré de pureté de celles-ci.

C. La chromatographie d'exclusion

1. *Principe*

Il s'agit d'une méthode de chromatographie en phase liquide permettant de séparer des macromolécules en fonction de leur masse moléculaire. Le principal phénomène physique permettant la séparation des différentes macromolécules n'est pas basé sur l'affinité chimique avec le support, mais idéalement sur la taille des macromolécules en solution. La colonne étant constituée de billes poreuses, les grosses molécules dont le diamètre est supérieur à celui des pores sont exclues et éluées en premier. Les molécules de plus petites tailles parcourent un trajet plus important au sein de la colonne du fait de leur passage dans les pores de tailles variables (Figure 37). En fonction de la taille des pores, il est possible de séparer les macromolécules entre elles (colonne de filtration sur gel) ou bien les protéines des petites molécules comme les sels (colonne de dessalage).

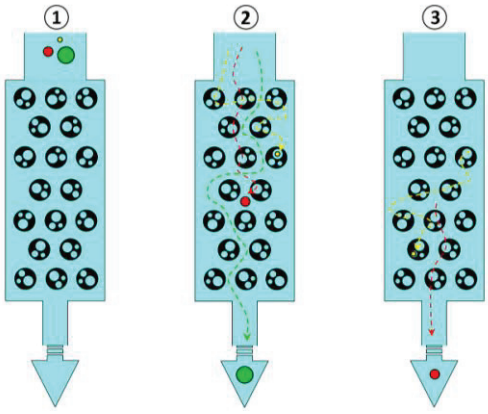


Figure 37 : Principe de la chromatographie d'exclusion. (1) Les protéines sont injectées en haut de la colonne et poussées par un débit constant. (2) Les grosses protéines ont un faible volume accessible et sont éluées rapidement. (3) Les protéines de plus petites tailles parcourent un plus grand volume de colonne et sont donc éluées plus tard.

2. *Protocole*

Dans le cas de séparation de protéines, nous utilisons une colonne analytique Superdex 200 10/300 GL (GE Healthcare) recommandée pour les protéines dont la masse moléculaire est comprise entre 10 et 600 kDa, montée sur le système FPLC Äkta *purifier* (GE Healthcare). La colonne est lavée avec 5 volumes d'eau puis équilibrée avec 5 volumes du tampon A de la protéine étudiée (Tableau 7). Le volume maximal injectable est 500 µl pour favoriser la résolution de séparation des protéines. L'échantillon est injecté sur la colonne à faible débit (0,5 ml/min) et élué par 1,5 volume de tampon A.

Dans le cas d'élimination des sels, nous utilisons une colonne de dessalage HiTrap Desalting HiPrep 26/10 (GE Healthcare) montée sur le système FPLC Äkta *purifier* (GE Healthcare). La colonne est lavée avec 3 volumes d'eau puis équilibrée avec 3 volumes du tampon A de la protéine étudiée (Tableau 7). Le volume maximal injectable d'échantillon est de 15 ml pour assurer la séparation entre les protéines et les sels présents dans le tampon initial. Dans notre étude, cette étape est réalisée pour éliminer l'imidazole des fractions retenues après l'étape de chromatographie d'affinité sur colonne de nickel. L'avantage de la colonne utilisée est de pouvoir travailler à débit important, l'échantillon étant injecté à 5 ml/min et élué par 1,5 volume de tampon A.

Toutes les fractions sont analysées par SDS-PAGE dans le but de déterminer les fractions les plus riches en protéines d'intérêt et le degré de pureté de celles-ci.

III. Révélation des protéines

A. Électrophorèse en gel de polyacrylamide en condition dénaturante

1. *Principe*

L'électrophorèse consiste en la migration de particules chargées sous l'influence d'un champ électrique.
Cette technique permet de séparer les protéines en fonction de leur poids moléculaire par migration au travers d'un gel de polyacrylamide en présence de Sodium Dodécyl Sulfate (SDS) et est appelée SDS-PAGE (SDS-PolyAcrylamide Gel Electrophoresis). Le SDS présent permet de charger négativement et uniformément les différentes protéines qui ne migreront au travers des mailles de polyacrylamide qu'en fonction de leur taille.

2. *Protocole*

Le gel est discontinu et constitué de deux parties : un gel de concentration et un gel de séparation qui différent par leurs pH et leurs concentrations en acrylamide (Tableau 8). Le but du premier gel est de concentrer les protéines qui sont déposées en faible quantité mais en volume conséquent. Au terme de cette première étape, les macromolécules sont massées en une zone très fine avant de migrer au travers du gel de séparation.

GEL DE SEPARATION					**GEL DE CONCENTRATION**	
	8%	10%	12%	15%		4%
Bis-Acrylamide 30%	2,6 ml	3,3 ml	4 ml	5 ml	Bis-Acrylamide 30%	660µl
Tris HCl 3 M pH 8,8	1,25 ml	1,25 ml	1,25 ml	1,25ml	Tris HCl 1 M pH 6,8	625 µl
H_2O	5,6 ml	5 ml	4,24 ml	3,24 ml	H_2O	3,48 ml
SDS 10%	400 µl	400 µl	400 µl	400 µl	SDS 10%	200 µl
APS 10%	100 µl	100 µl	100 µl	100 µl	APS 10%	500 µl
TEMED	10 µl	10 µl	10 µl	10 µl	TEMED	5 µl

Tableau 8 : *Composition des gels de séparation et de concentration*.

Les échantillons à déposer sont préalablement dénaturés en étant mélangés à un tampon de charge dénaturant concentré cinq fois (Tableau 9), puis chauffés pendant 5 minutes à 95°C. La migration est réalisée dans le tampon d'électrophorèse (Tris-base 25 mM, glycine 192 mM, SDS 0,1%) à 90 V pour le gel de concentration, puis à 180 V pour celui de séparation.

TAMPON DE CHARGE POUR GEL DENATURANT	
Tris HCl 1 M pH 6,8	0,25 M
SDS 10%	4%
Bleu de bromophénol	0,02%
β-mercaptoéthanol 14,2 M	0,6 M
Glycérol	40%
H_2O	q.s.p. 20 ml

Tableau 9 : *Composition du tampon de charge 5X pour gel dénaturant*.

B. Révélation des protéines

1. Révélation au bleu de Coomassie colloïdal

Cette méthode de coloration est indépendante de la composition en acides aminés, le bleu de Coomassie se liant aux acides aminés basiques et aromatiques. Une fois la migration des protéines achevée, le gel est démoulé puis rincé à l'eau distillée pendant 10 minutes. Celui-ci est ensuite plongé dans la solution PageBlue™ Protein Staining (Fermentas) pendant une heure à température ambiante ou légèrement chauffé aux micro-ondes pour une révélation plus rapide. Pour finir, le gel est rincé à l'eau distillée durant 10 à 20 minutes. Cette méthode permet de révéler, au minimum, 5 ng de protéine et est 10 fois plus sensible qu'une méthode dite classique au bleu de Coomassie.

2. Révélation au nitrate d'argent

Cette méthode de coloration est, elle aussi, indépendante de la composition en acides aminés.

A chaque étape, le gel est rincé deux fois à l'eau distillée. Une fois la migration des protéines achevée, le gel est démoulé, rincé à l'eau puis incubé 10 minutes dans une solution A (40% éthanol, 0,05% (v/v) formaldéhyde). Ensuite, le gel est incubé 10 minutes dans une solution B (0,2 g/l de $Na_2S_2O_3$ dissous dans de l'eau) puis 10 minutes dans une solution C (0,05 à 0,08 g de nitrate d'argent dans 50 ml d'eau). Pour finir, les bandes de protéines sont révélées dans 50 ml d'une solution D (30 g/l de $NaCO_3$, 4 mg/ml de $Na_2S_2O_3$, 0,05% (v/v) de formaldéhyde). La coloration est arrêtée en ajoutant 3 ml d'acide citrique 2 M. Cette méthode est aussi sensible que la méthode précédente.

3. Western-Blot

a. Principe

Le Western-blot (ou transfert de protéines) est une méthode permettant la détection et l'identification de protéines spécifiques à partir d'un échantillon.

Après migration sur gel dénaturant, les protéines sont transférées sur une membrane (typiquement en nitrocellulose) où elles sont exposées à un anticorps reconnaissant spécifiquement la protéine d'intérêt. Cet anticorps dit primaire est ensuite lui-même reconnu par un anticorps secondaire qui est couplé à une enzyme. La membrane est alors incubée dans un bain de substrat coloré soluble qui, au contact de l'enzyme, sera hydrolysé en un produit insoluble d'une autre couleur et teintera la membrane à l'endroit même où se trouve l'enzyme. Dans notre cas, l'anticorps utilisé reconnait l'étiquette $(His)_6$.

b. Protocole

Une fois la migration des protéines achevée, le gel est rincé à l'eau distillée puis mis en contact avec une membrane de nitrocellulose préalablement activée dans de l'eau puis du tampon de transfert (CAPS 10 mM pH 11, éthanol 10%). Les protéines sont transférées sur la membrane à 400 mA durant 45 minutes dans le tampon de transfert froid sous agitation et en présence d'un bloc réfrigérant.

Après transfert, la membrane est incubée une heure à température ambiante (ou une nuit à 4°C) dans une solution de saturation (PBS 1X, Tween 0,1%, lait en poudre écrémé 5% (w/v)). Les protéines de cette solution se lient à la membrane sur tous les sites non-occupés par la protéine cible. Ainsi, les anticorps ajoutés à l'étape suivante ne se fixeront que sur les sites de liaison de la protéine cible.

La membrane est ensuite mise en contact, pendant une heure à température ambiante, avec l'anticorps primaire monoclonal de souris anti-$(His)_6$ dilué (1/3000) dans la solution de saturation.

Rincée trois fois dans une solution de lavage (PBS 1X, Tween 0,1%) pour éliminer tous les anticorps primaires non fixés, la membrane est ensuite incubée en présence de l'anticorps secondaire couplé à une phosphatase alcaline dilué dans de la solution de lavage pendant une heure à température ambiante. L'anticorps secondaire est un anticorps de bouc reconnaissant la partie constante de l'anticorps de souris.

La révélation est réalisée après rinçage au tampon PBS (pour éliminer les anticorps non fixés) en utilisant une pastille de SIGMAFAST™BCIP®/NBT (5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium, Sigma-Aldrich®) dissoute dans 10 ml d'eau distillée.
La réaction est stoppée par rinçage à l'eau distillée.

IV. Digestion par la thrombine

L'étiquette $(His)_6$ peut être éliminée par une étape de protéolyse par la thrombine (Amersham Bioscience). La protéine est incubée à 4°C toute une nuit en présence de thrombine (3 unités de thrombine par mg de protéine à digérer), dans le tampon adapté à la protéine étudiée. La thrombine peut éventuellement être éliminée par une étape de chromatographie sur résine de Benzamidine Sépharose™, ce qui n'a pas été le cas dans cette étude car la quantité ajoutée est négligeable.

V. Analyse de l'état oligomérique des protéines

Il est possible de déterminer l'état oligomérique d'une protéine en utilisant différentes techniques. Les deux méthodes les plus simples sont la chromatographie d'exclusion et le pontage covalent par une molécule réactive analysé par SDS-PAGE.

Dans le premier cas, il est nécessaire de réaliser au préalable une calibration de la colonne d'exclusion de taille (Superdex 200 10/300 GL, GE Healthcare) grâce à des protéines « standard ». Ces protéines, de poids moléculaires connus, permettent d'établir une gamme de volume d'élution. Le volume d'élution est inversement proportionnel au logarithme du volume hydrodynamique moléculaire et donc au poids moléculaire pour une protéine considérée comme globulaire. Ainsi, en déterminant le volume d'élution de la protéine d'intérêt en solution et si sa masse moléculaire théorique est connue, il est aisé de déduire son poids moléculaire et donc d'en déduire l'état oligomérique.

La technique du pontage covalent (Phizicky & Fields, 1995) nécessite une molécule réagissant avec certains groupements des acides aminés de la protéine d'intérêt. Durant ces travaux, la molécule utilisée est le BS^3 (bis(sulfosuccinimidyl)suberate). Cette molécule contient une amine réactive N-hydroxysulfosuccinimide (NHS) ester de part et d'autre d'un bras espaceur de huit carbones réagissant avec les groupements amines primaires des chaines latérales des lysines et des extrémités N-terminales des polypeptides. Ainsi le BS^3 permet d'identifier les interactions inter-protéiques et même de figer celles qui sont faibles ou transitoires grâce au pontage intermoléculaire de lysines se trouvant à la surface des monomères distincts. Ceci n'est possible que s'il existe bel et bien un état oligomérique. La méthode utilisée a été l'approche en excès de 20:1 (BS^3:protéine) avec un temps de réaction de 45 et 60 min avant que la réaction ne soit stoppée avec 0,5 M de Tris pH 8,5. Les échantillons sont ensuite analysés par SDS-PAGE en présence d'un témoin négatif qui n'est autre que la protéine d'intérêt sans agent pontant.

Chapitre 3 : Études fonctionnelles

I. Principe :

Une enzyme est un catalyseur biologique qui accélère une réaction chimique. Le principe de l'étude fonctionnelle de ce biocatalyseur est de déterminer ses conditions optimales de réaction. En parallèle, il est nécessaire d'analyser les substrats sur lesquels elle est active pour ensuite définir ses différentes constantes catalytiques.

A. Modèle de Michaelis-Menten

Dans le cas d'une enzyme suivant le formalisme Michaelien, il est possible grâce à l'équation de Michaelis et Menten de déterminer les paramètres cinétiques propres à cette enzyme. Cette équation décrit la cinétique d'une réaction catalysée par une enzyme agissant sur un substrat unique pour donner un produit et relie la vitesse de réaction à la concentration de substrat et des constantes cinétiques. Le modèle comprend deux étapes, avec la formation d'une complexe Enzyme-Substrat (ES) et ensuite la libération du Produit (P) avec dissociation du complexe ES.
L'équation décrivant la vitesse de réaction enzymatique est la suivante :

$$V_i = \frac{V_{max} \cdot [S]}{K_m + [S]}$$

avec V_i la vitesse initiale de réaction pour une concentration en substrat [S] donnée, V_{max} la vitesse maximale de réaction et K_m la constante de Michaelis.

La V_{max} est atteinte lorsque la concentration en substrat est saturante. Le K_m est la concentration en substrat pour laquelle la vitesse de réaction est la moitié de la V_{max} et représente l'affinité de l'enzyme pour son substrat. Ces deux paramètres sont spécifiques d'un substrat donné et varient d'un substrat à un autre pour une même enzyme.
Graphiquement, l'équation de Michaelis-Menten représente une branche d'hyperbole (Figure 38).

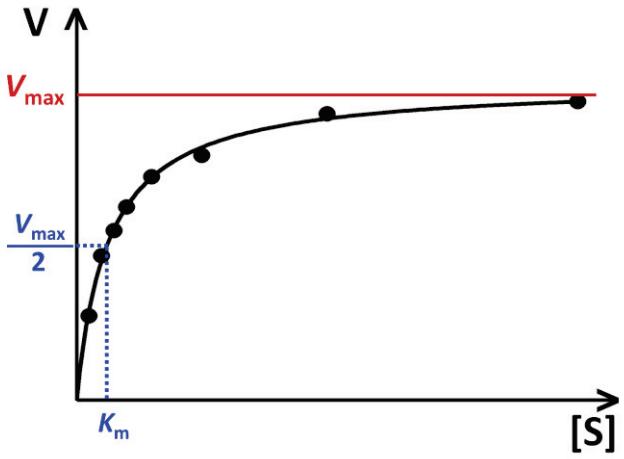


Figure 38 : Représentation de Michaelis-Menten montrant la vitesse initiale de réaction V_i (•) en fonction de la concentration en substrat [S]. Avec V_{max} la vitesse maximale de réaction et K_m la constante de Michaelis.

Expérimentalement, il n'est pas possible de déterminer exactement la vitesse maximale de réaction, seule son approximation est possible.

Il existe plusieurs méthodes pour déterminer les paramètres de Michaelis-Menten d'une enzyme. Dans la pratique, on réalise une série de mesures de la vitesse initiale de la réaction pour différentes valeurs de la concentration de substrat [S]. Afin de déterminer numériquement les paramètres K_m et V_{max} avec une plus grande précision, il est possible de linéariser l'équation de Michaelis-Menten.

Parmi les linéarisations les plus courantes, on peut citer les représentations de Lineweaver-Burk (ou double inverse) (Lineweaver & Burk, 1934), de Eadie-Hofstee (Hofstee, 1959) ou encore de Hanes-Woolf (Hanes, 1932) (Tableau 10).

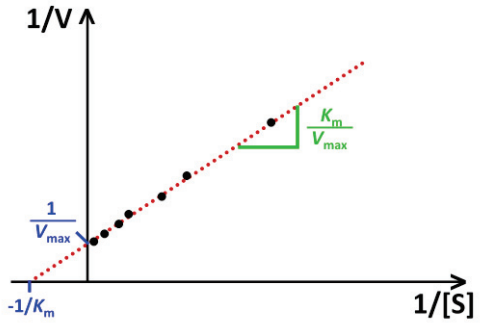
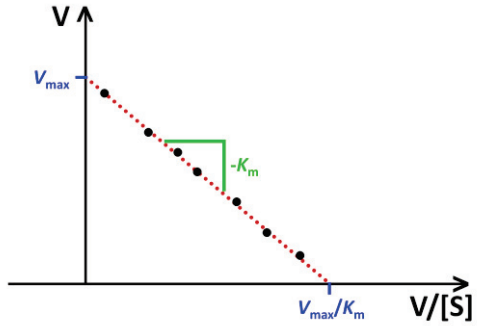
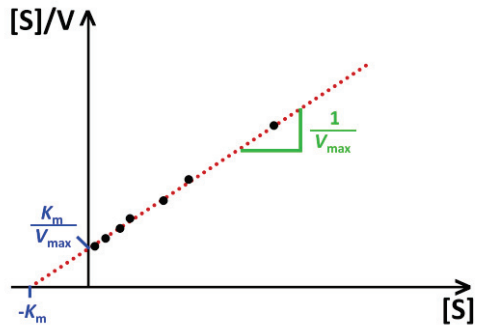
Représentation	Equation	Graphique
Lineweaver-Burk	$\frac{1}{V_i} = \frac{K_m}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}}$	
Eadie-Hofstee	$V_i = -K_m \cdot \frac{V_i}{[S]} + V_{max}$	
Hanes-Woolf	$\frac{[S]}{V_i} = \frac{[S]}{V_{max}} + \frac{K_m}{V_{max}}$	

Tableau 10 : Représentations des différentes formes de linéarisation les plus courantes de l'équation de Michaelis-Menten.

Dans le but de comparer différentes expériences tout en s'affranchissant de la quantité d'enzyme utilisée lors des cinétiques, on divisera la vitesse maximale de réaction par la concentration molaire enzymatique du milieu réactionnel. Ceci donne la constante catalytique k_{cat}. Pour faire une étude comparative de substrats, il est préférable de calculer une efficacité catalytique k_{cat}/K_m, également appelée constante de spécificité.

Désormais, ces techniques peuvent être remplacées par la régression non linéaire. Cette dernière, beaucoup plus précise et robuste (Tommasini *et al.*, 1985), est réalisée grâce au logiciel GraphPad Prism (version 6.0, GraphPad Software, La Jolla California USA, www.graphpad.com) ou au serveur http://zunzun.com.

B. Inhibition de l'activité enzymatique

Il est également possible de déterminer une constante d'inhibition pour un inhibiteur d'une enzyme donnée. Il suffit de procéder à des séries de mesures de vitesses initiales de réaction (comme précédemment) à différentes concentrations de substrats, avec plusieurs concentrations d'inhibiteur (ou ligand). Il existe différents types d'inhibition pouvant faire varier la V_{max}, le K_m ou les deux en même temps. Les données sont également traitées par régression non linéaire en utilisant le logiciel GraphPad Prism.

II. Mise en application

L'activité enzymatique des α-glucosidases peut être étudiée en utilisant différentes méthodes qui sont complémentaires.

A. Hydrolyse du *para*-nitrophényl-α-D-glucopyranoside (*p*NPG)

Cette technique consiste à utiliser un hétéroside de type glucoside constitué d'une molécule de glucose liée covalemment à un *para*-nitrophénol (*p*NP). Le *p*NPG est un substrat incolore non naturel des α-glucosidases mimant une liaison glucosidique de type α-(1,4) qui, lorsqu'elle est hydrolysée, permet de libérer le *para*-nitrophénol - molécule absorbant à une longueur d'onde de 410-420 nm (jaune) (Figure 39).

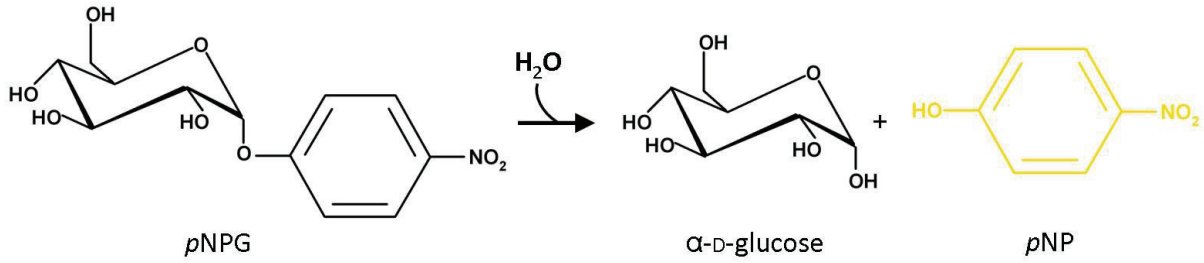


Figure 39 : Schéma de la réaction d'hydrolyse du *p*NPG en présence d'α-glucosidase.

1. Détermination des conditions optimales de réaction

Ces expériences ont été menées dans le but de déterminer la température et le pH optimaux, les effets des ions sur l'enzyme ainsi que les limites de stabilité enzymatique.

Une large gamme de pH a été couverte de 3,5 à 9 en utilisant différents tampons (sodium acétate 100 mM, pH 3,5 à 4,5 ; MES 100 mM, pH 5 à 6,5 ; MOPS 100 mM, pH 7 à 8,5 ; glycine 100 mM, pH 9) et les températures testées sont comprises entre 4°C et 65°C.

Le milieu réactionnel contient 1,25 mM de *p*NPG, 100 µl d'extrait enzymatique complété avec le tampon approprié pour un volume final de 1 ml et est ensuite incubé à la

température voulue pendant un temps déterminé en fonction de l'enzyme étudiée. La réaction est stoppée par adjonction d'une solution de carbonate de sodium à une concentration finale de 0,7 M.

Pour chaque essai enzymatique, un témoin négatif a été réalisé en parallèle dans les mêmes conditions à la différence que le substrat et l'enzyme ont été incubés séparément.

Dans le cadre des mesures de stabilité, l'enzyme a été préalablement incubée à la température et au pH appropriés.

Les échantillons sont ensuite dosés « au point final » par spectrophotométrie et la valeur la plus grande correspond à 100% d'activité.

Toutes les mesures ont été réalisées trois fois avec le même lot d'enzyme pour s'assurer de l'homogénéité des résultats.

2. Détermination des paramètres cinétiques de l'α-glucosidase de Shewanella sp.ANA-3 pour le pNPG et les ligands

Dans ce cadre, des mesures d'absorption en temps réel ont été réalisées avec un spectrophotomètre Helios Gamma (ThermoFisher) possédant une cuve thermostatée et piloté par le logiciel VISION Lite (version 2.2, ThermoFisher). Il a ainsi été possible de mesurer la vitesse initiale de réaction grâce à l'absorbance à 420 nm sur une période de 4 minutes.

Pour l'étude du *p*NPG seul, les mesures ont été réalisées dans une cuve contenant entre 0,01 mM et 2 mM de *p*NPG et 0,14 µM d'enzyme pour un volume final de 700 µl dans les conditions optimales (100 mM MES pH 6,5 ; 1 mM Ca^{2+}, à 35°C).

Dans le cas des ligands, les vitesses initiales de réaction ont été mesurées en présence de plusieurs concentrations de ces derniers (acarbose : 0,01 mM, 0,05 mM, 0,075 mM ; castanospermine : 0,05 mM, 0,1 mM, 0,15 mM ; déoxynojirimycine : 0,025 mM, 0,05 mM, 0,1 mM ; glucose : 1 mM, 5 mM, 10 mM) sur une gamme de concentrations de *p*NPG allant de 0,3 mM à 2,5 mM en présence de 0,14 µM d'enzyme pour un volume final de 700 µl.

En pratique, l'enzyme est ajoutée extemporanément dans la cuve au moment du lancement de la mesure de vitesse initiale.

3. Détermination des paramètres cinétiques de l'α-glucosidase de Shewanella sp.ANA-3 pour les substrats naturels

Dans le cadre des substrats naturels, nous avons utilisé selon les instructions du fournisseur un kit de réactifs permettant de doser le glucose libéré lors de la réaction enzymatique (Glucose Hexokinase FS, DiaSys).

Les mesures ont été effectuées sur des points de prélèvements de réactions enzymatiques réalisées avec les différents substrats (entre 0,3 mM et 10 mM de substrat) et 0,14 µM d'enzyme pour un volume final de 50 µl dans les conditions optimales (100 mM MES pH 6,5 ; 1 mM Ca^{2+}, à 35°C). Les prélèvements ont été réalisés toutes les 30 secondes à partir du lancement de la réaction par ajout de l'enzyme et la réaction a été stoppée par une incubation à 95°C pendant 5 minutes.

B. Criblage des substrats de l'enzyme par chromatographie sur couche mince

1. Principe

Le criblage a été réalisé en utilisant la méthode de chromatographie sur couche mince (CCM).

Il s'agit d'une technique de chromatographie partageant plusieurs principes théoriques des chromatographies sur colonne bien que le support soit différent.

Dans cette technique, la phase de séparation est solide et attachée à une matrice adaptée revêtant en couche mince une plaque de métal.

Le mélange à analyser est appliqué en « spots » près d'un bord de la plaque qui sera ensuite placée dans un réservoir de phase mobile.

La phase mobile est liquide et diffuse linéairement le long de la plaque par capillarité provoquant ainsi la migration des différents composés selon leurs caractéristiques. Ce type de chromatographie a un avantage pratique majeur par rapport aux autres chromatographies : plusieurs échantillons peuvent être analysés simultanément. Comme la résistance à l'écoulement est faible, la phase mobile migre rapidement à travers la couche.

Dans notre cas, la phase stationnaire est constituée de silice. Il s'agit d'un matériau polymérique produit à partir d'acide orthosilicique. Les groupements silanol (Si-OH) présents en grand nombre rendent cette matrice hydrophile.

2. *Protocole*

Les réactions enzymatiques ont été réalisées sur différents oligosaccharides dans les conditions optimales de réaction de l'enzyme étudiée. Au temps de réaction considéré comme final, les échantillons sont déposés en « spots » sur une même ligne de dépôt d'une plaque de silice (TLC Aluminium oxide 60 F_{254} neutral, Merck). Cette dernière est ensuite déposée dans une cuve saturée en solvant de migration (1-propanol-eau-éthyl acétate (6:3:1, v/v/v)). La migration est effectuée deux fois pour améliorer la séparation des mono- et oligosaccharides. Après séchage de la plaque, la révélation s'effectue par projection d'une solution de révélation (diphénylamine 1% (w/v) et 1% aniline (v/v) dans de l'acétone et ajout de 0,1% (v/v) d'acide phosphorique 85%) puis chauffage à 160°C pendant 5 minutes (Angelov *et al.*, 2006).

C. Détermination de l'anomérie de produit libéré

1. *Principe*

Cette détermination a été réalisée par RMN (Résonance Magnétique Nucléaire) du proton. L'échantillon est placé dans un champ magnétique statique intense (B0). Les molécules de cet échantillon sont soumises à un second champ magnétique B1 (orthogonal à B0) qui provoque une perturbation des noyaux considérés. L'application de ce champ de radiofréquence est de courte durée (impulsion de quelques microsecondes). La perturbation de l'aimantation de ces noyaux est captée par une bobine réceptrice : c'est le signal RMN. Ces données sont analysées et transformées en signal. Le traitement mathématique (transformée de Fourier) permet de transformer le signal en spectre RMN. Chaque signal RMN est caractérisé par plusieurs grandeurs caractéristiques de chaque noyau considéré. Ces paramètres donnent des informations sur l'atome en question et son environnement. Dans notre cas, nous nous intéressons à l'anomérie du carbone C1 du glucose libéré après la réaction enzymatique en vue de déterminer le mécanisme d'hydrolyse utilisé par l'enzyme.

2. *Protocole*

Des spectres RMN des protons ont été acquis sur un spectromètre Bruker 500 MHz à 25°C.

Le signal anomérique du proton lié au C1 du glucose possédant une fréquence proche de celle de l'eau (solvant), il a été nécessaire d'utiliser de l'eau deutérée (D_2O) pour éliminer au maximum le signal de l'eau, le deutérium (D) n'étant pas visible en RMN du proton.

Les échantillons protéiques ont donc été lyophilisés puis solubilisés dans du D_2O avant l'expérience. Le *p*NPG, le tampon et le chlorure de calcium ont été également

préparés en eau deutérée. Les essais ont été réalisés dans du tampon MES 100 mM pH 6,5 supplémenté de 1 mM de chlorure de calcium et avec une concentration finale de *p*NPG à 1 mM. L'hydrolyse du substrat est initiée par ajout de l'enzyme à une concentration finale d'environ 50 nM. Les spectres (64 scans et 1,5 secondes de relaxation pure entre chaque scan pour une relaxation totale de 3,7 secondes par scan) ont été acquis à différents intervalles. Des spectres de contrôle sur le tampon en présence de glucose 1 mM en milieu deutéré ont également été enregistrés.

III. Immobilisation d'enzyme

L'immobilisation d'enzyme nécessite de mélanger la suspension bactérienne à un même volume de solution d'alginate (6% (w/v) dans le tampon adéquat). Cette dernière a été préalablement autoclavée pour permettre sa solubilisation complète. Ce mélange est injecté au goutte-à-goutte dans une solution de chlorure de calcium 1 M provoquant ainsi la formation de billes d'alginate polymérisé contenant l'enzyme (Figure 40). Les billes sont ensuite rincées dans le tampon adéquat afin d'éliminer l'excédent de chlorure de calcium.

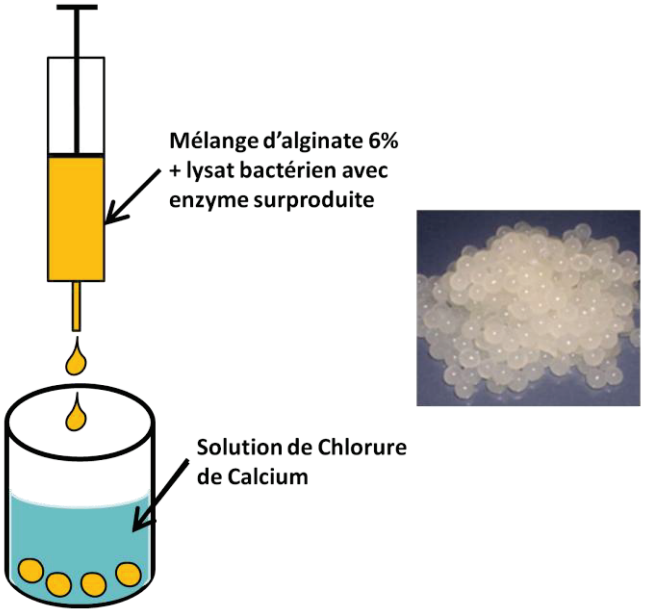


Figure 40 : Schéma de l'immobilisation d'enzyme dans des billes d'alginate et photo de ces billes.

Chapitre 4 : Études structurales

I. Dichroïsme circulaire

A. Principe

Le dichroïsme circulaire (CD) est une spectroscopie d'absorption dans l'UV proche et lointain qui donne des informations quantitatives sur les éléments de structures secondaires des macromolécules biologiques en solution. Un échantillon possédant un chromophore chiral ou placé dans un environnement asymétrique, absorbe différemment une onde polarisée circulaire gauche et une onde polarisée circulaire droite. Le dichroïsme circulaire permet de mesurer cette différence. L'évaluation de la structure secondaire est la principale utilisation du CD dans l'étude des protéines. Le chromophore dans ce cas est le groupement amide de la liaison peptidique. Il est placé entre deux carbones asymétriques; les carbones α des deux acides aminés voisins. Les groupements peptidiques (ou plans peptidiques) sont orientés suivant des angles phi et psi caractéristiques des différents types de structures secondaires. L'allure des spectres de CD varie donc suivant le type de repliement en hélice α, feuillet β, coude et apériodique (« random »). Des méthodes de déconvolution des spectres de CD permettent de quantifier ces différents éléments de structure secondaire. Les résultats obtenus après un traitement des spectres (qui nécessite plusieurs étapes) sont jugés fiables lorsque l'analyse par les trois programmes « SELCON », « CONTIN » » et « CDSSTR », disponible *via* le serveur DichroWeb (http://dichroweb.cryst.bbk.ac.uk) sont similaires (Sreerama & Woody, 2000). L'interprétation des résultats nécessite de connaitre la concentration exacte de l'échantillon analysé et la masse moléculaire moyenne des résidus composant la protéine étudiée.

Une des difficultés de cette technique réside dans le fait que les tampons « classiques » absorbent souvent aux UV lointains. Ainsi, les tampons de phosphates, de sulfates (HEPES, MOPS), de carbonates et d'acétates sont difficilement utilisables. Pour limiter ces absorptions parasites il est courant d'utiliser des cuves très fines.

B. Protocole

Les expériences ont été menées avec le dichrographe Chirascan (Applied Photophysics Ltd) qui est équipé d'un porte échantillon polyvalent pour cuves cylindriques ou cuves droites avec régulation de température par effet Peltier (déplacement de chaleur en présence d'un courant électrique). Les mesures ont été réalisées dans une gamme de

longueurs d'onde de 260 à 180 nm avec des pas de 0,2 nm avec des cuves de 0,1 mm, à 10°C. Les concentrations protéiques finales utilisées ont varié de 0,114 à 0,235 mg/ml.

II. Cristallogenèse

A. Principe

La cristallogenèse est la formation d'un cristal, correspondant au passage d'une molécule d'un état désordonné liquide à un état ordonné solide avec un motif périodique répété dans les trois dimensions de l'espace. Dans le cas des macromolécules comme les protéines, ceci est obtenu par une augmentation progressive de leur concentration. La nucléation (formation de germes cristallins) correspond à l'apparition d'une phase cristalline stable à partir d'une solution sursaturée (Figure 41).

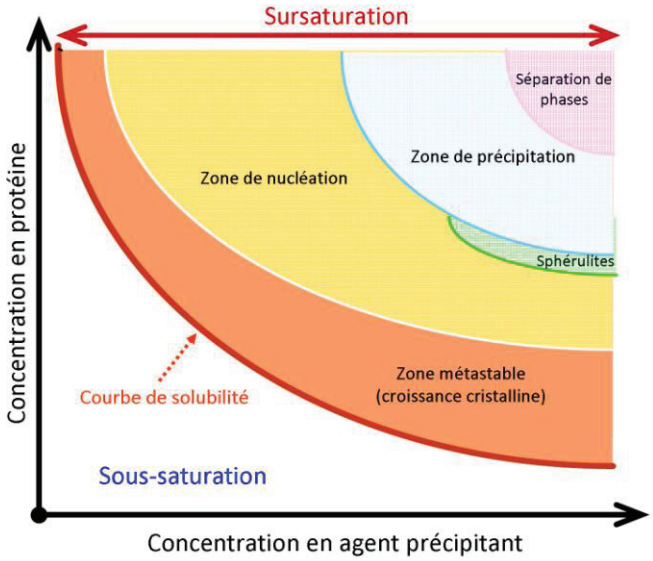


Figure 41 : Diagramme de solubilité suivant la concentration en protéine et en agent précipitant. La courbe de solubilité sépare la zone où la protéine est sous-saturée en solution des zones de sursaturation (métastable, nucléation ou précipitation) - adapté de *cristech.cnrs.fr/fenetre/doc_pdf/Madeleine_Ries.pdf*.

B. Méthode

La méthode la plus couramment utilisée est la diffusion en phase vapeur. Elle a été décrite pour la première fois en 1973 (Reid *et al.*, 1973). Une solution d'agent précipitant est déposée dans réservoir et sous forme de goutte dans un système hermétiquement scellé. La protéine est ajoutée à la goutte de solution de cristallisation, entrainant une dilution de cette dernière. Des échanges de vapeur d'eau ont alors lieu jusqu'à atteindre un équilibre thermodynamique. Du fait de la déplétion en eau, il s'ensuit une augmentation progressive de la concentration en protéine conduisant à un état de saturation. Les agents précipitants

les plus couramment utilisés peuvent être de différentes natures comme des polymères (les PEGs), des solvants (MPD : 2-méthyl-2,4-pentanediol) ou encore des sels (sulfate d'ammonium ou chlorure de sodium).

Il est possible de réaliser des gouttes suspendues ou assises dans le but d'obtenir des cristaux de nos protéines d'intérêt (Figure 42). Pour les expériences de cristallisation décrites dans ces travaux, les gouttes suspendues sont réalisées sur lamelles en utilisant des boites Linbro 24 puits (Hampton Research) et les gouttes assises à l'aide d'un robot de cristallisation Mosquito® (TTP Labtech) en utilisant des boites 96 puits TTP Labtech 3 cupules ou Greiner Low Profile.

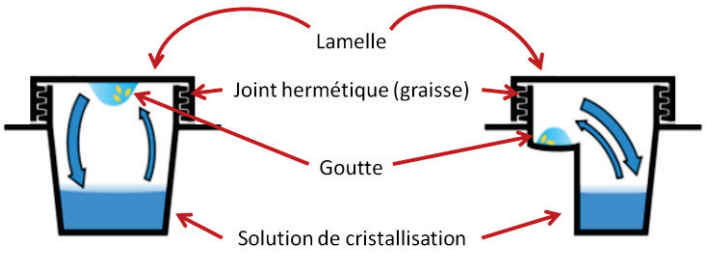


Figure 42 : Schéma du principe de la diffusion en phase vapeur, selon les techniques de la goutte suspendue (à gauche) ou de la goutte assise (à droite).

C. Trempage et co-cristallisation

Dans le but d'obtenir des cristaux de protéine en complexe avec un ligand, deux méthodes peuvent être mises en œuvre : le trempage et la co-cristallisation.

1. Trempage

Cette technique consiste à introduire par diffusion la molécule d'intérêt dans le cristal de protéine déjà formé. Le ligand se propage par les canaux de solvants présents dans l'empilement cristallin. Si les sites de fixation sont accessibles, la molécule diffusante peut alors les atteindre et s'y fixer - ceci dépend également de la taille du ligand. Cette méthode consiste à immerger le cristal dans la solution de cristallisation supplémentée de la substance choisie. La concentration de cette substance et le temps de trempage sont déterminés empiriquement. Le risque potentiel de cette technique est d'endommager le cristal sous l'effet de la diffusion de la molécule utilisée. Une diminution de temps de trempage ou de la concentration de la molécule permet souvent de minimiser ce risque.

2. Co-cristallisation

Il est également possible d'obtenir des complexes protéine-ligand par co-cristallisation. La (ou les) molécules ligand(s) sont ajoutées à la solution de protéine avant la formation de cristaux. En effet, il arrive que le ligand, ajouté préalablement à la solution protéique, facilite l'apparition de cristaux en stabilisant la protéine d'intérêt. Cette technique est également utilisée lorsque le ligand est trop volumineux pour diffuser dans les canaux de solvant d'un cristal déjà formé. Il est possible cependant que les conditions de cristallisation diffèrent de celles de la protéine seule.

D. Ensemencement

Lorsque la nucléation n'est pas optimale, il est possible d'utiliser la technique d'ensemencement. On peut distinguer deux méthodes : le micro-ensemencement (« micro-seeding ») et le macro-ensemencement (« macro-seeding »).

Le micro-seeding consiste à broyer un cristal, à prélever des points de nucléation grâce, par exemple, à une moustache de chat et à les placer dans une goutte pré-équilibrée.

Le macro-seeding quant à lui, consiste à pêcher un cristal et le placer dans une goutte pré-équilibrée dans le but d'obtenir un cristal de plus grande taille. Au préalable, le cristal aura été passé dans une solution de cristallisation légèrement moins concentrée.

E. Cryoprotection des cristaux

Lors d'une expérience de diffraction, un cristal de protéine est soumis à un faisceau intense de rayons X. Généralement, cette irradiation entraine une dégradation irréversible du cristal (Garman, 1999) causée par la libération de radicaux libres. Pour minimiser ce problème, on utilise des conditions cryogéniques : le cristal est préparé et gelé instantanément dans de l'azote liquide (avoisinant -196°C). Cette étape, appelée « flash cooling », permet de transformer le solvant présent à l'intérieur (dans les canaux de solvant) et à l'extérieur du cristal en glace amorphe. Cette technique possède un grand nombre d'avantages (Garman & Schneider, 1997; Garman, 1999) mais nécessite de « cryoprotéger » les cristaux afin d'éviter la formation de microcristaux de glace qui endommageraient le cristal de protéine et feraient apparaitre des tâches de diffraction parasites.

Cette cryoprotection peut s'effectuer en utilisant des concentrations de 5 à 30% de glycérol, d'éthylène glycol ou encore d'autres composés comme des sucres. Le temps et la composition de la cryoprotection est à déterminer empiriquement.

F. Déshydratation

Il a été montré qu'il est possible d'améliorer la qualité de diffraction d'un cristal en le soumettant à une étape préalable de déshydratation (Heras *et al.*, 2003). Cette technique consiste à transférer le cristal dans une solution de déshydratation (solution de cristallisation supplémenté d'un agent déshydratant comme les PEG, le glycérol ou des sels). Il est également possible d'agir sur la déshydratation des cristaux de façon précise par l'intermédiaire d'un dispositif de contrôle d'humidité (HC1) (Sanchez-Weatherby *et al.*, 2009) implémenté sur les lignes synchrotron et, par exemple, retrouvé sur la ligne BM14 de l'ESRF.

III. Cristallographie rayons X

A. Obtention et enregistrement des données de diffraction

1. Le principe de diffraction

a. Facteur de structure et densité électronique

Un atome irradié par des rayons X se comporte comme une source secondaire : il diffuse le rayonnement X. Cette diffusion résulte de l'interaction des rayons X avec les électrons des atomes. Si l'atome en question appartient à un cristal, les ondes diffusées interagissent entre elles pour créer des interférences. Les rayons X sont alors émis dans des directions privilégiées : c'est le phénomène de diffraction. L'onde diffractée par un cristal est décrite par son facteur de structure F(hkl) où h, k, l sont les indices du réseau réciproque du cristal, encore appelés indices de Miller. F(hkl) est un nombre complexe, qui est la somme des facteurs de diffusion atomique des atomes de la maille cristalline en prenant en compte les décalages de phase des amplitudes diffusées par les différents atomes :

$$F(hkl) = |F(hkl)|.e^{i\varphi_{hkl}} \quad \text{ou} \quad F(hkl) = \sum_{j=1}^{j=n} f_j . e^{2\pi i(hx_j + ky_j + lz_j)}$$

$|F(hkl)|$ est le module et ϕ_{hkl} la phase du facteur de structure pour la réflexion hkl. f_j représente le facteur de diffusion de l'atome j de coordonnées (x_j, y_j, z_j) dans le cristal et dépend de son nombre d'électrons.

Si l'on arrive à mesurer expérimentalement les facteurs de structure d'un cristal, il est possible de calculer sa fonction densité électronique (répartition des électrons dans le cristal par unité de volume). En effet la densité électronique $\rho(x,y,z)$, en un point quelconque de

coordonnées (x,y,z) de ce cristal, est la transformée de Fourier inverse du facteur de structure :

$$\rho(x,y,z) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} F(hkl) \cdot e^{-2\pi i(hx+ky+lz)}$$ où V est le volume du cristal.

b. Intensité diffractée

Pour mesurer l'intégralité des réflexions, nous avons utilisé la méthode du cristal tournant et imposé au cristal une série de rotations de faibles amplitudes, autour d'un axe perpendiculaire à celui du faisceau de rayons X monochromatique incident.
Pour une réflexion hkl donnée, l'intensité I(hkl) des ondes diffractées par un cristal qui tourne à une vitesse angulaire ω dans un faisceau de rayons X d'intensité I_0 et de longueur d'onde λ est donnée par la relation :

$$I(hkl) \propto (I_0/\omega) \times \lambda^3 \times (V_c/V_m^2) \times |F(hkl)|^2 \times P(hkl)L(hkl)A(hkl)$$

avec P(hkl) est le facteur de polarisation correspondant à l'angle de Bragg θ de la réflexion d'indice hkl ; L(hkl) est le facteur de Lorentz ou facteur cinématique ; A(hkl) le facteur d'absorption ; V_c le volume irradié du cristal ; V_m le volume de la maille élémentaire du cristal.

I(hkl), quantité mesurée expérimentalement par un détecteur, est donc proportionnelle au carré du module du facteur de structure $|F(hkl)|^2$. Ainsi, seules les amplitudes des ondes diffusées sont accessibles à l'expérience car les ondes diffractées par un cristal ne peuvent être refocalisées et les phases sont perdues. La détermination de la phase ϕ_{hkl} de chaque réflexion constitue donc l'une des difficultés majeures de la cristallographie. Nous verrons plus loin que trois techniques peuvent être utilisées pour la détermination des phases en cristallographie des macromolécules biologiques.

2. Enregistrement des données

Au cours de cette étude, nous avons fait des mesures de diffraction soit au laboratoire, soit sur les lignes de synchrotrons FIP BM30A, ID14-2, ID23-1, ID23-2 et ID29 de l'ESRF à Grenoble et PROXIMA 1 de SOLEIL à Gif-sur-Yvette.

a. Le générateur de rayons X à anode tournante

L'anode tournante du laboratoire (Nonius *FR* 591) est équipée d'une cible en cuivre, dont la longueur d'onde d'émission est de 1,5418 Å (CuK_α). Le cristal est monté sur une tête goniométrique, la distance entre le détecteur et le cristal est choisie de manière à séparer correctement les tâches de diffraction. Le cristal est irradié par les rayons X et les intensités des ondes résultantes sont mesurées par un détecteur bidimensionnel de type « Image Plate » (Marresearch MAR345), numérisées et enregistrées.

b. Le synchrotron ESRF

La **ligne FIP BM30A** (« French beamline for Investigation of Proteins ») est spécialement dédiée à la cristallographie des macromolécules biologiques. Elle peut être utilisée pour des expériences de diffraction à longueurs d'onde fixes ou variables (dispersion anomale, MAD). L'optique de la ligne délivre un faisceau d'une taille de 300x300 µm² avec un large éventail d'énergies accessibles (de 8 à 18 keV) correspondant à des longueurs d'onde comprises entre 0,69 et 1,55 Å. La ligne possède un détecteur de type CCD (« Charge-Coupled Device ») ADSC Q315r, un diffractomètre avec un goniomètre mini-Kappa, un passeur d'échantillons/goniomètre robotisé (NatX-ray G-Rob) et un équipement cryogénique complet (OXFORD system). Il est également possible d'y effectuer des mesures de fluorescence (détecteur de fluorescence Roentec Multi-Channel Analyser).

La **ligne ID14** est dédiée à la cristallographie macromoléculaire. Cette ligne possède quatre sous-stations indépendantes (de ID14-1 à ID14-4). ID14-1 et ID14-2 délivrent des longueurs d'onde fixes, respectivement de 0,934 Å (13,274 keV) et 0,933 Å (13,289 keV) ce qui correspond à des résolutions limites de 0,94 Å et 1,00 Å. ID14-4, quant à elle, permet d'avoir un faisceau de longueur d'onde variable (0,86 à 1,3 Å soit 14,5 à 9,6 keV) ce qui autorise ainsi les expériences SAD (« Single wavelength Anomalous Diffraction ») ou MAD (« Multiple-wavelength Anomalous Diffraction »). Sa limite de résolution est de 0,88 Å. ID14-1 est équipée d'un détecteur CCD ADSC Q210, ID14-2 d'un CCD ADSC Q4 et ID14-4 d'un ADSC CCD Q315r. Un système cryogénique complet y est également installé. Enfin, notons qu'ID14-3, initialement dédiée aux expériences de SAXS (« Small Angle X-ray Scattering »), a été relocalisée sur la ligne BM29 dans le courant de l'année 2012. La taille du faisceau de ID14-1 est de 50x200 µm² et il est possible d'enregistrer 25 images par minute avec un temps d'exposition d'une seconde.

La ligne **ID23** est également dédiée aux expériences de cristallographie macromoléculaire et possède deux sous-stations indépendantes (ID23-1 et ID23-2). ID23-1 permet d'avoir un faisceau de 30x40 µm² et de longueur d'onde variable (autorisant les expériences MAD) de 5 à 20 keV soit de 0,62 à 2,48 Å. ID23-2, quant à elle, a une énergie fixe de 14,2 keV correspondant à 0,873 Å permettant les enregistrements de données anomales avec une taille de faisceau de moins de 10x10 µm². Depuis Octobre 2012, ID23-1 est équipée d'un détecteur Pilatus 6M-F et peut enregistrer 25 images par seconde. Cette sous-station possède également un détecteur « Rontec Silicon drift diode » permettant d'effectuer des mesures de fluorescence. ID23-2 est équipée d'un détecteur CCD MarMosaic 225 et les deux sous-stations possèdent un équipement cryogénique complet (OXFORD system).

La ligne **ID29** est prévue pour les expériences de MAD et de SAD sur une large gamme d'énergie de 6 à 20 keV équivalant à des longueurs d'onde de 0,62 à 2 Å. La taille du faisceau de cette ligne dédiée à la cristallographie macromoléculaire est de 10x10 à 75x30 µm². ID29 est équipée d'un détecteur Pilatus 6M-F permettant de collecter 12 images par seconde ainsi que d'un équipement cryogénique complet, d'un dispositif de déshydratation et d'un détecteur de fluorescence Rontec.

c. Le synchrotron SOLEIL

La **ligne PROXIMA 1** est l'une des deux lignes dédiées à la cristallographie macromoléculaire à SOLEIL (l'autre étant encore en construction). Cette ligne délivre un faisceau intense parallèle et ajustable pour des expériences de diffraction de haute résolution ou sur des cristaux ayant de grands paramètres de maille. Elle peut être utilisée pour des expériences de diffraction à longueurs d'onde fixes ou variables (permettant des mesures de dispersion anomale, MAD). Elle est équipée d'un détecteur Pilatus 6M et d'un diffractomètre avec un goniomètre 3 cercles. L'optique de la ligne délivre un faisceau d'une taille de 100x100 µm² à 250x250 µm² avec une large gamme d'énergies accessibles (de 5 à 15 keV) correspondant à des longueurs d'onde comprises entre 0,83 et 2,48 Å. Un système cryogénique complet y est également installé (OXFORD system) et il est également possible d'y effectuer des mesures de fluorescence.

Les données enregistrées sont ensuite indexées, intégrées et mises à l'échelle avec les programmes XDS et XSCALE (Kabsch, 1993).

3. Le problème des phases

Comme nous l'avons vu précédemment, les données de diffraction donnent directement accès à l'intensité I(hkl) et, par conséquent, à l'amplitude du facteur de structure F(hkl). Cependant, l'information concernant les phases est perdue au moment de l'enregistrement des données sur le détecteur. Nous sommes donc dans l'impossibilité de calculer la fonction de densité électronique. Trois méthodes principales (MIR, MAD et remplacement moléculaire) sont utilisées en cristallographie biologique pour résoudre le problème des phases.

- La méthode MIR (« *Multiple Isomorphous Remplacement* ») consiste à comparer les clichés de diffraction des cristaux dérivés par rapport à ceux des cristaux natifs. Cette technique nécessite de fixer de manière spécifique, au sein de la protéine, des atomes lourds localisables dans le cristal (Harker, 1956; Perutz, 1956). La contribution de ces atomes, beaucoup plus riches en électrons que les atomes composant classiquement une protéine, détermine un pouvoir de diffraction plus important, la rendant relativement simple à détecter. Toutefois, cette technique requiert d'obtenir des cristaux dérivés isomorphes des cristaux natifs, ce qui se révèle être souvent difficile. L'isomorphisme implique que les paramètres de maille ainsi que la position des macromolécules n'ont pas subi de variations trop importantes.

- La méthode MAD (« *Multiple Anomalous Diffraction* ou *Dispersion* ») exploite la variation de facteur de diffusion d'un seul type d'atome en fonction de la longueur d'onde des rayons X. En effet, lorsque la longueur d'onde incidente des rayons X correspond à une transition entre deux niveaux d'énergie atomiques, il se produit un phénomène de résonance se traduisant par l'apparition d'un terme complexe dans le facteur de diffusion de l'atome résonant : il s'agit du phénomène de diffusion anomale.

- La méthode du remplacement moléculaire nécessite l'enregistrement des données de diffraction d'un cristal de la protéine et l'existence d'une structure tridimensionnelle d'une protéine suffisamment homologue. Il est alors possible d'estimer les phases de la macromolécule étudiée en plaçant la protéine homologue (modèle-guide) à partir des facteurs de structure calculés depuis le modèle guide. Une recherche sur six dimensions est nécessaire afin de trouver le meilleur placement du modèle-guide. Il est possible de répartir cette recherche en deux étapes à trois dimensions chacune : la fonction de rotation suivie de la fonction de translation.

a. Fonction de rotation

A partir des amplitudes observées (F_{obs}) et de celles calculées (F_{calc}), il est possible d'établir les fonctions de Patterson du modèle-guide et des données de diffraction :

$$P(t) = \frac{1}{V} \sum F(h)^2 . e^{(-2\pi i h.t)}$$

avec F = F_{calc} ou F_{obs}, t représente les valeurs interatomiques et h un vecteur dans l'espace réciproque.

Les vecteurs interatomiques peuvent être intramoléculaires ou intermoléculaires. Un rayon d'intégration (généralement le rayon du modèle-guide) sera choisi de manière à inclure un maximum de vecteurs intramoléculaires tout en excluant un maximum de vecteurs intermoléculaires.

La fonction de rotation permet de déterminer la matrice de rotation. Elle est calculée sur les vecteurs intramoléculaires par superposition de la Patterson du modèle-guide sur celle des données. Les orientations correspondant aux meilleures corrélations seront sélectionnées. L'algorithme utilisé pour calculer cette fonction a été proposé par Rossmann & Blow (Rossmann & Blow, 1962).

b. Fonction de translation

Ce calcul consiste à rechercher par des mouvements de translation, dans une unité asymétrique, les positions correspondant aux meilleures corrélations entre les F_{obs} et les F_{calc}. Les déplacements sont basés sur l'utilisation d'une grille adaptée à la résolution maximale (généralement ¼ à ½ de la résolution maximale). Plusieurs méthodes de calcul peuvent être utilisées pour cette recherche.

Le programme BALBES (Long *et al.*, 2008) de la suite logicielle CCP4 (Collaborative Computational Project, 1994), correspond à un enchaînement de routines permettant de rechercher automatiquement les solutions de rotation puis de translation présentant la meilleure corrélation après avoir déterminé la meilleure structure homologue disponible dans sa propre base de données interne.

Un facteur R de confiance représentant l'accord existant entre le modèle (F_{calc}) et les données cristallographiques (F_{obs}) peut être calculé à la fin du processus. Il permet de juger de la vraisemblance du modèle déterminé par le programme. Ce facteur se définit comme suit :

$$R = \frac{\sum_h \left\| F_{obs}(h) \right| - k \left| F_{calc}(h) \right\|}{\sum_h \left| F_{obs}(h) \right|}$$ avec k, facteur d'échelle.

La méthode du remplacement moléculaire est de plus en plus utilisée, compte tenu du nombre grandissant de structures expérimentales connues.

4. Affinement du modèle : vers la structure finale

Le modèle de la protéine étudiée (F_{calc}) doit expliquer au maximum les données expérimentales (F_{obs}). Pour obtenir le meilleur accord entre les données expérimentales et calculées, il est nécessaire, pour chaque atome, de modifier sa position spatiale (x,y,z), son facteur d'agitation thermique B et éventuellement son facteur d'occupation, si les données le permettent. Pour affiner tous ces paramètres, le nombre d'observations doit normalement être supérieur au nombre de paramètres à affiner d'au moins un facteur 3 à 4, du fait des erreurs sur les mesures, et dans tous les cas supérieur à 1. Ce nombre d'observations est rarement atteint en cristallographie des macromolécules biologiques (résolution insuffisante) et l'affinement utilise des restrictions stéréochimiques en plus des contraintes cristallographiques pour minimiser l'énergie totale du système selon l'équation :

$$\mathbf{E_{totale}} = E_{cristallographique} + [\, E_{bond} + E_{angl} + E_{impr} + E_{elec} + E_{vdw} + E_{pelec} + E_{pvdw}]$$

ou de manière simplifiée :

$$\mathbf{E_{totale}} = E_{cristallographique} + E_{empirique}$$

Le terme $E_{empirique}$ représente les termes d'énergie liés aux contraintes stéréochimiques. Elle est égale à la somme des énergies liées aux interactions covalentes (longueur E_{bond}, angles de liaison E_{angl}, angles impropres E_{impr}) et non covalentes (potentiels électrostatiques E_{elec} et interactions de Van der Waals E_{vdw}). Pour les atomes reliés par symétrie cristalline, ces interactions non covalentes sont représentées par les termes E_{pelec} et E_{pvdw}. L'évolution de l'énergie empirique est suivie en considérant la différence entre la valeur actuelle et la valeur idéale. Ainsi, si les atomes dévient de la géométrie idéale, la quantité d'énergie augmente.

Le terme $E_{cristallographique}$ est la différence pondérée existant entre les facteurs de structure observés et calculés. Il s'écrit :

$$E_{\text{cristallographique}} = \frac{W_a}{N_a} \sum_{hkl} W_{hkl} \left[\left| F_{obs}(hkl) \right| - k. \left| F_{calc}(hkl) \right| \right]$$

Le facteur W_a est un facteur de pondération appliqué à toutes les réflexions de manière à ce que le gradient de l'énergie soit égal au gradient de l'énergie empirique. Il est déterminé par une étape de dynamique moléculaire. Le facteur N_a est un facteur de normalisation, W_{hkl} le poids associé à une réflexion donnée, et k est le facteur de mise à l'échelle entre les facteurs de structure observés et calculés.

Les cycles d'affinement permettent, par de subtiles variations de coordonnées atomiques, de faire évoluer le système vers une conformation plus stable et donc de moindre énergie. Il existe trois protocoles principaux de minimisation d'énergie : l'affinement en corps rigide, par gradient conjugué de Powell ou encore par recuit simulé.

Pour réaliser les cycles d'affinement, nous avons utilisé le logiciel REFMAC 5.7 (Vagin *et al.*, 2004) de la suite CCP4 (Collaborative Computational Project, 1994). Ce programme permet d'appliquer des contraintes stéréochimiques au modèle de manière à maintenir la longueur des liaisons, les angles de valence, etc. proches des valeurs canoniques. Aux résolutions usuelles pour les structures de protéines (3,0 à 1,5 Å), le nombre d'observations est insuffisant pour effectuer un affinement sans contraintes. Tout au long de l'affinement, la qualité du modèle est contrôlée par l'évolution des facteurs R et R_{free} (définis ci-après) mais aussi par la valeur des écarts types par rapport aux distances et angles stéréochimiques idéaux.

$$R = \frac{\sum_h \left\| F_{obs}(h) \right| - k \left| F_{calc}(h) \right\|}{\sum_h \left| F_{obs}(h) \right|} \qquad \text{et} \qquad R_{free} = \frac{\sum_{h \in T} \left\| F_{obs}(h) \right| - k \left| F_{calc}(h) \right\|}{\sum_{h \in T} \left| F_{obs}(h) \right|}$$

où k est un facteur d'échelle et T, un jeu de réflexions qui comprend, dans cette étude, une sélection aléatoire de 5% de toutes les réflexions observées (Brunger, 1992).

Le facteur R_{free} se calcule donc comme le facteur R mais porte sur des réflexions non biaisées par le modèle dans la mesure où elles ne sont pas prises en compte dans l'affinement. Un modèle correct doit pouvoir rester en accord avec ces réflexions tout au long de l'affinement. Pour l'affinement de nos structures, 5% des réflexions ont été aléatoirement mises de côté.

5. Les cartes de densité électronique

La seule application des méthodes d'affinement ne suffit pas à obtenir une structure affinée et un certain nombre de corrections doivent être faites manuellement sur le modèle. De plus, ce dernier est souvent partiel en début d'affinement et doit être complété. Les modifications et les ajouts sont effectués à partir de l'interprétation de différents types de cartes de densité électronique, notamment des cartes dites de « Fourier difference ». Celles-ci sont calculées par transformée de Fourier de différentes combinaisons entre facteurs de structures observés et ceux calculés à partir du modèle :

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}(m|F_{obs}| - n|F_{calc}|)e^{i\varphi_{calc}}e^{-2\pi i(hx+ky+lz)}$$

où φ_{calc} représente la phase calculée et (m, n) appartient à {(3, 2), (2, 1), (1, 1)}.

Les cartes calculées avec le couple (3, 2) et (2, 1) sont en général utilisées pour visualiser la densité électronique de la structure et construire le modèle car elles donnent plus de poids aux observations expérimentales et réduisent le biais du modèle. Les différentes cartes de densité électroniques ont été observées avec le programme COOT (Emsley *et al.*, 2010).

Les cartes de « Fourier difference », c'est-à-dire celles calculées avec les coefficients (1, 1) permettent de localiser les différences entre le modèle et la structure. Les pics positifs de cette carte indiquent la position d'électrons (donc d'atomes) manquants dans le modèle. A l'inverse, les pics négatifs indiquent les positions d'électrons excédentaires. Une paire de pics positif-négatif signale qu'un atome, présent dans le modèle, est mal placé et qu'il doit être déplacé du pic négatif vers le pic positif. Cependant, dans un cas réel, l'interprétation de ces cartes est plus complexe car un certain nombre d'erreurs ou d'artéfacts peuvent entacher la carte. On parle alors de « bruit ». À terme, il aurait été nécessaire de réaliser en alternance plusieurs séries de cycles d'affinement par REFMAC5 et des corrections manuelles grâce au programme COOT.

IV. « Docking »

Le « docking » moléculaire est une méthode *in silico* qui a pour but de déterminer le mode d'interaction d'un complexe généralement formé de deux molécules, en cherchant des orientations dans l'espace et des conformations favorables pour la fixation d'un ligand à un récepteur.

Le plus souvent, le récepteur est une protéine qui possède un ou plusieurs sites actifs spécifiques, plus ou moins accessibles selon les cas. Le ligand est généralement, une molécule flexible étrangère de petite taille. Cependant, le ligand peut également être une molécule de plus grande taille comme un polymère d'acides nucléiques, un polypeptide, voire une autre protéine.

Par extension, cette technique permet de prédire, pour une petite molécule donnée complexée sur sa protéine cible, son affinité et son activité (substrat, cofacteur, inhibiteur…). Ainsi, le « docking » joue un rôle important dans la conception rationnelle de médicaments.

A. Méthode de « docking »

Deux approches complémentaires sont employées pour les expériences de « docking ». La première utilise la complémentarité de surface entre le récepteur et le ligand, la seconde est basée sur le calcul de l'énergie du complexe en utilisant une complémentarité d'interaction fondée sur des règles physico-chimiques.

- <u>Complémentarité de surface</u>

Cette méthode décrit la protéine (récepteur) en termes de surface moléculaire accessible au solvant et le ligand par des sphères de surface d'occupation. Le calcul de « docking » consiste à remplir les surfaces accessibles du récepteur avec les sphères du ligand. Il s'agit d'une occupation stérique, géométrique dans l'espace.

- <u>Calcul de l'énergie du complexe</u>

Cette méthode utilise des descripteurs correspondants à des points dans l'espace auxquels certaines propriétés physico-chimiques sont assignées. Les descripteurs du ligand et du récepteur doivent alors coïncider géométriquement et chimiquement. Dans cette approche, la protéine et le ligand sont séparés par une distance physique et le ligand « trouve » sa position dans un espace déterminé de la protéine (le site actif par exemple) après un certain nombre de « mouvements » dans son espace conformationnel. Ces mouvements comprennent tant les translations et rotations de la molécule ligand que des

changements structuraux internes comme des rotations d'angles de torsion. L'ensemble de ces mouvements induisent un cout énergétique au système.

Le ligand peut ainsi être perçu et utilisé comme une entité rigide ou flexible. Dans ce dernier cas, il est nécessaire d'identifier les liaisons possédant une liberté de rotation. Il est également possible de décrire le site de fixation du récepteur avec des degrés de flexibilité (définition d'angles de torsion pour quelques chaines latérales). Toutefois, l'augmentation du nombre de degrés de liberté accroit les temps de calculs.

B. La fonction de score

La fonction de score estime la complémentarité récepteur-ligand. Le plus souvent, elle prédit le gain d'énergie du ligand en interaction avec le récepteur par rapport à la forme libre. Le « docking » génère de nombreux faux positifs, cependant leur nombre peut être réduit par des données empiriques connues au préalable. Les solutions proposées sont classées en fonction de leur score. Cependant, ce classement doit être utilisé avec prudence et les solutions retenues doivent être examinées en détail.

C. Les logiciels majeurs

Il existe des logiciels avec un accès payant comme FlexX (www.biosolveit.de/flexx/) et GOLD (www.ccdc.cam.ac.uk/Solutions/GoldSuite/Pages/GOLD.aspx).

Il existe également des logiciels sous licence académique, comme par exemple AutoDOCK (http://autodock.scripps.edu/).

Dans le cadre de cette thèse, les structures du ligand et du récepteur ont été préparées avec le programme AutoDOCK Tools (Morris *et al.*, 2009). Le ligand (panose) a été traité comme une entité flexible, tout comme deux résidus du site actif de SHWaglu (Val 438 et Phe 370) et le « docking » a été réalisé avec AutoDOCK sur un espace de recherche réduit autour du site actif.

Résultats et Discussion

Chapitre 1 : Étude de l'α-glucosidase de *Lactobacillus bulgaricus* ATCC11842

I. Étude bioinformatique

A. Séquence complète

Une analyse bioinformatique a été réalisée pour déterminer si la protéine était potentiellement membranaire mais aussi si elle possédait un domaine de type CBM. Ainsi, différents programmes répertoriés dans ExPASy (http://www.expasy.org) ont été utilisés pour prédire la présence éventuelle d'un (ou plusieurs) passage(s) transmembranaire(s). Les programmes TopPred (Claros & von Heijne, 1994), TMpred (Hofmann & Stoffel, 1993) et DAS (Cserzo *et al.*, 1997) ont ainsi présagé un passage transmembranaire compris entre les résidus 410 et 430, alors que les serveurs TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/), SOSUI (Mitaku *et al.*, 2002) et HMMTOP (Tusnady & Simon, 1998) n'en ont prédit aucun.

Des modèles tridimensionnels ont été générés sur la base de la séquence entière de 11842aglu en utilisant le serveur « I-TASSER ». Les cinq modèles (Figure 43) obtenus ont été construits par modélisation moléculaire par homologie à partir de différentes structures tridimensionnelles résolues par cristallographie aux rayons X issues de la PDB : 1XSI (Lovering *et al.*, 2005), 1XSK (Lovering *et al.*, 2005), 2F2H (Kim *et al.*, 2006), 2QMJ (Sim *et al.*, 2008) et 2G3N (Ernst *et al.*, 2006). Elles correspondent respectivement à l'α-xylosidase YicI issue d'*E.coli* pour les 3 premières, à la sous-unité N-terminale de la maltase-glucoamylase humaine et à l'α-glucosidase de *Sulfolobus solfataricus* avec 16,6%, 15% et 14,9% d'identité, respectivement, et toutes appartenant à la famille GH31. Les modèles diffèrent entre eux, tout particulièrement du coté C-terminal qui n'est pas ou peu replié. La région N-terminale, structurellement conservée, correspond aux domaines N et catalytique. Parmi les cinq modèles générés, la région C-terminale d'une des solutions (modèle 3) n'a pas été modélisée (résidus 590 à 1001). Ces observations tendent à montrer que cette région, chez 11842aglu, ne serait pas un domaine conservé chez toutes les enzymes issues de la famille GH31. De plus, selon ces modèles, les résidus se trouvant aux positions 410 à 430, prédits comme étant un passage transmembranaire, sont inclus dans une hélice α du domaine catalytique (Figure 43). Il apparait donc plus raisonnable de privilégier les résultats des serveurs TMHMM, SOSUI et HMMTOP.

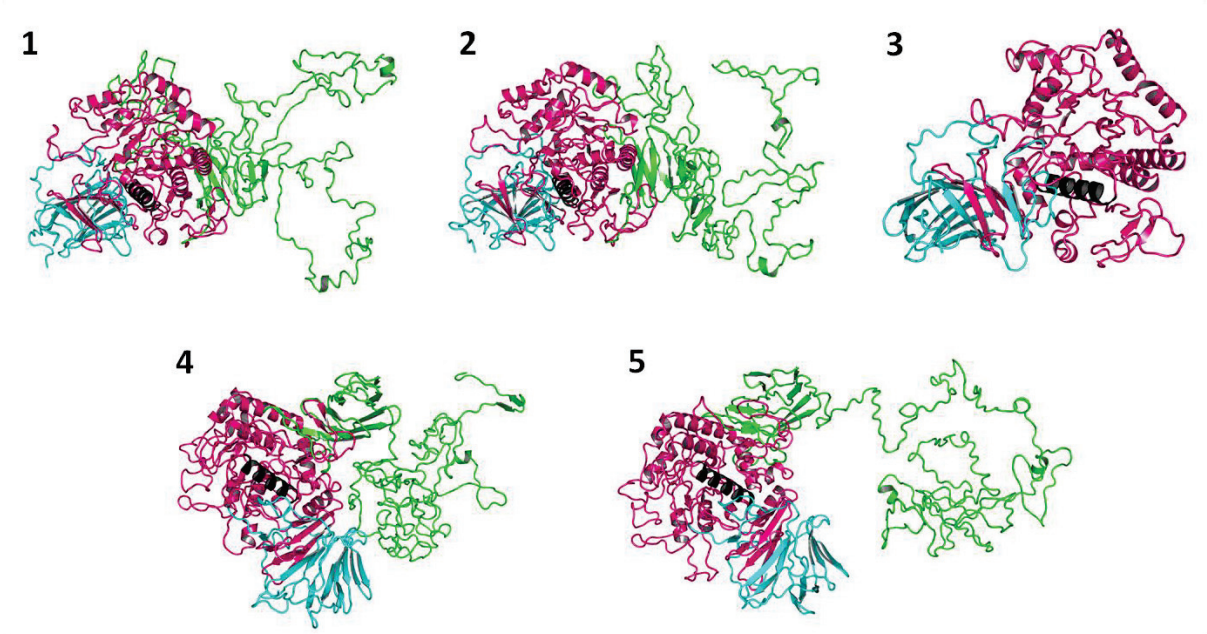


Figure 43 : Modèles structuraux de 11842aglu générés par I-TASSER. Les domaines N et catalytique sont colorés en cyan et magenta, respectivement. La région C-terminale est colorée en vert. La portion (résidus 410 à 430) prédite par certains programmes comme potentiellement membranaire est colorée en noir. Le modèle 3 est tronqué après le domaine catalytique, l'extrémité C-terminale n'ayant pas été modélisée.

B. Région C-terminale

Une étude bioinformatique plus approfondie de la région C-terminale a ensuite été réalisée. Cette région peut en fait être scindée en deux, avec une partie proximale et distale correspondant respectivement aux domaines C et D (décrits dans la partie bibliographique). En effet, la partie proximale est modélisée comme un ensemble de brins β, tel qu'attendu. Seule la partie distale n'a pas du tout pu être modélisée correctement. Ainsi, une deuxième série de modèles se basant uniquement sur les 190 derniers résidus de la séquence a été générée par « I-TASSER ». Cette fois-ci, cinq modèles avec un repliement correct ont pu être obtenus avec la présence d'un feuillet β (Figure 44). Les cinq modèles se superposent correctement à l'exception de l'extrémité N-terminale qui diffère notablement. Comme précédemment, ceux-ci ont été générés à partir de différentes structures tridimensionnelles résolues par cristallographie aux rayons X ayant pour codes PDB : 2V5D (Ficko-Blean *et al.*, 2009), 2V72 (Boraston *et al.*, 2007), 2ZXQ (Suzuki *et al.*, 2009), 3HNM (non publié) et 4A41 (Ficko-Blean *et al.*, 2012) avec 23,3%, 11%, 15,6%, 16,1% et 16,2% d'identité, respectivement. Ces structures correspondent toutes à des glycoside hydrolases possédant un module CBM32, suggérant que le domaine D de 11842aglu en serait également un. Cette famille de CBM est connue pour être l'une des plus diversifiée et est retrouvée dans différentes familles de pectate lyases et de glycoside hydrolases comme les GHs 8, 10, 16, 29, 84, 85, 89 et 31 (Abbott *et al.*, 2008). Le « sous-groupe GH31 » comporte des α-

glucosidases provenant de plusieurs souches de *Bacteroides* et *Lactobacillus*. Ceci renforce l'hypothèse que le domaine D est bien un CBM32. Les CBM32 ont montré leur capacité à fixer une grande diversité de ligands oligosaccharidiques avec, toutefois, la présence consensus d'une unité de galactose (Abbott *et al.*, 2008). Cette famille de CBM adopte un repliement en feuillet β avec la présence d'un ion structural, modélisé comme un ion calcium (Ficko-Blean *et al.*, 2012). La définition même d'un domaine CBM étant de « concentrer » l'enzyme qui le possède vers ses substrats de prédilection (Boraston *et al.*, 2004), il est fort probable que 11842aglu soit active sur des substrats possédant au moins une unité de galactose. Cependant, la famille CBM32 est étendue et son spectre de spécificité de ligands est susceptible d'être bien plus vaste que ce qui a déjà été décrit (Abbott *et al.*, 2008). Une autre théorie sur la présence de ce module chez 11842aglu peut être avancée : en effet, il a été rapporté que le CBM32 de la β-*N*-acetylglucosaminidase GH89 de *Clostridium perfringens* pouvait interagir avec les mucines de classe III (retrouvées à la surface des muqueuses gastro-intestinales) dont le motif comporte un galactose (Ficko-Blean *et al.*, 2009). Ainsi, il est probable que le CBM32 de 11842aglu puisse conférer une capacité d'adhésion aux mucines de classe III, ceci permettant à la bactérie une interaction étroite avec l'hôte dans le cas où l'enzyme est immobilisée à la surface cellulaire bactérienne (Ficko-Blean *et al.*, 2009). Ceci pourrait en partie expliquer l'action probiotique de *Lactobacillus bulgaricus* et plus particulièrement le rôle de protection contre les pathogènes par un mécanisme d'adhésion compétitive *via* le mucus. Cet effet de barrière a plusieurs fois été décrit (Candela *et al.*, 2008; Van Tassell & Miller, 2011; Turpin *et al.*, 2012).

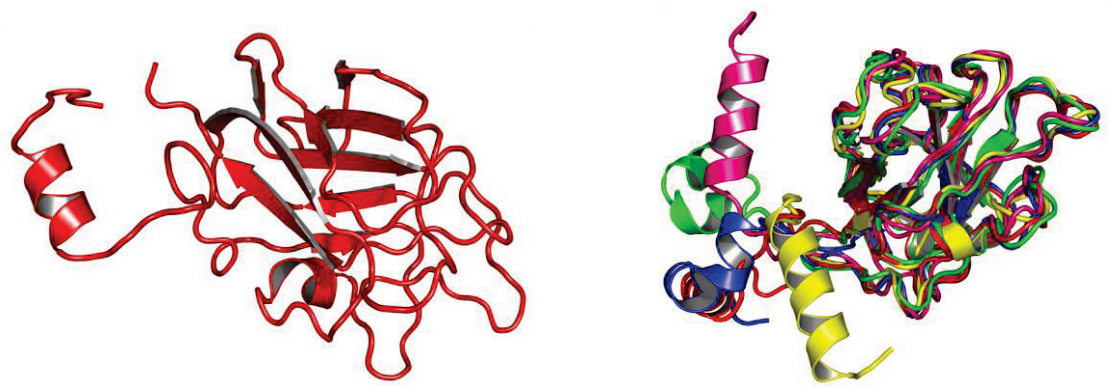


Figure 44 : A gauche, représentation en ruban d'un des modèles de la région C-terminale générés par I-TASSER. A droite, superposition des cinq modèles obtenus.

Tous ces résultats sont en accord avec les prédictions annoncées dans la base données Uniprot. En effet, bien qu'elle soit annotée comme protéine non caractérisée, elle est tout de même classéee comme appartenant à la famille GH31 et possédant un domaine-*like* de fixation au galactose. L'organisation modulaire de 11842aglu basée sur les prédictions et les modélisations est schématisée en Figure 45.

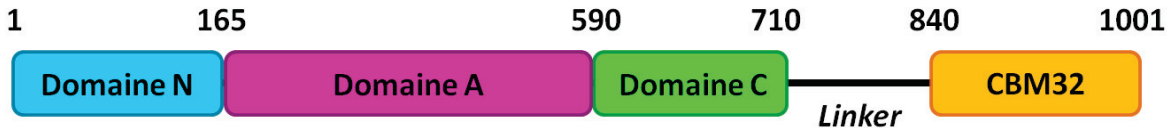


Figure 45 : Schéma général de l'organisation modulaire de 11842aglu d'après les prédictions réalisées.

II. Biologie moléculaire

Le génome de *Lactobacillus bulgaricus* ATCC11842 étant gracieusement fourni par le Dr. E. Maguin (Département « Microbiologie et Chaîne Alimentaire », INRA, Jouy-en-Josas), nous avons cloné le gène d'intérêt dans le vecteur pET15b apportant une étiquette His_6 N-terminale clivable à la thrombine. En utilisant cet ADN génomique comme matrice, nous avons amplifié le gène *11842aglu* par PCR grâce à des amorces construites de sorte à inclure les sites de restriction compatibles pour le clonage dans le vecteur choisi, à savoir *Nde*I en 5' en amont du gène et *Cla*I en 3' après le codon stop.

NOM	SEQUENCE DE L'AMORCE 5'→ 3'
OL-Direct-11842	GTCGTC**CATATG**ACTGCAGACCTGCAAAATTCAAGACATC
OL-Reverse-11842	GACGAC**ATCGAT**TTATTATTTAAGGGGCCGGTAGAAG

Tableau 11 : Amorces utilisées pour l'amplification par PCR du gène *1403aglu*. Les sites de restriction *Nde*I et *Cla*I sont surlignés en gris.

L'insert amplifié a ensuite été extrait puis purifié après digestion par les enzymes *Nde*I et *Cla*I, tout comme le vecteur pET15b. Après ligature et transformation dans la souche *E. coli* DH5α, les plasmides ont été extraits et digérés par des enzymes de restriction appropriées (*Nde*I, *Cla*I et *Eco*RI) afin de vérifier la présence de l'insert. En parallèle, une PCR a également été réalisée sur ces mêmes vecteurs pour confirmer le choix des plasmides recombinants. Au final, le seul clone apparaissant comme recombinant a été envoyé pour séquençage chez la société GATC afin de valider sa séquence. Une fois approuvé, le vecteur sélectionné a été transformé dans une souche d'expression *E. coli* BL21 (DE3).

III. Expression et purification

A. Expression et extraction de la protéine soluble

Un premier test d'expression a été effectué en induisant un petit volume de culture (5 ml) à une DO_{600} ≈ 0,6 avec 0,5 mM d'IPTG à 18°C, 25°C et 37°C. Les culots bactériens obtenus après centrifugation ont été lysés par sonication après resuspension dans du tampon MES 100 mM, pH 6. Le lysat bactérien a ensuite été centrifugé pour séparer la fraction soluble de la fraction insoluble. La protéine d'intérêt comportant une étiquette His_6, une révélation a été réalisée par Western-Blot afin de vérifier son expression. A l'issue de ce premier test, les résultats montrent une induction de l'expression de la protéine uniquement à 18°C. Il en ressort également que la protéine est soluble car retrouvée dans le surnageant et qu'elle migre sur SDS-PAGE à la taille attendue. Une analyse de la séquence par spectrométrie de masse a permis de confirmer que la protéine entière était exprimée.

L'expression a ensuite été testée sur de plus gros volume de culture (300 ml) en induisant par 0,8 mM d'IPTG à 18°C. Différentes méthodes de lyse ont ensuite été testées en parallèle : la sonication, la presse de French et le broyeur manuel. Dans les trois cas, la protéine a toujours été retrouvée dans le culot, ceci attestant de son insolubilité. D'autres tests ont ensuite été réalisés en variant la DO_{600} d'induction, mais aussi la concentration finale d'IPTG. Ainsi, les DO_{600} de 0,4 et 0,8, ainsi que les concentrations de 0,25 et 0,5 mM d'IPTG ont été testées en parallèle sur des cultures de 300 ml. Cette fois encore, pour ces quatre conditions d'expression, la protéine n'a été révélée par Western-Blot que dans le culot. Toutes ces expériences ont été réalisées avec du tampon MES 100 mM pH 6.

En reprenant la condition d'induction de l'expression initiale (0,8 mM IPTG à DO_{600} ≈ 0,6 à 18°C), d'autres tampons de resuspension ont été expérimentés : HEPES 100 mM pH 8, HEPES 100 mM pH 8 avec 150 mM NaCl, HEPES 100 mM pH 8 avec 150 mM NaCl et 10% de glycérol. La présence de sels dans le tampon permet d'augmenter la solvatation de la protéine tandis que le glycérol est parfois utilisé comme agent stabilisant. Une analyse par Western-Blot montre qu'en aucun cas la solubilité de la protéine n'a été améliorée par ces protocoles.

Etant donné les problèmes de solubilité, il a été supposé que la protéine formait des corps d'inclusion dans les bactéries lors de son expression. Ainsi, un protocole permettant d'extraire et de resolubiliser les corps d'inclusion a été testé. Dans un premier temps, ce protocole consiste à former des sphéroplastes en utilisant un tampon composé de Tris-HCl 20 mM pH 7,5, de 20% de saccharose et 5 mM d'EDTA. Ce tampon permet en fait de fragiliser les bactéries et casser leur paroi, libérant ainsi les protéines périplasmiques. L'étape suivante consiste à rompre la membrane plasmique des bactéries en utilisant de

l'eau glacée qui provoque un choc thermique et osmotique sur ces cellules. Les étapes suivantes permettent d'extraire les corps d'inclusion, de les dénaturer puis de tenter de les renaturer progressivement. Ce nouveau protocole de lyse bactérienne a permis d'obtenir une partie de la protéine dans une fraction soluble (Figure 46).

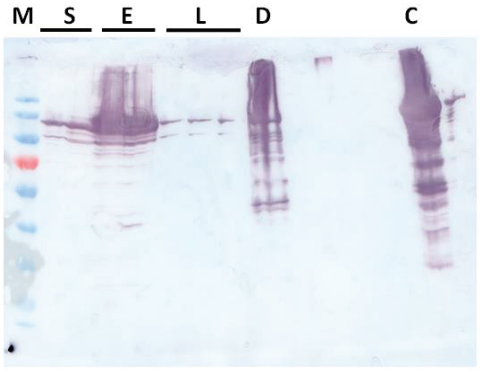


Figure 46 : Analyse par Western-Blot des différentes étapes du protocole d'extraction des corps d'inclusion. M : marqueurs de masse moléculaire. S : étape de libération des protéines périplasmiques ; E : surnageant obtenu après l'étape de lyse à l'eau glacée ; L : lavages des culots ; D : resuspension du culot dans le tampon de dénaturation ; C : dépôt de la fraction insoluble obtenue après tentative de renaturation et centrifugation.

B. Purification

Comme la protéine exprimée avec le système pET15b porte une étiquette His_6 du coté N-terminal, il est possible d'effectuer une chromatographie d'affinité sur colonne de nickel IMAC dont le protocole est décrit dans le chapitre « Matériel et Méthodes ». Lorsque l'on constate l'élution de protéines, le gradient est alors mis en pause pour obtenir un échantillon homogène en concentration d'imidazole et ainsi limiter une possible élution concomitante d'autres protéines.

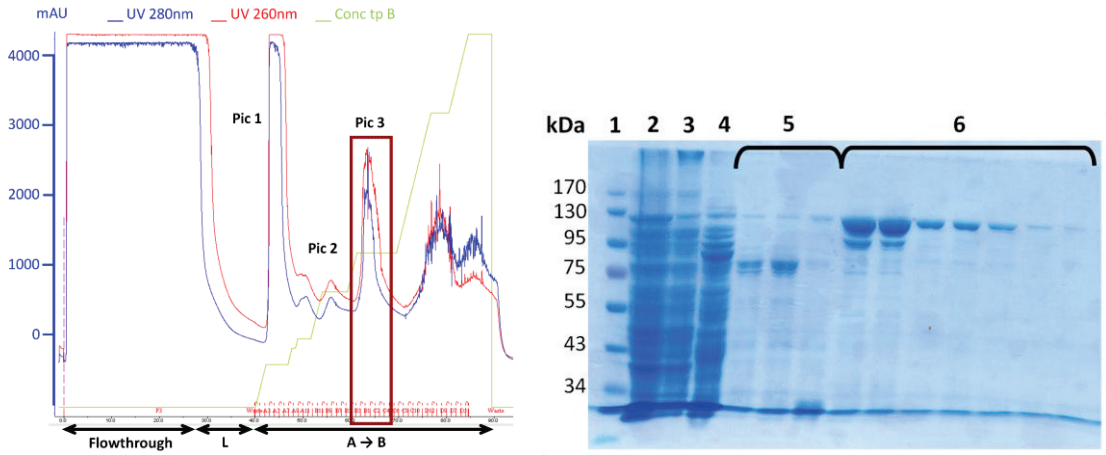


Figure 47 : Purification de 11842aglu par chromatographie d'affinité. A gauche, chromatogramme : les courbes bleue, rouge et verte représentent respectivement l'absorbance à 280 nm, à 260 nm et le gradient d'imidazole. *Flowthrough*, L et A→B correspondent respectivement à l'injection sur colonne, au lavage et au gradient d'imidazole. Le cadre rouge représente les fractions contenant la protéine d'intérêt. A droite, analyse par SDS-PAGE des fractions collectées. Piste 1 : marqueurs de masse moléculaire, piste 2 : surnageant avant purification, piste 3 : *Flowthrough*, piste 4 : pic 1, pistes 5 : fractions du pic 2, pistes 6 : fractions du pic 3.

Il a été observé sur les différents SDS-PAGE et Western-Blots réalisés que la protéine existait sous deux formes : une de taille correspondant à la protéine entière (environ 115 kDa), mais aussi une de taille inférieure (environ 100 kDa). Les bandes révélées par coloration au bleu de Coomassie ont donc été analysées par spectrométrie de masse. Ainsi, la bande de plus haut poids moléculaire contient bien la protéine entière puisque des peptides correspondant au début et à la fin de la séquence protéique ont été détectés. Pour la bande de plus bas poids moléculaire, aucun peptide n'a été détecté pour l'extrémité C-terminale. L'analyse semble démontrer la perte des 130 derniers résidus de la protéine, ce qui correspondrait au domaine CBM32.

En concentrant la protéine à 2,8 mg/ml à partir de ce lot de production, il a été possible d'effectuer les premiers essais de cristallogenèse. Ainsi, des gouttes assises de 300 nl (150:150) dans des boites TTPLabtech 3 cupules ont été réalisées par Mosquito® en utilisant les deux kits PEG et PEGII (Qiagen). Deux conditions protéiques ont été testées : la protéine seule et en présence de 10 mM d'acarbose. La majorité des gouttes sont restées claires laissant présumer que la protéine n'est pas assez concentrée pour atteindre le seuil de nucléation.

Chapitre 2 : Étude de l'α-glucosidase de *Lactococcus lactis* IL1403

I. Étude bioinformatique

Un modèle tridimensionnel a été généré sur la base de la séquence entière de 1403aglu en utilisant le serveur « I-TASSER » (Figure 48). Il a été construit par modélisation moléculaire par homologie à partir de différentes structures tridimensionnelles résolues par cristallographie aux rayons X issues de la PDB : 2ZE0 (Shirai *et al.*, 2008), 1UOK (Watanabe *et al.*, 1997), 2ZIC (Hondoh *et al.*, 2008), 1M53 (Zhang *et al.*, 2003), 1ZJA (Ravaud *et al.*, 2007). Elles correspondent respectivement à l'α-glucosidase de *Geobacillus* HTA-462, à l'oligo-1,6-glucosidase de *Bacillus cereus*, à la dextrane 1,6-α-glucosidase de *Streptococcus mutans*, à l'isomaltulose synthase de *Klebsiella* sp. LX3 et à la trehalulose synthase de *Pseudomonas mesoacidophila* MX-45 avec 43,5%, 39,1%, 36,2%, 37%, et 36,4% d'identité. Toutes ces enzymes qui ont servi comme modèle-guide appartiennent à la famille GH13.

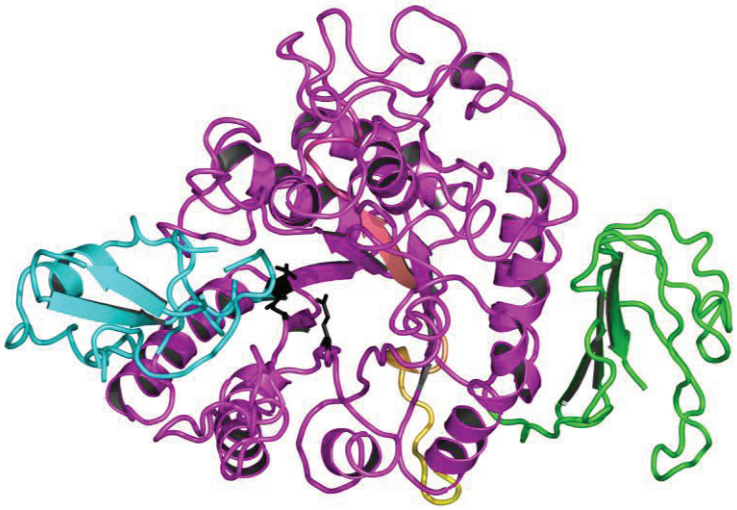


Figure 48 : Modèle structural de 1403aglu généré par I-TASSER. Les domaines A, B et C sont colorés en magenta, cyan et vert, respectivement. L'extrémité N-terminale et les résidus catalytiques sont colorés en jaune et noir, respectivement.

II. Biologie moléculaire

Nous avons cloné le gène d'intérêt dans le vecteur pET15b apportant une étiquette His_6 N-terminale clivable à la thrombine à partir du génome complet de *Lactococcus lactis* IL1403, déjà accessible au laboratoire. En utilisant comme matrice cet ADN génomique, nous avons amplifié par PCR le gène *1403aglu* grâce à des amorces construites de sorte à inclure les sites de restriction compatibles pour le clonage dans le vecteur choisi, à savoir *Nde*I en 5' en amont du gène et *Eco*RI en 3' après le codon stop (Tableau 12).

NOM	SEQUENCE DE L'AMORCE 5'→ 3'
OL-Direct-1403	GTCGTC**CATATG**AAAAAAATAAAAACAAATGATTTTC
OL-Reverse-1403	G**GAATTC**TTACTATTTTTTATAAACCCGATATTCCCAG

Tableau 12 : Amorces utilisées pour l'amplification par PCR du gène *1403aglu*. Les sites de restriction *Nde*I et *Eco*RI sont surlignés en gris.

L'insert amplifié a ensuite été extrait puis purifié après digestion par les enzymes *Nde*I et *Eco*RI, tout comme le vecteur pET15b. Après ligature et transformation dans la souche *E. coli* DH5α, les plasmides ont été extraits et digérés par une enzyme de restriction appropriée (*Eco*RI) afin de vérifier la présence de l'insert. En parallèle, une PCR a également été réalisée sur ces mêmes vecteurs pour confirmer le choix des plasmides recombinants. Avec une taille de vecteur linéarisé de près de 7000 pb, les clones 2 et 3 ont été présélectionnés. Seul le vecteur du clone 3 présente une amplification d'un fragment avec une taille proche de 1500 pb correspondant à la taille du gène d'intérêt. Pour distinguer sans ambiguïté les clones 2 et 3, une digestion par une autre enzyme de restriction a été réalisée (*Nco*I). Cette dernière possède un seul site de restriction dans pET15b et un autre dans le gène d'intérêt. Ainsi dans le cas d'un clone positif, un fragment de 700 pb serait libéré. C'est le cas pour le clone 3 mais pas pour le clone 2, confirmant bien les résultats de PCR (Figure 49). Seul le clone 3 a donc été envoyé en séquençage pour valider sa séquence. Une fois validé, le vecteur sélectionné a été transformé dans une souche d'expression *E. coli* BL21 (DE3).

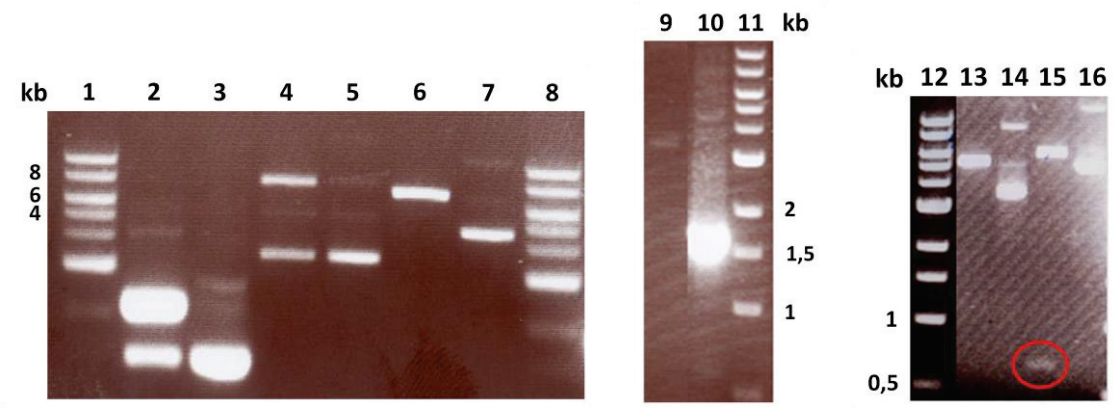


Figure 49 : A gauche, digestion par *Eco*RI des vecteurs pET15b pouvant contenir le gène *1403aglu*. Au centre, PCR du gène *1403aglu* sur les vecteurs sélectionnés. A droite, digestion par *Nco*I des vecteurs pET15b des clones 2 et 3 pouvant contenir le gène *1403aglu*. Pistes 1, 8, 11 et 12 : marqueurs de masse moléculaire. De 2 à 7 et de 13 à 16, pistes paires : vecteurs digérés, pistes impaires : vecteurs non digérés. Pistes 2 et 3 : clone 1, pistes 4 et 5 : clone 2, pistes 6 et 7 : clone 3, piste 13 et 14 : clone 2, piste 15 et 16 : clone 3. Piste 9 : clone 2, piste 10 : clone 3. Cercle rouge : fragment d'ADN de 700 pb.

III. Expression et purification

Un premier test d'expression a été effectué en induisant la culture à une DO_{600} ≈ 0,6 avec 0,5 mM d'IPTG à 37°C. La protéine d'intérêt comportant une étiquette His_6, une révélation a été réalisée par Western-Blot afin de vérifier son expression. Ce premier résultat indique que l'α-glucosidase IL1403 est effectivement bien exprimée mais retrouvée majoritairement dans le culot et non dans le surnageant, ceci montrant qu'elle est en grande partie insoluble (Figure 50).

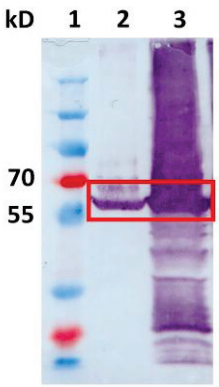


Figure 50 : Révélation de l'étiquette His_6 de la protéine d'intérêt (cadre rouge). Piste 1 : marqueurs de masse moléculaire, piste 2 : surnageant, piste 3 : culot.

Une analyse en spectrométrie de masse a permis de confirmer qu'il s'agissait bien de l'α-glucosidase IL1403 exprimée sous sa forme entière avec un très bon taux de recouvrement.

MKKIKTNDFHKMVGYQIYPKSFKDTNGDGIGDINGIIEKIPYLRELGIDFIWINPIYKSPQVDGGYDISDYQVIDEM
FGSLEDFKKLLDKAHENGLKVIMDLVVNHTSDQHPWFLESQKSKDNPYRDYYLWEDASPNQMPNDWISFFGDS
TWAYDEKTKQAYFHVFAKEQPDLNWKNPKVRQEIYDMIRWWLDLGIDGFRLDAISHIQKMPWNYKIKSWKGD
GPWVPFMNVAGIETYMSDLKAIFDEYGALTVGEASGVRSKQAPDWTDEAGYINMIFELEHNVRKSNGNGQGS
IYGYKKVIMRWQEDLLERGWNALYIENHDQPRSIDTFGDGSIESAKALAVSYMLLRGTPFIYQGQEFGMTNFPFT
DAKQLRADEIKKKYEMLMTELSANQALAEVTQNSREHSRTPMQWDNSKNAGFTSGQAWMEINPNFSVISSRT
NQELIELYKKLIELRHQEPAISEGSIKFHFKRHPQVLVYERAEFLIIVNLADKKAHLDFTKIDSFTDSSAESSVINYQQV
LGSSNQKFQEKMVLSAWEYRVYKK

Figure 51 : Séquençage par spectrométrie de masse de la bande protéique correspondant à celle révélée par Western Blot. Vert : >95% de sûreté, jaune : >50% de sûreté, gris : non détecté.

Au regard de ces premières observations d'expression, plusieurs températures (20°C, 25°C et 37°C) et DO_{600} d'induction (0,6, 0,7, 0,8 et 0,9) ont été testées en parallèle dans le but d'améliorer la solubilité de la protéine surexprimée. A partir de ces observations, un protocole d'expression a été établi. Une colonie unique de la souche recombinante est inoculée dans 5 ml de préculture et mise sous agitation toute une nuit à 37°C. Cette dernière est ensuite inoculée dans 300 ml de LB supplémenté en ampicilline à une concentration de 150 µM. La culture est alors mise sous agitation à 25°C pendant quatre heures, puis refroidie à 20°C avant d'induire l'expression sur la nuit par 0,5 mM d'IPTG lorsque la DO_{600} atteint ≈ 0,9.

Les cellules bactériennes sont culottées par centrifugation puis resuspendues dans un tampon de lyse MOPS 100 mM pH 7 contenant du lysozyme à hauteur de 10 mg/ml et de la ribonucléase bovine. Cette suspension est ensuite placée à -80°C sur la nuit. Le lendemain, le lysat bactérien est décongelé puis soniqué trois fois 30 secondes. Il est ensuite centrifugé afin de culotter tous les débris cellulaires et filtré à 0,45 µm ; l'extrait protéique obtenu étant alors prêt pour l'étape de purification.

Comme la protéine exprimée avec le système pET15b porte une étiquette His_6 du coté N-terminal, il est possible d'effectuer une chromatographie d'affinité sur colonne de nickel IMAC dont le protocole est décrit dans le chapitre « Matériel et Méthodes ». Lorsque l'on constate l'élution de protéines, le gradient est alors mis en pause afin d'obtenir un échantillon homogène en concentration d'imidazole et limiter une possible élution concomitante d'autres protéines. La protéine d'intérêt est éluée à différentes concentrations d'imidazole (pic 1 à 200 mM, pic 2 à 240 mM et pic 3 à 290 mM) avec un degré de pureté croissant (Figure 52).

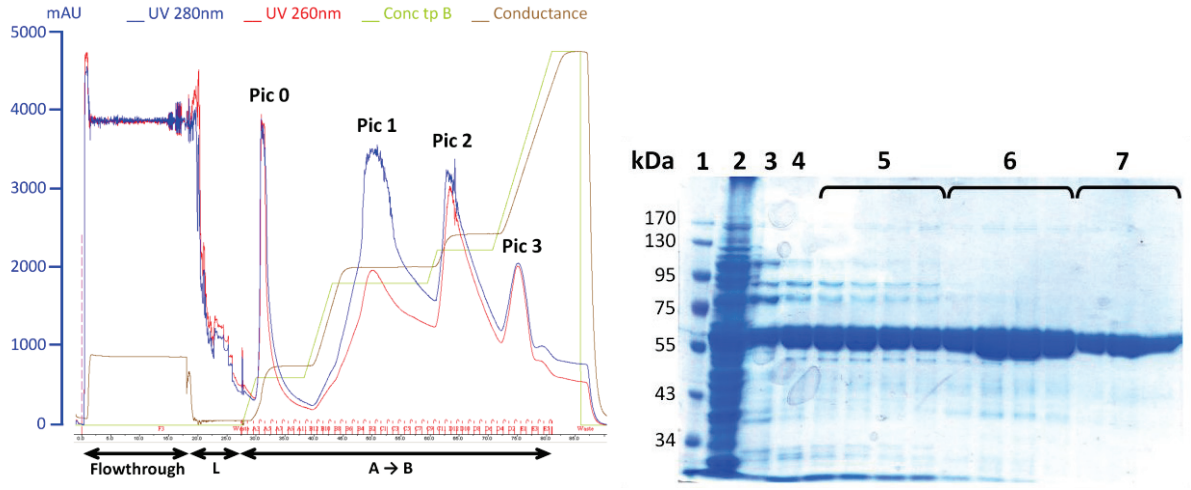


Figure 52 : Purification de 1403aglu par chromatographie d'affinité. A gauche, chromatogramme, les courbes bleue, rouge, verte et marron représentent respectivement l'absorbance à 280 nm, à 260 nm, le gradient d'imidazole et la conductance. *Flowthrough*, L et A→B correspondent respectivement à l'injection sur colonne, au lavage et au gradient d'imidazole. A droite, analyse par SDS-PAGE des fractions collectées. Piste 1 : marqueurs de masse moléculaire, piste 2 : surnageant avant purification, piste 3 : *Flowthrough*, piste 4 : pic 0, pistes 5 : fractions du pic 1, pistes 6 : fractions du pic 2, pistes 7 : fractions du pic 3.

Vu le degré de pureté des différents pics élués (pics 1 à 3), l'étape de purification par chromatographie d'affinité se suffit à elle-même. Plusieurs débits d'injection de l'échantillon sur la colonne ont été testés (1, 0,8 et 0,5 ml/min). Ainsi, il apparait que la réduction du débit permet d'augmenter le taux de pureté des pics élués. A partir de ces échantillons, il a été possible de réaliser les premières études fonctionnelles ainsi que des tests préliminaires de cristallogenèse.

IV. Caractérisation biochimique

A. Caractérisation préliminaire

Un test préliminaire d'activité sur le *p*NPG avec la protéine purifiée portant encore l'étiquette His_6 a été réalisé à 35°C pendant 30 minutes dans une gamme de pH allant de 3,5 à 9. Il s'avère qu'aucune activité n'a été détectée sur toute cette gamme.

Un clivage de l'étiquette par la thrombine, possible grâce à la présence du site de coupure entre l'étiquette et l'α-glucosidase, a donc été réalisé à 4°C pendant 18 h. Cette enzyme de clivage étant sensible à l'imidazole, un dessalage sur colonne (HiPrep™ 26/10 Desalting) a été nécessaire au préalable. Il en résulte un clivage total comme en attestent le SDS-PAGE et le Western-Blot anti-His_6 (Figure 53). En effet, on ne constate aucune révélation d'étiquette sur le Western-Blot après clivage et un léger déplacement de la bande sur le SDS-PAGE démontrant une diminution de la taille de la protéine d'intérêt. Ces résultats ont été

confirmés par analyse par spectrométrie de masse (résultats non montrés). Un test colorimétrique sur *p*NPG avec ce lot de protéine entièrement débarrassée de son étiquette a été réalisé à 35°C à pH 4 et 7 en présence (à hauteur de 1 mM final) ou en absence d'ions calcium. Cette fois-ci, une activité a été détectée démontrant que l'enzyme est plus active à pH 7 qu'à pH 4 et que l'ajout d'ions calcium réduit l'activité hydrolytique (résultats non montrés).

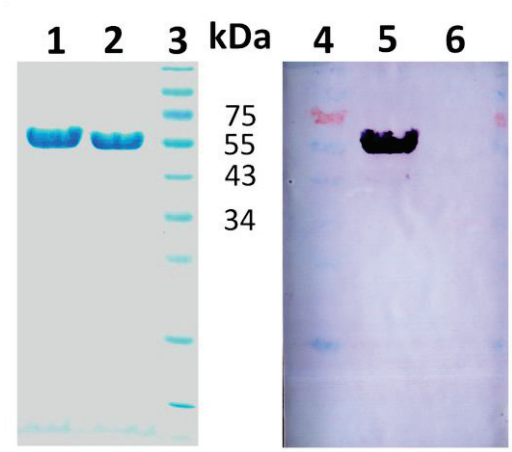


Figure 53 : SDS-PAGE (à gauche) et Western-Blot (à droite) de l'étape de clivage à la thrombine de l'étiquette His_6. Pistes 3 et 4 : marqueurs de masse moléculaire. Pistes 1 et 5 : 1403aglu avant clivage. Pistes 2 et 6 : 1403aglu après clivage.

1. Détermination du pH optimum

Des tests enzymatiques sur l'α-glucosidase libre ont été réalisés à différents pH dans une gamme allant de 3,5 à 9,5. Il en ressort que le pH optimal de l'enzyme est de 7 et que l'activité enzymatique chute assez rapidement de part et d'autre de ce pH avec toutefois une conservation de plus de 50% de son activité aux pH 6 et 8 (Figure 54). L'activité détectée est quasi nulle aux pH extrêmes de la gamme testée.

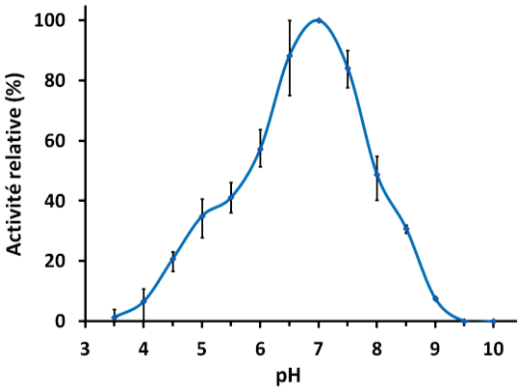


Figure 54 : Profil d'activité relative de l'enzyme 1403aglu en fonction du pH. Les barres représentent les déviations standards de trois mesures indépendantes.

2. *Détermination de la température optimale*

Des tests enzymatiques sur enzyme libre ont été également réalisés entre 4°C et 50°C afin de déterminer la température optimale d'activité. Nous constatons que l'enzyme présente un maximum d'activité à 30°C avec une diminution drastique de l'activité relative aux températures supérieures (Figure 55). Cette température optimale d'activité est en accord avec la température optimale de croissance de *Lactococcus lactis* qui est également de 30°C (Panoff *et al.*, 1994).

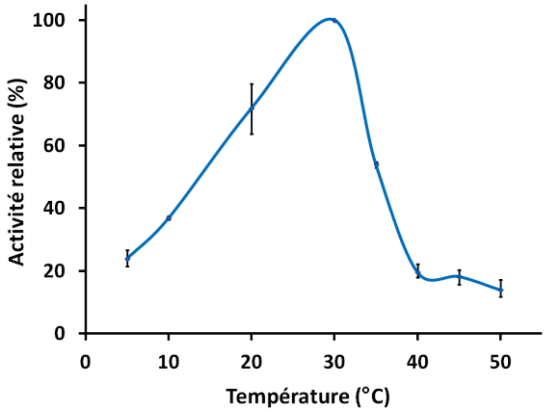


Figure 55 : Profil d'activité relative de l'enzyme 1403aglu en fonction de la température. Les barres représentent les déviations standards de trois mesures indépendantes.

3. *Effets des ions divalents sur l'activité catalytique*

Nous avons étudié les effets de différents cations divalents (1 mM final) sur l'activité de l'enzyme purifiée et clivée par la thrombine. Tous les ions testés ont eu un effet inhibiteur plus ou moins marqué. L'activité est réduite d'environ 20% en présence de Ca^{2+}, de Mg^{2+} ou de Mn^{2+}. Les ions Cu^{2+} inhibent de plus de la moitié l'activité enzymatique alors que les ions Zn^{2+} l'abolissent totalement (Figure 56).

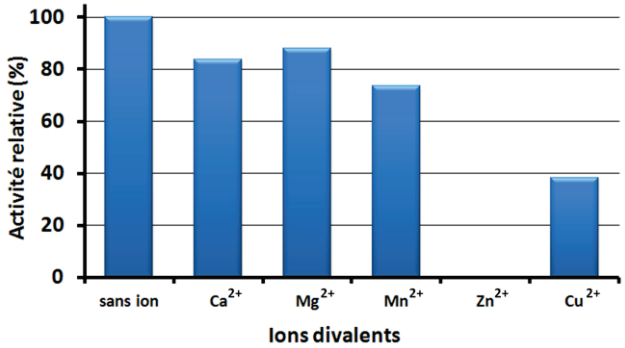


Figure 56 : Effets des cations divalents sur l'activité de l'enzyme purifiée et clivée par la thrombine. Les ions ont été ajoutés à hauteur de 1 mM. « sans ion » indique l'activité mesurée sans ajout de cations. L'activité relative est normalisée par rapport à l'activité mesurée sans ion.

4. Criblage des substrats naturels de 1403aglu

Cette étape de caractérisation enzymatique de spécificité de substrat a été réalisée en utilisant la technique de chromatographie sur couche mince (CCM).

Comme les α-glucosidases (EC 3.2.1.20), par définition, catalysent l'hydrolyse de liaisons inter-glycosidiques α-(1,4) du coté non-réducteur du substrat, les premiers tests enzymatiques sur des substrats naturels ont été réalisés avec le substrat préférentiel de l'enzyme qui est le maltose mais aussi avec différents maltooligosaccharides.

La CCM révèle que le maltose, le maltotriose, le maltotetraose, le maltopentaose, le maltohexaose et le maltoheptaose sont tous hydrolysés en glucose et sous-produits intermédiaires correspondant à tous les DP (degré de polymérisation) possibles liés au substrat initial. Ils sont donc bien des substrats naturels de 1403aglu (Figure 57). Il apparait toutefois que l'enzyme semble plus efficace sur les substrats ayant des DP supérieurs au maltose.

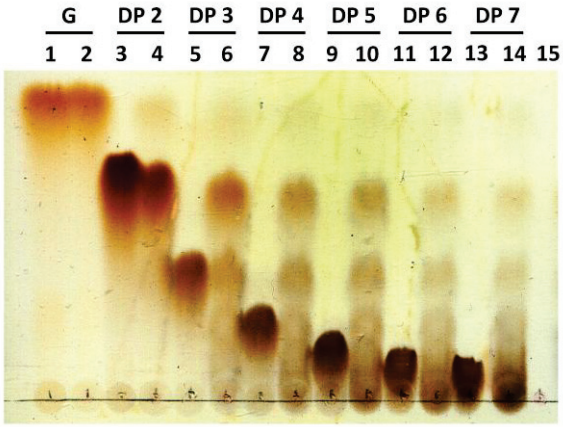


Figure 57 : Analyse par CCM après réaction enzymatique de 1403aglu clivée par la thrombine sur différents maltooligosaccharides. Pistes 1 et 2 : glucose en absence ou présence d'enzyme, respectivement. Piste 3 et 4 : maltose en absence ou présence d'enzyme, respectivement. Pistes 5 et 6 : maltotriose en absence ou présence d'enzyme, respectivement. Pistes 7 et 8 : maltotetraose en absence ou présence d'enzyme, respectivement. Pistes 9 et 10 : maltopentaose en absence ou présence d'enzyme, respectivement. Pistes 11 et 12 : maltohexaose en absence ou présence d'enzyme, respectivement. Pistes 13 et 14 : maltoheptaose en absence ou présence d'enzyme, respectivement. Piste 15 : dépôt de l'enzyme seule.

Dans la littérature, il a plusieurs fois été montré que les α-glucosidases pouvaient être actives sur d'autres types de liaisons interglycosidiques comme, par exemple, les α-(1,6) (Alarico *et al.*, 2008; Kitamura *et al.*, 2008; Kang *et al.*, 2009). Ainsi, d'autres tests enzymatiques ont été réalisés sur des substrats variés - comme par exemple le sucrose, le trehalose, l'isomaltose, l'isomaltotriose, le palatinose, le panose - dont les types de liaisons interglycosidiques sont répertoriés dans le Tableau 13.

Sucre	Nombre d'oses	Liaison
Isomaltose	2	Glc α-(1,6) Glc
Isomaltotriose	3	Glc α-(1,6) Glc
Trehalose	2	Glc α-(1,1)-α-Glc
Kojibiose	2	Glc α-(1,2) Glc
Nigerose	2	Glc α-(1,3) Glc
Panose	3	Glc α-(1,6) Glc α-(1,4) Glc
Saccharose	2	Glc α-(1,2)-β-Fru
Turanose	2	Glc α-(1,3) Fru
Palatinose	2	Glc α-(1,6) Fru
Lactose	2	Gal β-(1,4) Glc
Melibiose	2	Gal α-(1,6) Glc
Raffinose	3	Gal α-(1,6) Glc α-(1,2)-β-Fru
Stachyose	4	Gal α-(1,6) Gal α-(1,6) Glc α-(1,2)-β-Fru

Tableau 13 : Récapitulatif des oligosaccharides testés avec rappel du nombre d'unités osidiques et du type de liaisons les composant. Glc : glucose, Fru : fructose, Gal : galactose.

Il s'avère que l'enzyme n'hydrolyse aucun autre substrat que les maltooligosaccharides et apparait donc très spécifique (Figure 58) de la liaison glucosidique α-(1,4).

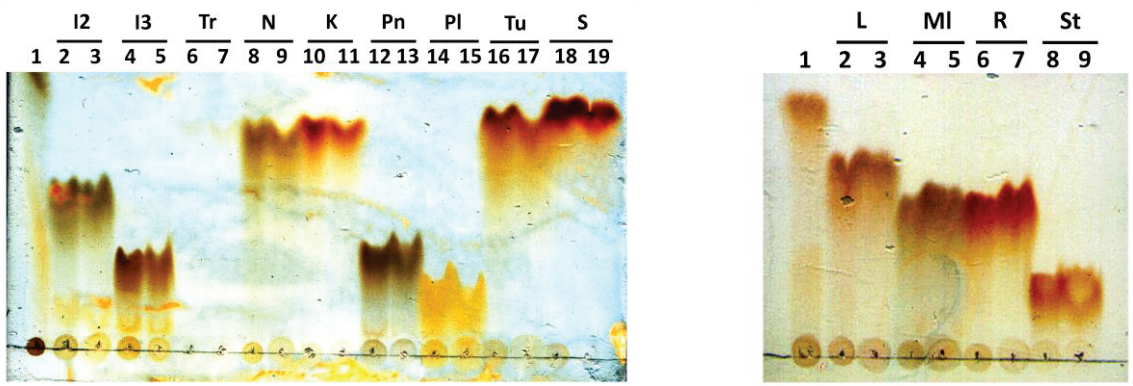


Figure 58 : Analyse par CCM après réaction enzymatique de 1403aglu sur différents oligosaccharides. *A gauche.* Piste 1 : glucose témoin. Pistes 2 et 3 : isomaltose en absence ou présence d'enzyme, respectivement. Piste 4 et 5 : isomaltotriose en absence ou présence d'enzyme, respectivement. Pistes 6 et 7 : trehalose en absence ou présence d'enzyme, respectivement. Pistes 8 et 9 : nigerose en absence ou présence d'enzyme, respectivement. Pistes 10 et 11 : kojibiose en absence ou présence d'enzyme, respectivement. Pistes 12 et 13 : panose en absence ou présence d'enzyme, respectivement. Pistes 14 et 15 : palatinose en absence ou présence d'enzyme, respectivement. Pistes 16 et 17 : turanose en absence ou présence d'enzyme, respectivement. Pistes 18 et 19 : saccharose en absence ou présence d'enzyme, respectivement. ***A droite.*** Piste 1 : glucose témoin. Pistes 2 et 3 : lactose en absence ou présence d'enzyme, respectivement. Piste 4 et 5 : melibiose en absence ou présence d'enzyme, respectivement. Pistes 6 et 7 : raffinose en absence ou présence d'enzyme, respectivement. Pistes 8 et 9 : stachyose en absence ou présence d'enzyme, respectivement.

5. *Stabilité de 1403aglu dans le temps*

Deux mois après la production, la purification, le clivage par la thrombine et le stockage à 4°C, la protéine a été analysée par SDS-PAGE. Il s'avère qu'une bande de masse moléculaire inférieure à celle de la protéine d'intérêt apparaissait sur le gel. Le résultat obtenu après analyse par LC/MS/MS montre qu'il s'agit de 1403aglu tronquée d'une partie de sa région C-terminale, soit une perte de 170 résidus. Bien qu'aucun site de clivage à la thrombine ne soit présent dans la séquence, nous avons déposé sur gel une fraction de la protéine du même lot de production/purification mais qui n'a pas été mis en contact avec la thrombine. Ainsi, de la même manière, la bande de plus bas poids moléculaire apparait. Ces observations disculpent donc la thrombine.

B. Détermination des constantes catalytiques

En utilisant la méthode directe de détection de produit libéré (protocole détaillé dans le chapitre « Matériel et Méthodes »), nous avons établi pour le *p*NPG une constante catalytique (k_{cat}) de 41±7 min^{-1} et une constante de Michaelis (K_m) de 36,5±8,5 mM pour une efficacité catalytique (k_{cat}/K_m) de 1,12 $min^{-1}.mM^{-1}$. Ces valeurs, déterminées par régression non linéaire grâce au serveur http://zunzun.com, montrent une affinité et une efficacité catalytique vraiment très faibles. Des mesures ont également été réalisées pour essayer de déterminer les constantes de l'enzyme pour ses substrats naturels. Cette fois-ci, il n'a pas été possible d'établir des vitesses initiales cohérentes et donc d'en déduire des cinétiques et les constantes associées. Au regard de ces constats, on peut présumer que la protéine est soit mal repliée ou bien qu'elle exige un partenaire protéique pour être entièrement et optimalement active.

V. État oligomérique de la protéine en solution

L'état oligomérique de 1403aglu (clivée par la thrombine) en solution a été analysé sur SDS-PAGE après pontage covalent au BS^3 (agent de pontage homo-bifonctionnel). Cette molécule de 570 Da réagit en formant des liaisons covalentes avec les amines primaires des chaines latérales des lysines et/ou avec les extrémités N-terminales des protéines. Grâce à ses propriétés, cette molécule est utilisée pour identifier, sur gel dénaturant, des interactions protéiques et même « figer » celles qui sont faibles ou transitoires. Après les réactions de pontage menées sous plusieurs conditions, l'analyse des échantillons par SDS-PAGE montrent que la protéine en présence de BS^3 migre à la même hauteur que le témoin négatif sans agent pontant. Etant donné que tous les échantillons déposés ont une masse moléculaire proche de 60 kDa et que la masse moléculaire d'un monomère de 1403glu est

déterminée à 65 kDa, ceci indique que la protéine en solution est donc monomérique (Figure 59). Il a toutefois été observé l'apparition d'une bande « parasite » correspondant à une masse d'environ 70 kDa dont l'importance est proportionnelle au temps d'incubation et à la concentration de BS^3 utilisée. Ceci est sans doute dû à la présence de nombreuses lysines à la surface de l'enzyme comme le démontre le modèle obtenu par I-TASSER (Figure 59). Elles sont donc susceptibles de réagir avec l'agent pontant, ce qui augmenterait la masse apparente de l'enzyme sur le gel.

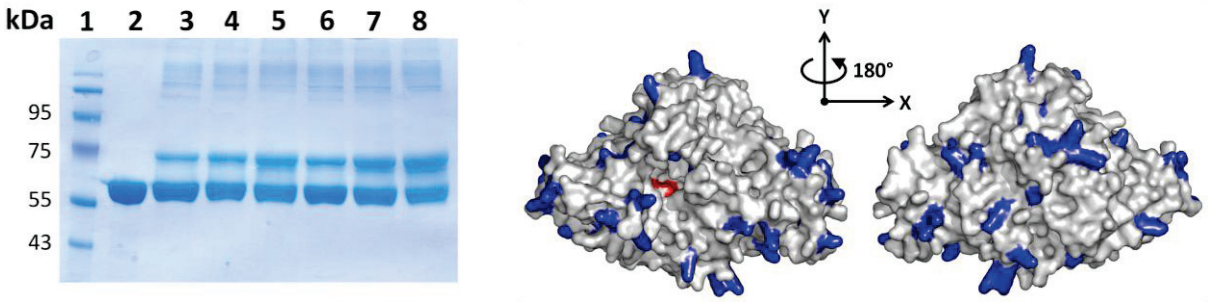


Figure 59 : A gauche : Analyse par SDS-PAGE après pontage covalent au BS^3. Piste 1 : marqueurs de masse moléculaire, piste 2 : 1403aglu sans agent de pontage, piste 3 : 1403aglu avec 0,5 mM BS^3 incubée 45min, piste 4 : 1403aglu avec 1 mM BS^3 incubée 45 min, piste 5 : 1403aglu avec 2 mM BS^3 incubée 45 min, piste 6 : 1403aglu avec 0,5 mM BS^3 incubée 60 min, piste 7 : 1403aglu avec 1 mM BS^3 incubée 60min, piste 8 : 1403aglu avec 2 mM BS^3 incubée 60 min. Les réactions de pontage ont été réalisées à température ambiante et avec 25 µM 1403aglu (purifiée et clivée par la thrombine). **A droite : Représentation de la surface moléculaire du modèle de 1403aglu généré par I-TASSER.** Les lysines et les résidus catalytiques sont colorés en bleu et rouge, respectivement.

VI. Étude structurale

A. Cristallogenèse

La stratégie adoptée pour les expériences de cristallogenèse a été de faire un criblage large des conditions grâce à plusieurs kits commerciaux proposés par MDL, Qiagen ou Hampton Research.

Les premiers essais ont été réalisés en goutte assise selon la méthode de diffusion de vapeur grâce au robot Mosquito® (TTP Labtech) dans des boites 96 puits TTPLabtech à 3 cupules (TTP Labtech). Cet appareil permet de prélever de très petits volumes (jusqu'à 100 nl) de façon précise et reproductible. Ainsi, près de 3000 conditions ont été testées en utilisant la protéine préalablement concentrée à 4 et 6 mg/ml dans un tampon MOPS 100 mM pH 7. Des gouttes de 300 nl (1:1, $v_{protéine}$:$v_{solution}$) ont été réalisées avec 100 µl de solution de cristallisation dans le puits. Ce premier criblage à 19°C n'a donné aucun résultat positif.

Une deuxième série d'essais à 19°C a été réalisée, cette fois-ci avec des gouttes de 400 nl (3:1) dans les mêmes 3000 conditions, toujours avec la protéine en tampon MOPS 100 mM pH7. Des cristaux (Figure 60) bipyramidaux aux sommets tronqués sont apparus dans la condition 63 du kit $AmSO_4$ de Qiagen ($AmSO_4$ 2,4 M, MES 100 mM pH6). Des gouttes suspendues de 1,2 µl (3:1) ont ensuite été réalisées dans le but d'obtenir de plus gros cristaux. Cette étape s'est révélée reproductible. Ensuite, plusieurs cryoprotégeants comme le glycérol, l'éthylène glycol, le sucrose et le glucose, à différentes concentrations, ont été testés. Les cristaux, de nature fragile, se dissolvent rapidement après ajout de l'un de ces trois derniers. Des cristaux sans cryoprotection et ceux ayant résisté au trempage dans le glycérol ont alors été congelés et testés sur ligne synchrotron. Seules quelques tâches de diffraction à très basse résolution (30 Å) sont visibles sur les clichés indiquant, malgré tout, qu'il s'agit bien de cristaux de protéine. Le très faible pouvoir de diffraction de ces cristaux peut signifier plusieurs choses : soit les cristaux possèdent un très haut pourcentage de solvant, soit l'empilement cristallin est peu ordonné, ou bien encore la protéine est sous plusieurs états oligomériques ou n'est pas uniformément et correctement repliée.

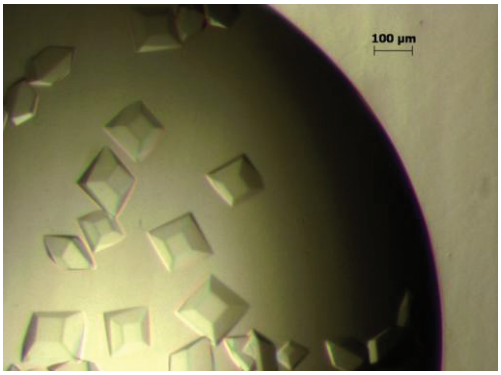


Figure 60 : Cristaux bipyramidaux de 1403aglu obtenus dans la condition 63 de cristallisation du kit $AmSO_4$ de Qiagen ($AmSO_4$ 2,4 M, MES 100 mM pH 6).

Une autre forme cristalline, cette fois de morphologie hexagonale (Figure 61), est apparue dans la condition 93 du kit Crystal Screen (Hampton Research) : NiCl 0,01 M, Tris 100 mM pH 8,5, PEG MME 20%. Cette fois encore, seules quelques tâches de diffraction aux très basses résolutions (30 Å) étaient visibles sur les clichés confirmant qu'il s'agit de cristaux de protéines, mais toujours avec un pouvoir de diffraction très faible.

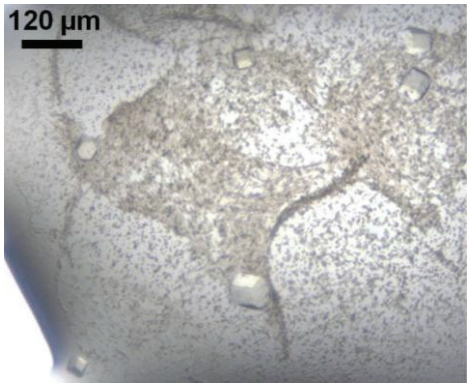


Figure 61 : Cristaux de morphologie hexagonale de 1403aglu obtenus dans la condition de cristallisation 93 du kit Crystal Screen d'Hampton Research (NiCl 0,01 M, Tris 100 mM pH 8,5, PEG MME 20%).

Trois autres séries de criblage ont cette fois été réalisées en changeant le tampon dans lequel se trouve la protéine au profit de l'HEPES 100 mM pH 7 à 19°C ou de l'HEPES 20 mM pH 7, NaCl 20 mM à 19°C ou 4°C, en absence ou encore en présence de déoxynojirimycine 1 mM. Mais jusqu'à présent, aucune amélioration de diffraction des cristaux obtenus n'a été observée.

Pour essayer d'expliciter ces problèmes de diffraction récurrents et vérifier l'intégrité du repliement de notre protéine, des études par dichroïsme circulaire ont été réalisées.

B. Dichroïsme circulaire

L'analyse de spectres de dichroïsme circulaire aux UV lointains permet d'estimer les proportions de structures secondaires au sein d'une protéine. Les résultats obtenus après un traitement des spectres sont jugés fiables lorsque l'analyse par les trois programmes « SELCON », « CONTIN » et « CDSSTR » sont similaires (Sreerama & Woody, 2000).

Plusieurs spectres de dichroïsme ont été acquis pour la protéine se trouvant dans différents tampons : HEPES 100 mM pH 7 et HEPES 20 mM pH 7 - NaCl 20 mM. Les données ont ensuite été analysées par les trois programmes pré-cités, disponibles *via* le serveur DichroWeb (http://dichroweb.cryst.bbk.ac.uk) permettant la déconvolution des spectres. Plusieurs séries de mesures n'ont pas abouti à des résultats, l'analyse étant avortée au cours des étapes de traitement des données. Toutefois, parmi tous les spectres acquis, certains ont pu être traités et validés par les trois méthodes. Ainsi, ces résultats tendent vers des proportions de structures secondaires suivantes : 6 à 10% d'hélices α, 30 à 40% de brins β, 15 à 25% de tours et 30 à 35% de structures désordonnées. Au vu des pourcentages normalement attendus d'après le modèle (environ 35% d'hélices α et 30% de brins β) mais et des difficultés pour acquérir des spectres de dichroïsme circulaire de qualité, il est fort probable que la protéine ne soit pas correctement repliée.

Chapitre 3 : Étude de l'α-glucosidase de *Shewanella* sp. ANA-3

I. Biologie moléculaire

Nous avons cloné le gène d'intérêt dans le vecteur pET15b apportant une étiquette His_6 N-terminale clivable à la thrombine à partir du génome complet de *Shewanella* sp. ANA-3, déjà accessible au laboratoire. En utilisant comme matrice cet ADN génomique, nous avons amplifié par PCR le gène *SHWaglu* grâce à des amorces construites de sorte à inclure les sites de restriction compatibles pour le clonage dans le vecteur choisi, à savoir *Nde*I en 5' en amont du gène et *Eco*RI en 3'après le codon stop (Tableau 14).

NOM	SEQUENCE DE L'AMORCE 5'→ 3'
OL-Direct-SHW	GTCGTC**CATATG**TTCAGCAATCCTCTTACCCACGTACG
OL-Reverse-SHW	G**GAATTC**TTATTACAGCGCTTTAAAGCGGATCGCTGTG

Tableau 14 : Amorces utilisées pour l'amplification par PCR du gène *SHWaglu*. Les sites de restriction *Nde*I et *Eco*RI sont surlignés en gris.

Après ligature et transformation dans la souche *E. coli* DH5α, les plasmides ont été extraits et digérés par une enzyme de restriction appropriée (*Eco*RI) pour vérifier la présence de l'insert. En parallèle, une PCR a également été réalisée sur ces mêmes vecteurs afin de confirmer le choix des plasmides recombinants (Figure 62). Ainsi, avec une taille de vecteur linéarisé proche de 8000 pb et un fragment de près de 2000 pb fortement amplifié semblable à la taille du gène d'intérêt, les clones 1, 3 et 4 ont été sélectionnés pour être envoyés en séquençage et valider leurs séquences. Une fois confirmés, les vecteurs sélectionnés ont été transformés dans une souche d'expression *E. coli* BL21 (DE3).

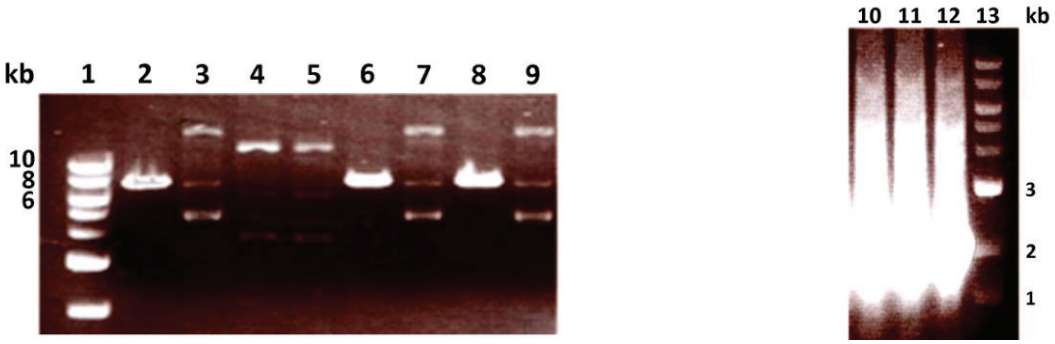


Figure 62 : A gauche, digestion par *Eco*RI des vecteurs pET15b pouvant contenir le gène *SHWaglu*. A droite, PCR du gène *SHWaglu* sur les vecteurs sélectionnés. Piste 1 et 13 : marqueurs de masse moléculaire. De 2 à 9, pistes paires : vecteurs digérés, pistes impaires : vecteurs non digérés. Pistes 2 et 3 : clone 1, pistes 4 et 5 : clone 2, pistes 6 et 7 : clone 3, pistes 8 et 9 : clone 4. Piste 10 : clone 1, piste 11 : clone 3, piste 12 : clone 4.

II. Expression et purification de SHWaglu

Des tests d'expression en milieu LB ont permis de démontrer que la température d'induction de 25°C était préférable à une température de 37°C (résultats non montrés) car la protéine était exprimée de façon plus soluble à température inférieure.

Ainsi, le protocole retenu pour une surexpression optimisée de l'enzyme soluble consiste à induire la culture bactérienne à DO_{600} ≈ 0,6 avec 0,5 mM d'IPTG à 25°C toute une nuit dans un volume de 300 ml de LB. Le lendemain, les bactéries sont culottées par centrifugation (5000 *g* pendant 12 minutes). Les cellules sont ensuite resuspendues dans un tampon de lyse MES 100 mM pH 6,5 contenant du lysozyme à hauteur de 10 mg/ml et de la ribonucléase bovine. Après incubation sur glace, le mélange est soniqué 3 fois pendant 30 secondes. Le lysat est alors centrifugé pour culotter tous les débris cellulaires et filtré à 0,45 µm. L'extrait protéique obtenu est alors prêt pour la première étape de purification.

Comme la protéine exprimée avec le système pET15b porte une étiquette His_6 du coté N-terminal, il est possible d'effectuer une chromatographie d'affinité sur colonne de nickel IMAC dont le protocole est décrit dans le chapitre « Matériel et Méthodes ». Lorsque l'on observe l'élution de protéines, le gradient est alors mis en pause pour obtenir un échantillon homogène en concentration d'imidazole et limiter une possible élution concomitante d'autres protéines. La protéine d'intérêt est éluée à environ 70 mM d'imidazole (Figure 63).

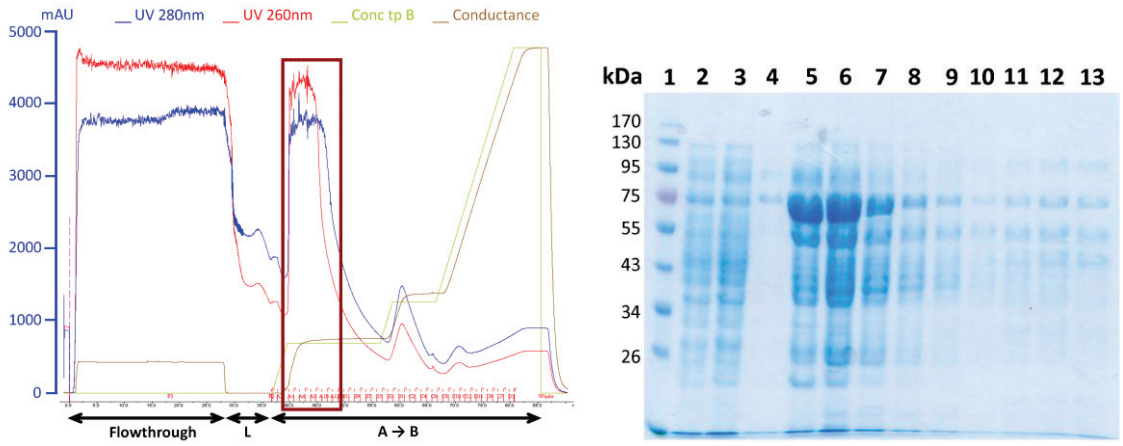


Figure 63 : Purification de SHWaglu par chromatographie d'affinité. A gauche, chromatogramme : les courbes bleue, rouge, verte et marron représentent respectivement l'absorbance à 280 nm, à 260 nm, le gradient d'imidazole et la conductance. *Flowthrough*, L et A→B correspondent respectivement à l'injection sur colonne, au lavage et au gradient d'imidazole. Le cadre rouge représente les fractions contenant la protéine d'intérêt. A droite, analyse sur SDS-PAGE des fractions collectées. Piste 1 : marqueurs de masse moléculaire, piste 2 : surnageant avant purification, piste 3 : *flowthrough*, pistes 4 à 9 : premier pic élué par le gradient, pistes 10 à 13 : deuxième pic élué par le gradient.

A ce stade, la protéine n'étant pas pure, une deuxième étape de purification par chromatographie d'exclusion (sur colonne Superdex S-200, GE Healthcare) permettant de séparer les protéines de différents poids moléculaires a ensuite été testée. Par cette méthode, un seul pic est élué sur le chromatogramme (résultats non montrés), la protéine d'intérêt n'étant pas séparée des contaminants (Figure 64).

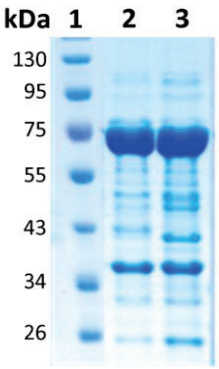


Figure 64 : Analyse par SDS-PAGE du pic élué par chromatographie d'exclusion. Piste 1 : marqueurs de masse moléculaire. Piste 2 : dépôt de fraction obtenue après chromatographie d'exclusion. Piste 3 : échantillon avant la chromatographie d'exclusion.

Parmi les méthodes de purification les plus couramment utilisées, on retrouve également la chromatographie échangeuse d'ions permettant de séparer les protéines en fonction de leur charge globale. Dans notre cas, le choix s'est porté sur une colonne échangeuse d'anions (Mono-Q™ 5/50, GE Healthcare). Le protocole de cette chromatographie est également décrit dans le chapitre « Matériel et Méthodes ». Durant

cette étape de purification, la protéine d'intérêt est entièrement retrouvée dans la fraction qui n'a pas été fixée sur la colonne (« *flowthrough* »), contrairement à tous les contaminants retenus puis libérés ultérieurement au cours du gradient de sel (Figure 65). Une protéine quasi pure (à environ 98%) a ainsi pu être obtenue grâce à deux techniques de chromatographie successives : celles d'affinité et échangeuse d'anions.

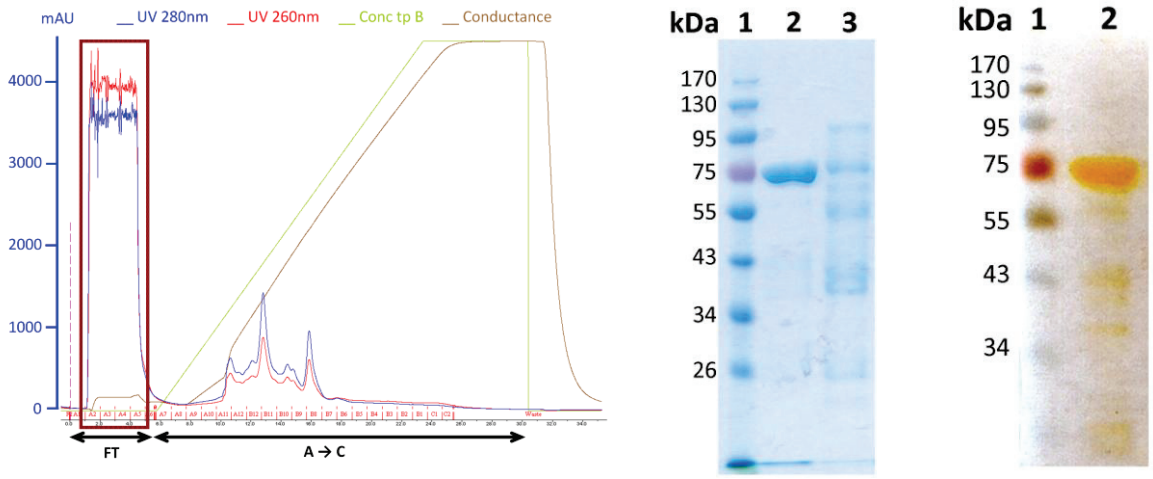


Figure 65 : Purification de SHWaglu par chromatographie échangeuse d'anions. A gauche, chromatogramme : les courbes bleue, rouge, verte et marron représentent respectivement l'absorbance à 280 nm, à 260 nm, le gradient de NaCl et la conductance. Le cadre rouge représente les fractions contenant la protéine d'intérêt. Au milieu, analyse par SDS-PAGE et coloration au bleu de Coomassie des fractions collectées. Piste 1 : marqueurs de masse moléculaire, piste 2 : flowthrough (FT), piste 3 : protéines éluées lors du gradient. Au milieu, analyse par SDS-PAGE et coloration au nitrate d'argent de la fraction contenant la protéine d'intérêt.

A partir de là, il a été possible de concentrer la protéine jusqu'à 5,5 mg/ml par l'intermédiaire de concentrateurs (Vivaspin®, Sartorius) ayant une membrane PES et un seuil de coupure de 50 kDa puis de lancer les premiers tests de cristallogenèse. En parallèle, une étude fonctionnelle de l'enzyme a été réalisée.

III. État oligomérique de la protéine en solution

L'état oligomérique de SHWaglu en solution a été analysé à l'aide de deux techniques : la chromatographie d'exclusion et le pontage covalent au BS^3. Avec un volume de rétention de 14,4 ml sur une colonne Superdex S-200, SHWaglu possède une masse moléculaire apparente d'environ 77 kDa correspondant à la masse d'un monomère en solution. Une analyse par SDS-PAGE après pontage covalent au BS^3 (agent de pontage homo-bifonctionnel) a permis de confirmer ce résultat (Figure 66). En effet, grâce à ses propriétés, cette molécule est utilisée pour identifier, sur gel dénaturant, des interactions protéiques et même « figer » celles qui sont faibles ou transitoires. Après les réactions de pontage menées

sous plusieurs conditions, l'analyse des échantillons par SDS-PAGE montrent que la protéine en présence de BS^3 migre à la même hauteur que le témoin négatif en absence de l'agent pontant. Etant donné que tous les échantillons déposés ont une masse moléculaire d'environ 75 kDa et que la masse moléculaire d'un monomère de SHWglu est déterminée à 80 kDa, ceci indique que la protéine en solution est bien monomérique.

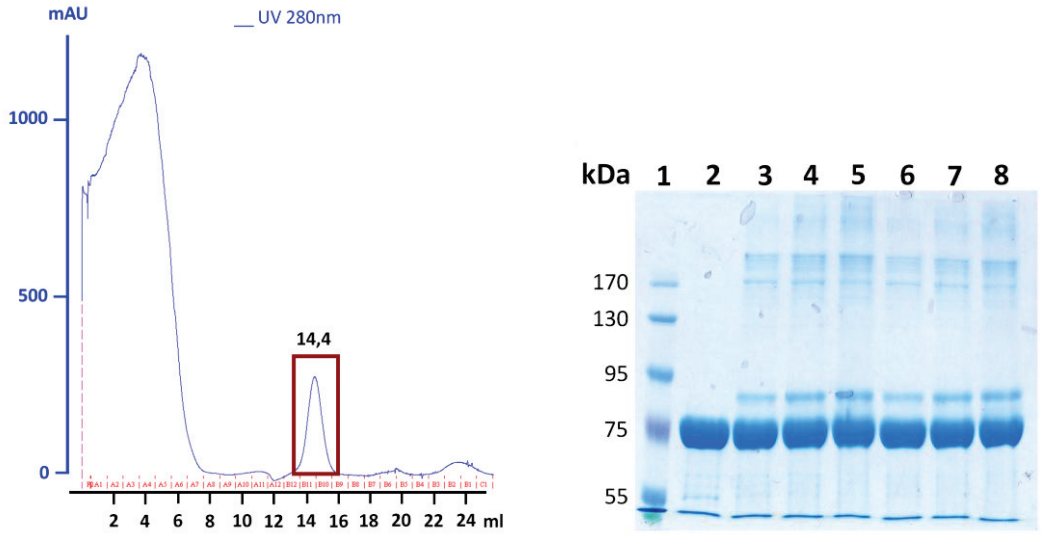


Figure 66 : Analyse de l'état oligomérique de SHWaglu. *A gauche* : par chromatographie d'exclusion sur colonne Superdex S-200, la protéine est éluée à 14,4 ml correspondant à une MM de 77 kDa. La courbe bleue représente l'absorbance à 280 nm. *A droite* : analyse par SDS-PAGE après pontage covalent au BS^3. Piste 1 : marqueurs de masse moléculaire, piste 2 : SHWaglu sans agent de pontage, piste 3 : SHWaglu avec 0,5 mM BS^3 incubée 45min, piste 4 : SHWaglu avec 1 mM BS^3 incubée 45 min, piste 5 : SHWaglu avec 2 mM BS^3 incubée 45 min, piste 6 : SHWaglu avec 0,5 mM BS^3 incubée 60 min, piste 7 : SHWaglu avec 1 mM BS^3 incubée 60min, piste 8 : SHWaglu avec 2 mM BS^3 incubée 60 min. Les réactions de pontage ont été réalisées à température ambiante et avec 25 µM SHWaglu purifiée.

IV. Étude fonctionnelle

A. Caractérisation préliminaire

Les tests préliminaires ont été réalisés avec un chromophore (le *p*NPG) sur un surnageant protéique non purifié. Un témoin négatif a été réalisé avec un extrait brut protéique d'une culture bactérienne qui n'a pas été induite à l'IPTG confirmant l'absence d'activité α-glucosidase et donc de libération de nitrophénol. L'enzyme a été testée sous forme libre mais aussi sous forme immobilisée dans des billes d'alginate. Aucune activité n'a été observée dans le cas de l'immobilisation de l'extrait cellulaire soluble dans les billes d'alginate. Ceci est très probablement dû au relargage des protéines hors du maillage des billes, comme l'ont déjà démontré des études précédentes (Tanaka *et al.*, 1984). La stratégie a donc été d'immobiliser les cellules bactériennes ayant exprimé la protéine d'intérêt.

1. *Impact de la concentration cellulaire dans les billes d'alginate*

L'activité catalytique relative a été mesurée en fonction de la concentration cellulaire dans les billes d'alginate. La concentration optimale s'étend sur une gamme de 30 à 36 g/l. Au-delà de ces concentrations, l'activité se réduit (Figure 67), probablement du fait de contraintes stériques ou de limitations de diffusion au sein de la matrice de la bille (Le-Tien *et al.*, 2004).

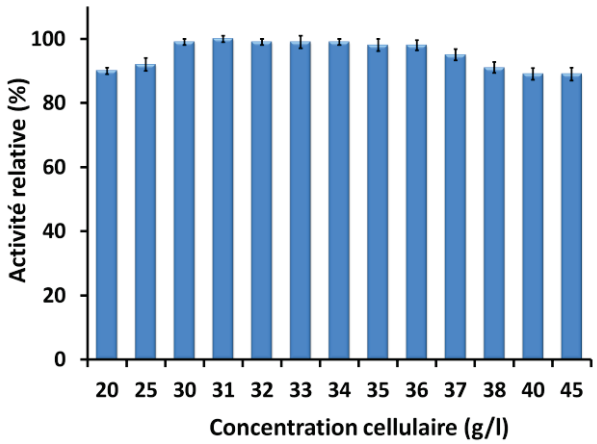


Figure 67 : Profil des activités relatives de l'enzyme immobilisée dans des billes d'alginate en fonction de la concentration cellulaire. Les barres représentent les déviations standards de trois mesures indépendantes.

2. *Détermination du pH optimal*

Des tests enzymatiques sur l'α-glucosidase libre et immobilisée ont été réalisés à différents pH dans une gamme allant de 4,5 à 9. Il en ressort que le pH optimal de l'enzyme libre est de 6,5 et que l'activité enzymatique chute très rapidement de part et d'autre de ce pH avec toutefois une conservation de plus de 50% à pH 6 et 7. L'activité détectée est quasi nulle aux pH inférieurs à 5 ou supérieurs à 7,5.

Le pH optimal de l'enzyme, lorsqu'elle est immobilisée, est décalé à pH 7 et celle-ci conserve plus de 60% de son activité entre pH 5 et 8,5. De part et d'autre de cette gamme, l'activité enzymatique chute assez brutalement (Figure 68). Ainsi, l'immobilisation de l'enzyme permet de maintenir de façon significative son activité autour de son pH optimum comparativement à la forme libre.

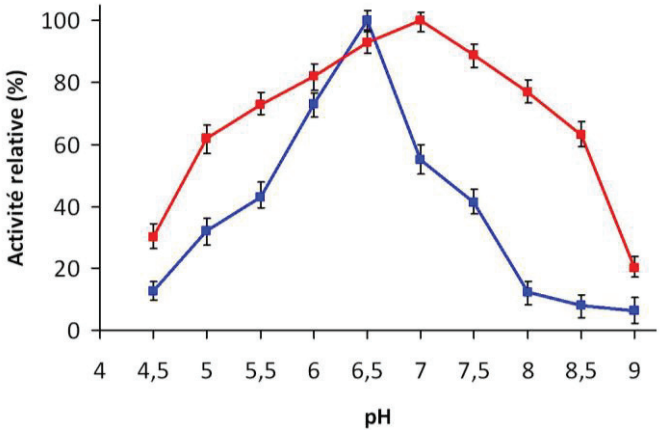


Figure 68 : Profil d'activité relative de l'enzyme libre et immobilisée en fonction du pH. Bleu : enzyme libre, rouge : enzyme immobilisée dans des billes d'alginate. Les barres représentent les déviations standards de trois mesures indépendantes.

3. Détermination de la température optimale

Des tests enzymatiques sur enzyme libre et immobilisée ont été réalisés entre 4°C et 65°C afin de déterminer la température optimale d'activité. Nous constatons que l'enzyme possède un maximum d'activité à 50°C avec une diminution relativement rapide de l'activité aux températures supérieures. Il est à noter que cette température optimale d'activité est largement supérieure à la température optimale de croissance de *Shewanella* sp. ANA-3 ou de toutes les autres souches *Shewanella* étudiées à ce jour (Karpinets *et al.*, 2010). En dessous de la température optimale, l'activité diminue progressivement jusqu'à 50% de l'activité maximale à partir de 30°C et cette activité relative est conservée jusqu'à 4°C. Jusqu'à présent, à notre connaissance, une seule autre α-glucosidase (chez *Aspergillus niger*) avait la caractéristique d'être active à 50% à 4°C (Brizova *et al.*, 1992). L'immobilisation permet de conserver une activité maximale jusqu'à 60°C (Figure 69).

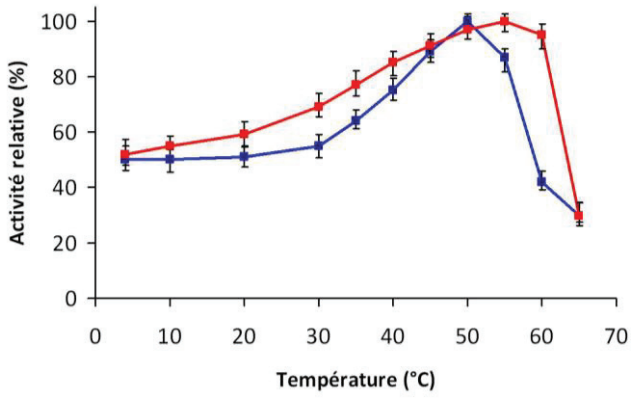


Figure 69 : Profils d'activité relative de l'enzyme libre et immobilisée en fonction de la température. Bleu : enzyme libre, rouge : enzyme immobilisée dans des billes d'alginate. Les barres représentent les déviations standards de trois mesures indépendantes.

4. *Effets des ions divalents sur l'activité catalytique*

Dans un premier temps, nous avons étudié les effets de différents cations divalents sur l'activité enzymatique de l'extrait brut. L'activité est augmentée de 230% en présence de Ca^{2+} et il apparait que les ions Mg^{2+} et Mn^{2+} sont également activateurs avec, respectivement, 147% et 122% d'activité relative par rapport l'activité mesurée sans ajout d'ions (Figure 70). Les autres cations testés sont tous inhibiteurs, les ions Zn^{2+} abolissant toute activité.

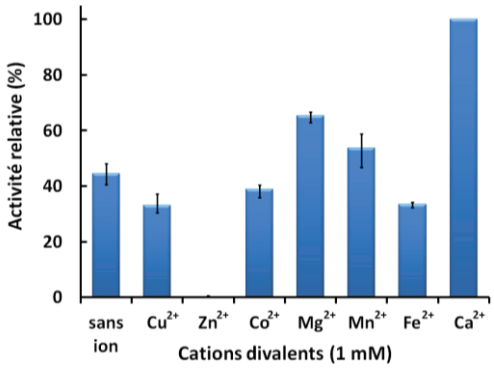


Figure 70 : Effets des cations divalents sur l'activité enzymatique de l'extrait brut. Les ions ont été ajoutés à hauteur de 1 mM. L'activité relative est normalisée par rapport à l'activité mesurée en présence de Ca^{2+}. « Sans ion » indique l'activité mesurée sans ajout de cations. Les barres représentent la déviation standard de trois mesures indépendantes.

Dans un second temps, l'enzyme purifiée a été traitée à l'EDTA. Dans ces conditions appauvries en cations, il apparait que SHWaglu perd la quasi-totalité de son activité. Comme attendu, l'addition dans le milieu réactionnel de 1 mM de $MgCl_2$, de $MnCl_2$ ou plus notablement de $CaCl_2$ restaure son activité (Figure 71).

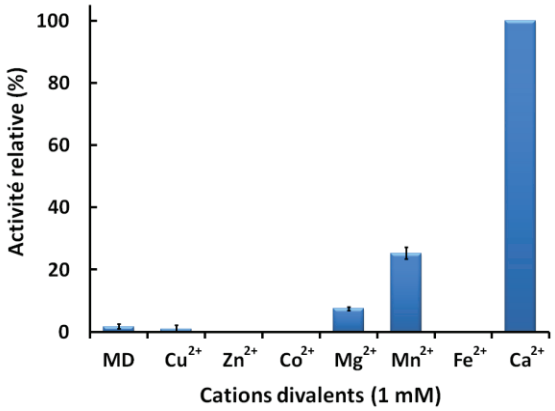


Figure 71 : Effets des cations divalents sur l'activité de l'enzyme purifiée, prétraitée à l'EDTA et dialysée deux fois dans 100 mM MES-milliQ (pH 6,5). Les ions ont été ajoutés à hauteur de 1 mM. L'activité relative est normalisée par rapport à l'activité mesurée en présence de Ca^{2+}. « MD » indique l'activité mesurée dans la solution déplétée en cations. Les barres représentent la déviation standard de trois mesures indépendantes.

Différentes concentrations de $CaCl_2$ testées indiquent que l'enzyme recouvre progressivement son activité et quasi totalement à 2 mM de $CaCl_2$. La forme sigmoïdale de la courbe observée indique que SHWaglu serait une enzyme activée de façon allostérique. Une analyse des données par régression non linéaire (http://zunzun.com) nous a permis de calculer un coefficient de Hill de 1,6, ceci traduisant une coopérativité positive liée à l'ajout de ces ions.

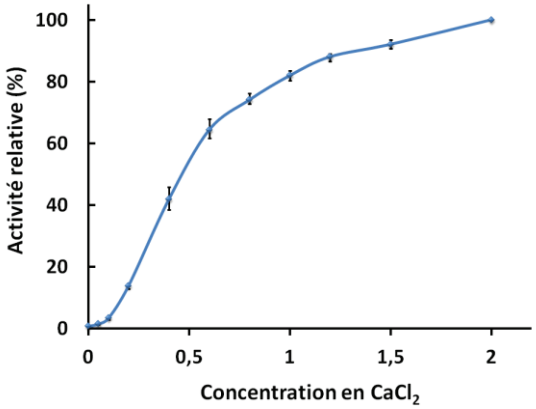


Figure 72 : Effets de différentes concentrations en ion Ca^{2+} sur l'activité de l'enzyme purifiée, prétraitée à l'EDTA et dialysée deux fois dans 100 mM MES-milliQ (pH 6,5). L'activité relative est normalisée par rapport à l'activité mesurée en présence de 2 mM Ca^{2+}. Les barres représentent la déviation standard de trois mesures indépendantes.

5. Le rôle des ions chlorure

Une analyse complémentaire a été réalisée dans le but de déterminer si l'augmentation de l'activité par le $CaCl_2$ ou le $MgCl_2$ dépendait sans ambiguïté de ces cations. Dans ce cadre, des dosages d'activité ont été réalisés en utilisant des composés autres que des sels de chlorure ($MgSO_4$, $CaSO_4$ et $Ca(CH_3COO)_2$). Ainsi, nous avons pu constater que ceux-ci n'entraînaient aucune restauration de l'activité catalytique par rapport à l'enzyme traitée à l'EDTA. L'ion chlore apparait donc indispensable à l'activité hydrolytique de l'enzyme.

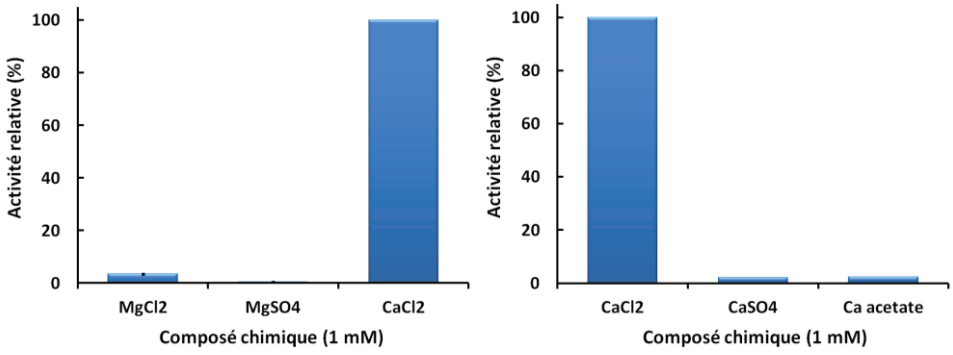


Figure 73 : Effets des ions Cl^- sur l'activité de l'enzyme purifiée, prétraitée à l'EDTA et dialysée deux fois dans 100 mM MES-milliQ (pH 6,5). Les ions ont été ajoutés à hauteur de 1 mM. L'activité relative est normalisée par rapport à l'activité mesurée en présence de Ca^{2+}. Les barres représentent la déviation standard de trois mesures indépendantes.

Cette étude de l'effet des ions sur l'activité enzymatique de SHWaglu montre très clairement que les ions calcium jouent un rôle clé pour cette dernière. Ce constat est en accord avec ce qui a été démontré dans des études précédentes suggérant que l'ion Ca^{2+} renforce le caractère du résidu catalytique acide durant l'hydrolyse (Gloster *et al.*, 2008; Kitamura *et al.*, 2008). Ce qui était plus inattendu est le besoin essentiel de cette enzyme psychrotolérante en ion Cl^-. Ces constations sont toutefois cohérentes avec le milieu naturel de *Shewanella* sp. ANA-3 qui est une eau saumâtre, avec la présence de calcium et de chlorure à hauteur de 400 mg/l et de 17-18 g/l, respectivement (Page, 1928). Il a également déjà été rapporté que le chlore était un activateur allostérique chez certaines enzymes issues de microorganismes adaptés au froid comme, par exemple, l'α-amylase de *Pseudoalteromonas haloplanktis* (Feller *et al.*, 1994; Feller *et al.*, 1996; Aghajari *et al.*, 1998; Aghajari *et al.*, 2002).

6. Thermostabilité

Une étude de la thermostabilité de l'enzyme libre et immobilisée dans des billes d'alginate a été réalisée à 45°C, 50°C et 55°C (Figure 74). L'enzyme libre parait entièrement stable au moins une heure à 45°C. Son activité résiduelle diminue de moitié au bout d'une heure à 50°C et est réduite de plus de deux tiers pour ce même temps d'incubation à 55°C. En comparaison, l'immobilisation de l'enzyme dans des billes d'alginate permet de conserver entièrement l'activité hydrolytique après une heure d'incubation à 50°C.
Ainsi, l'immobilisation de l'enzyme permet d'améliorer sa stabilité à des températures relativement importantes.

Toutefois, une chute brutale de l'activité est constatée lorsqu'elle est immobilisée et incubée à 55°C; elle pourrait être due à une détérioration des billes d'alginate à cette température.

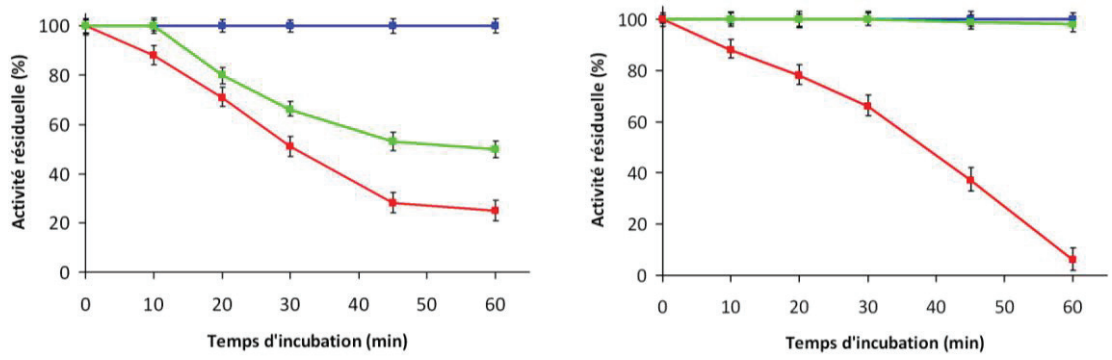


Figure 74 : Profils de l'activité résiduelle mesurée après incubation à différentes températures pendant une heure. *A gauche : incubation de l'enzyme SHWaglu libre, à droite : incubation de l'enzyme SHWaglu immobilisée. Courbe bleue : 45°C, courbe verte : 50°C, courbe rouge : 55°C. Les barres représentent la déviation standard de trois mesures indépendantes.*

7. *Stabilité vis-à-vis du pH*

Par la suite, nous avons étudié la stabilité dans le temps de l'enzyme libre et immobilisée dans des billes d'alginate, à différents pH (5,5, 6, 6,5 et 7) et aux températures limites de stabilité de l'enzyme (45°C et 50°C).

A 50°C, une diminution drastique de l'activité résiduelle est constatée dès 10 minutes d'incubation de l'enzyme libre à pH 5,5, 6 et 7 alors que l'activité est quasiment totalement conservée à pH 6,5 (Figure 75). Une amélioration considérable est constatée lorsque l'enzyme est immobilisée. En effet, avec ce même temps d'incubation, l'enzyme immobilisée révèle un gain d'activité d'au moins 40% pour ces pH comparativement à la forme libre. Toutefois, la perte d'activité demeure très importante au-delà de 10 minutes d'incubation, exception faite pour un pH de 6,5 où l'activité résiduelle est de 95% (contre 50% pour l'enzyme libre).

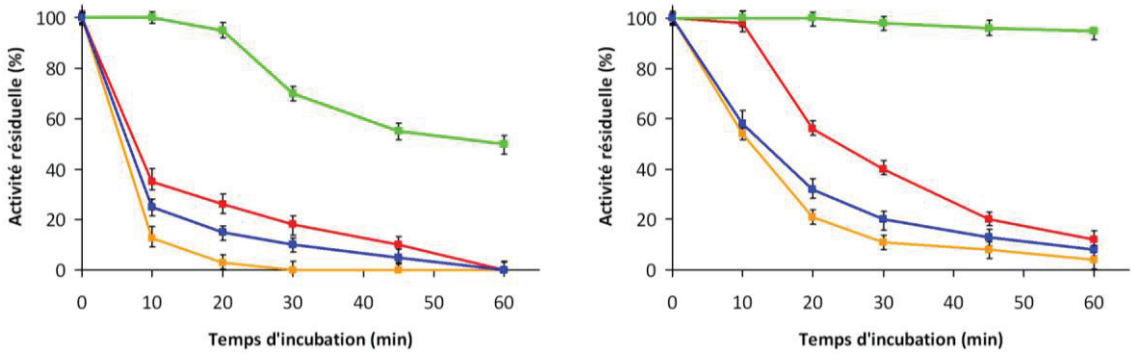


Figure 75 : Profils de l'activité résiduelle mesurée après incubation à différents pH à 50°C pendant une heure. *A gauche : incubation de l'enzyme SHWaglu libre, à droite : incubation de l'enzyme SHWaglu immobilisée. Courbe jaune : pH 5,5, courbe rouge : pH 6, courbe verte : pH 6,5, courbe bleue : pH 7. Les barres représentent la déviation standard de trois mesures indépendantes.*

A 45°C, le profil de l'activité résiduelle montre que l'enzyme est complètement stable pendant une heure à pH 6,5, que ce soit à l'état libre ou immobilisée. Aux autres pH, l'activité résiduelle décroit rapidement lorsque l'enzyme est libre ; pour l'enzyme immobilisée, l'amélioration est flagrante au terme de l'heure d'incubation avec au minimum 40% d'activité résiduelle à pH 5,5 (Figure 76).

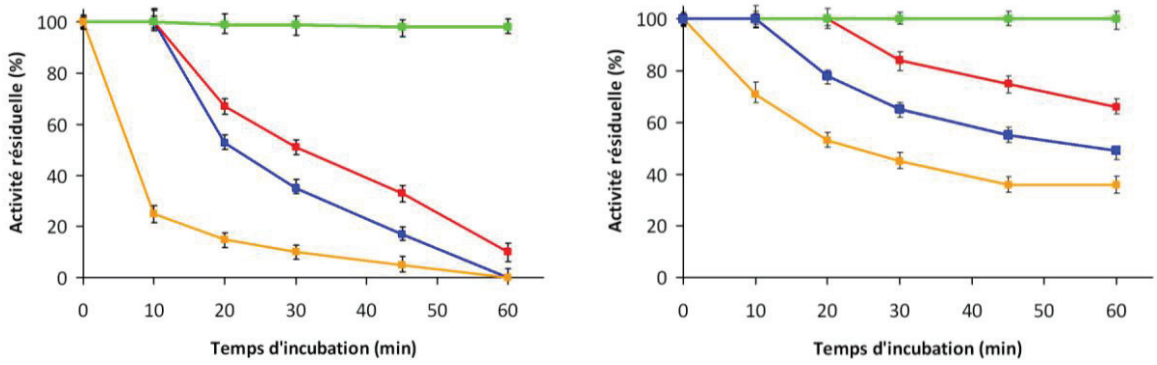


Figure 76 : Profils de l'activité résiduelle mesurée après incubation à différents pH à 45°C pendant une heure. *A gauche : incubation de l'enzyme SHWaglu libre, à droite : incubation de l'enzyme SHWaglu immobilisée. Courbe jaune : pH 5,5, courbe rouge : pH 6, courbe verte : pH 6,5, courbe bleue : pH 7. Les barres représentent la déviation standard de trois mesures indépendantes.*

8. Impacts de l'immobilisation sur l'activité catalytique

La grande amélioration de l'activité relative de l'enzyme immobilisée à tous les pH suggère que SHWaglu, sous cette forme, est moins sensible aux altérations induites par le pH que l'enzyme libre. Ceci est plus notable pour les pH alcalins avec un déplacement du pH optimum de 6,5 à 7. Ce déplacement peut être dû à des interactions secondaires (comme les interactions ioniques, polaires ou des liaisons hydrogène) entre l'enzyme et la matrice polymérique ou encore une augmentation de la concentration locale des ions hydrogène dans le microenvironnement (Munjal & Sawhney, 2002; Kumar *et al.*, 2009).

L'activité conservée pour l'enzyme immobilisée entre 50 et 60°C peut être expliquée par la nécessité d'une plus grande énergie d'activation ou par des changements conformationnels provoqués par la charge et les contraintes physiques de la matrice (Puri *et al.*, 1996). En outre, l'activité à basses températures n'a pas ou peu été modifiée contrairement à la réduction observée dans d'autres études (Puri *et al.*, 1996; Munjal & Sawhney, 2002). La perte en flexibilité causée par les basses températures est sans doute compensée par une meilleure efficacité catalytique de l'enzyme en dessous de 30°C.

Il existe un lien direct entre l'activité et la stabilité d'une enzyme. Les enzymes psychrophiles, reconnues pour être structuralement flexibles sont donc généralement thermolabiles (Feller *et al.*, 1996; Georlette *et al.*, 2004). Ainsi, la perte d'activité de la forme libre de l'α-glucosidase aux hautes températures est très probablement causée par la dénaturation progressive de l'enzyme (Iyer & Ananthanarayan, 2008). Il semblerait que l'emprisonnement de l'enzyme dans des billes d'alginate diminuerait sa dénaturation en limitant sa flexibilité. En outre, l'absorption d'une partie de la chaleur par la matrice pourrait également avoir un rôle coopératif et indirect dans la stabilisation constatée.

9. Criblage des substrats naturels de SHWaglu

Cette étape de la caractérisation enzymatique a été réalisée en utilisant la technique de chromatographie sur couche mince (CCM).

Comme les α-glucosidases (EC 3.2.1.20), par définition, catalysent l'hydrolyse de liaisons inter-glycosidiques α-(1,4) du coté non-réducteur, les premiers tests enzymatiques ont été réalisés avec le substrat considéré comme étant préférentiel de ces enzymes qui est le maltose mais aussi avec différents maltooligosaccharides (Figure 77).

La CCM révèle que le maltose, le maltotriose, le maltotetraose, le maltopentaose, le maltohexaose et le maltoheptaose sont tous hydrolysés en glucose. Ils sont donc bien des substrats naturels de SHWaglu.

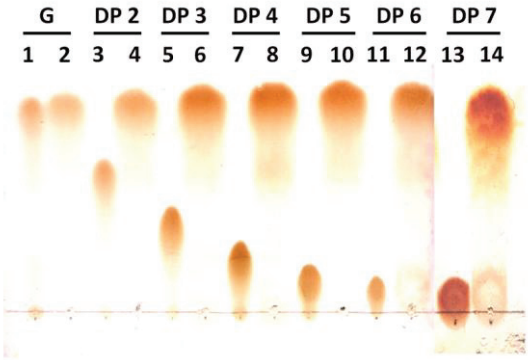


Figure 77 : Analyse par CCM après réaction enzymatique de SHWaglu sur différents maltooligosaccharides. Pistes 1 et 2 : glucose en absence ou présence d'enzyme, respectivement. Piste 3 et 4 : maltose en absence ou présence d'enzyme, respectivement. Pistes 5 et 6 : maltotriose en absence ou présence d'enzyme, respectivement. Pistes 7 et 8 : maltotetraose en absence ou présence d'enzyme, respectivement. Pistes 9 et 10 : maltopentaose en absence ou présence d'enzyme, respectivement. Pistes 11 et 12 : maltohexaose en absence ou présence d'enzyme, respectivement. Pistes 13 et 14 : maltoheptaose en absence ou présence d'enzyme, respectivement.

Dans la littérature, il a plusieurs fois été montré que les α-glucosidases pouvaient être actives sur d'autres types de liaisons interglycosidiques, comme par exemple les α-(1,6) (Alarico *et al.*, 2008; Kitamura *et al.*, 2008; Kang *et al.*, 2009). Ainsi, d'autres tests enzymatiques ont été réalisés sur des substrats variés, comme le sucrose, le tréhalose, l'isomaltose, l'isomaltotriose, le palatinose (ou isomaltulose), le panose, le turanose, le nigérose et le kojibiose, dont les types de liaisons interglycosidiques sont répertoriés dans le Tableau 15.

Sucre	Nombre d'oses	Liaison
Isomaltose	2	Glc α-(1,6) Glc
Isomaltotriose	3	Glc α-(1,6) Glc
Tréhalose	2	Glc α-(1,1) α-Glc
Kojibiose	2	Glc α-(1,2) Glc
Nigérose	2	Glc α-(1,3) Glc
Panose	3	Glc α-(1,6) Glc α-(1,4) Glc
Saccharose	2	Glc α-(1,2) β-Fru
Turanose	2	Glc α-(1,3) Fru
Palatinose	2	Glc α-(1,6) Fru

Tableau 15 : Récapitulatif des oligosaccharides testés avec rappel du nombre d'unités osidiques et du type de liaisons les composant. Glc : glucose, Fru : fructose.

Les résultats montrent que l'enzyme est capable d'hydrolyser l'isomaltose, l'isomaltotriose, le panose, le nigérose et le kojibiose et très faiblement le palatinose (Figure 78). Ainsi SHWaglu possède une grande diversité de substrats puisqu'elle est capable d'hydrolyser les liaisons interglucosidiques de type α-(1,4), α-(1,6), α-(1,2), α-(1,3) et dans une bien moindre mesure la liaison α-(1,6) entre un glucose et un fructose.

En outre, ces résultats indiqueraient que la spécificité de cette enzyme serait plus associée à la conformation de l'unité de sucre présente au niveau du sous-site +1 que le type de liaison interglycosidique impliqué.

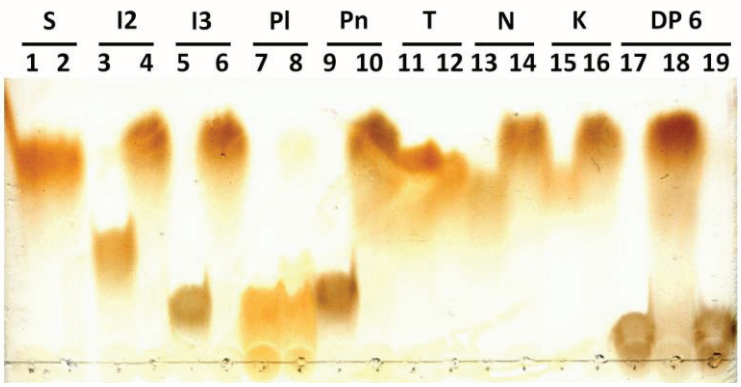


Figure 78 : Analyse par CCM après réaction enzymatique de SHWaglu sur différents oligosaccharides. Pistes 1 et 2 : saccharose en absence ou présence d'enzyme, respectivement. Piste 3 et 4 : isomaltose en absence ou présence d'enzyme, respectivement. Pistes 5 et 6 : isomaltotriose en absence ou présence d'enzyme, respectivement. Pistes 7 et 8 : palatinose en absence ou présence d'enzyme, respectivement. Pistes 9 et 10 : panose en absence ou présence d'enzyme, respectivement. Pistes 11 et 12 : turanose en absence ou présence d'enzyme, respectivement. Pistes 13 et 14 : nigérose en absence ou présence d'enzyme, respectivement. Pistes 15 et 16 : kojibiose en absence ou présence d'enzyme, respectivement. Pistes 17 et 18 : maltohexaose en absence ou présence d'enzyme, respectivement. Piste 19 : maltohexaose en présence d'enzyme sans ajout d'ions Ca^{2+}.

Tous les tests ont été réalisés avec ajout d'ions Ca^{2+} dans le milieu réactionnel (tampon MES 100 mM pH 6,5, $CaCl_2$ 1mM, enzyme 0,65 mg/ml, substrat 4 mM). En effet, des essais effectués sans ajout de $CaCl_2$ ont montré qu'il n'y avait aucune hydrolyse de substrats le cas échéant (Figure 78).

B. Détermination des constantes catalytiques

1. Les substrats

Une fois que les substrats ont été identifiés, nous avons pu déterminer les constantes catalytiques de l'enzyme pour ceux-ci. Après avoir dosé le glucose libéré lors des réactions enzymatiques (protocole détaillé dans le chapitre « Matériel et Méthodes »), nous avons établi les constantes catalytiques (k_{cat}), les constantes de Michaelis (K_m) et, ainsi, les efficacités catalytiques (k_{cat}/K_m) pour chacun des substrats (Tableau 16).

Constantes / Substrats	k_{cat} (s^{-1})	(+/-) k_{cat}	K_m (mM)	(+/-) K_m	k_{cat}/K_m ($s^{-1}.mM^{-1}$)
***p*NPG**	4,12	0,19	0,13	0,02	31,41
***p*NPG (10°C)**	16,1	0,85	0,085	0,02	189 ,4
Maltose	14,89	0,48	0,96	0,12	15,51
Maltotriose	19,32	0,49	0,33	0,04	59,12
Maltotétraose	41,85	2,29	0,86	0,15	48,57
Maltopentaose	44,07	2,16	1,37	0,19	32,22
Maltohexaose	29,89	2,6	2,84	0,6	10,54
Maltoheptaose	ND	-	ND	-	ND
Isomaltose	96,04	3,23	1,08	0,13	89,34
Isomaltotriose	44,06	1,95	0,3	0,06	144,55
Panose	32,89	0,56	0,1	0,02	313,56
Nigérose	365,71	20,61	1,68	0,22	217,69
Kojibiose	268,93	25,2	1,29	0,29	208,47

Tableau 16 : Récapitulatif des constantes catalytiques obtenues pour les différents substrats étudiés. Toutes les mesures ont été réalisées à 40°C, excepté lorsque la température est signalée entre parenthèse. Constantes obtenues par régression non linéaire (logiciel GraphPad Prism).

Les représentations de Michaelis et Menten obtenues pour le *p*NPG et le nigérose sont montrées en exemple (Figure 79).

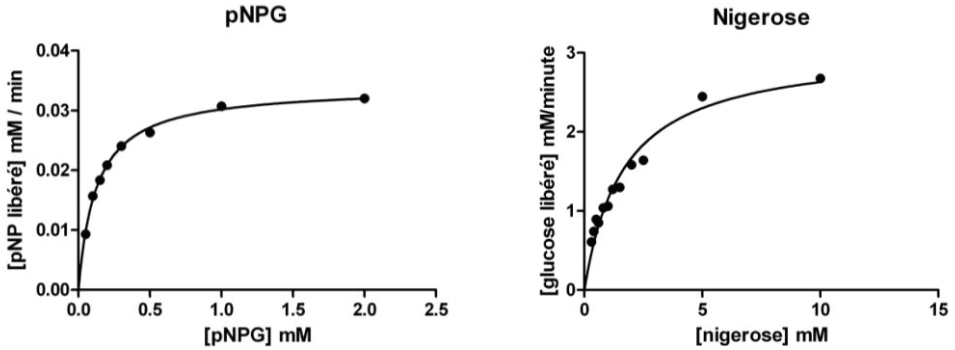


Figure 79 : Représentations de Michaelis et Menten obtenues pour le *p*NPG et le nigérose à 40°C.

Il ressort de cette étude que pour un même type de liaison (α-(1,4) ou α-(1,6)), l'efficacité catalytique de l'enzyme est meilleure sur un trisaccharide que sur les autres DP. Pour les maltooligosaccharides, cette efficacité diminue progressivement lorsque le DP augmente jusqu'à DP7 (maltoheptaose) où il n'a pas été possible de déterminer une quelconque constante. Cette réduction est due à la diminution graduelle d'affinité de l'enzyme pour ces substrats, le k_{cat} n'étant que moyennement altéré.

L'efficacité catalytique se révèle être bien meilleure pour les liaisons glucosidiques autres que α-(1,4) ; le meilleur substrat étant le panose qui est un trisaccharide (avec deux types de liaison interglucosidique différents), confirmant la préférence de l'enzyme pour ce degré de polymérisation. Le nigérose et le kojibiose sont hydrolysés avec une efficacité relativement comparable et avec des constantes catalytiques très élevées par rapport à tous les autres substrats testés. De plus, les constantes catalytiques mesurées pour le *p*NPG à 10°C et 40°C montrent une efficacité catalytique de l'enzyme 6 fois plus importante à basse température qu'à 40°C. Ces résultats sont cohérents avec le caractère psychrotolérant de SHWaglu puisque, très souvent, les enzymes adaptées au froid peuvent « optimiser » le rapport k_{cat}/K_m en augmentant le k_{cat}, diminuant le K_m, voire les deux simultanément (Georlette *et al.*, 2004).

2. Les inhibiteurs et autres ligands

Nous avons ensuite étudié l'affinité de l'enzyme pour différents ligands. Les mesures obtenues, analysées avec le logiciel GraphPad Prism, ont permis de déterminer qu'ils sont tous des inhibiteurs compétitifs. En effet, la V_m n'est pas modifiée confirmant qu'il s'agit bien d'inhibiteurs de type compétitif.

Il apparait que le glucose, qui est le produit d'hydrolyse de l'enzyme, a un effet inhibiteur avec un K_i de 1,2 mM. Il s'agit donc d'un inhibiteur avec une faible affinité par rapport à l'acarbose qui est 1000 fois plus affin. La castanospermine et la déoxynojirimycine ont, pour leur part, une constante d'inhibition intermédiaire (Tableau 17). Les représentations de Michaelis et Menten obtenues pour le *p*NPG en présence de différentes concentrations d'acarbose ou de glucose sont montrées en exemple (Figure 80).

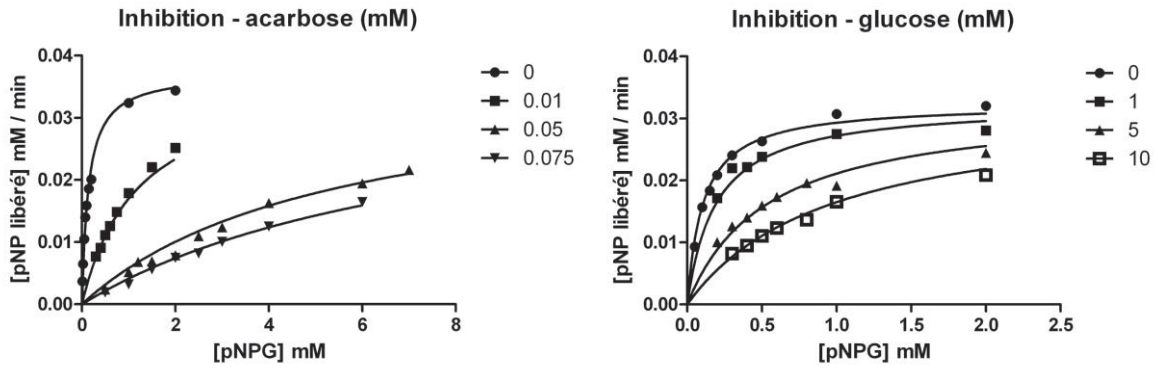


Figure 80 : Représentations de Michaelis et Menten obtenues pour le *p*NPG en présence de différentes concentrations d'inhibiteurs. *A gauche, en présence d'acarbose. Cercles : sans acarbose, carrés : 0,01 mM acarbose, triangles : 0,05 mM acarbose, triangles inversés : 0 ,075 mM acarbose. A droite, en présence de glucose. Cercles : sans glucose, carrés pleins : 1 mM glucose, triangles : 5 mM glucose, carrés vides : 10 mM glucose.*

	K_i (µM)	(+/-) K_i
Acarbose	1.3	0.07
Castanospermine	24.3	4.2
Déoxynojirimycine	7.2	0.6
Glucose	1230	95.8

Tableau 17 : Récapitulatif des constantes d'inhibition obtenues pour les différents ligands étudiés. *Constantes obtenues par régression non linéaire (logiciel GraphPad Prism).*

C. Confirmation du mécanisme catalytique

Etant donné que la famille GH97 a comme particularité de regrouper les deux mécanismes catalytiques, il est important de confirmer celui de SHWaglu. Même si les motifs HHET et K(T/S)GY (normalement retrouvés chez les enzymes de la famille GH97 agissant selon le mécanisme d'hydrolyse par inversion de la configuration du carbone anomérique) est présent dans la séquence protéique de l'enzyme (Figure 81), il est nécessaire de lever le doute. En effet, peu d'enzymes de cette famille ont été caractérisées (Gloster *et al.*, 2008; Kitamura *et al.*, 2008; Okuyama *et al.*, 2009).

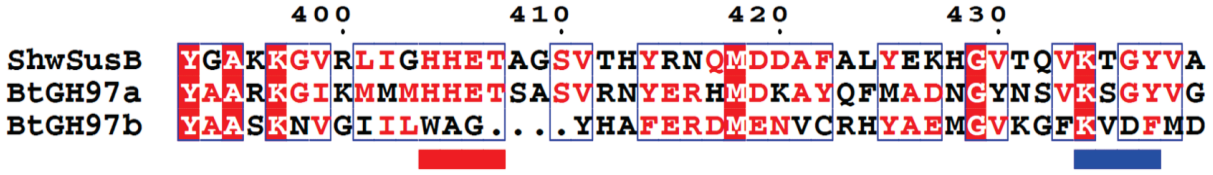


Figure 81 : Alignement de séquence partiel de l'α-glucosidase de *Shewanella* sp. ANA-3 (ShwSusB) avec l'α-glucosidase (BtGH97a) et l'α-galactosidase (BtGH97b) de *B. thetaiotaomicron*. *La numérotation correspond aux résidus de ShwSusB et les motifs HHET et K(T/S)GY sont soulignés en rouge et bleu, respectivement.*

La résonnance magnétique nucléaire (RMN) est la technique que nous avons utilisée pour définir l'anomérie du glucose libéré lors de la réaction enzymatique. Ainsi, au cours de la réaction enzymatique avec le *p*NPG, plusieurs spectres de RMN 1D du proton ont été acquis dans le temps. Le signal du proton spécifique du β-D-glucose (doublet à 4,55 ppm) apparait progressivement après 16 minutes de réaction. Ce signal témoigne de l'accumulation de cet anomère du glucose, démontrant que le β-glucose est le produit initial issu du substrat α-glucosidique. Notons un faible signal, attribué à l'α-D-glucose, qui apparait à partir de 2 h de réaction, et qui découle de la mutarotation naturelle du glucose (Figure 82). Etant donné que le substrat utilisé est le *p*NPG, molécule mimant la liaison glucosidique α-(1,4), nous pouvons conclure que SHWaglu agit bien selon le mécanisme d'inversion de la configuration du carbone anomérique. Toutefois, ceci va à l'encontre de la classification IUBMB de cette α-glucosidase (EC 3.2.1.20) qui, par définition, catalyse certes l'hydrolyse des liaisons glucosidiques α-(1,4) de l'extrémité non-réductrice, mais avec libération d'α-D-glucose. Ainsi, SHWaglu semble plutôt être une glucoamylase (EC 3.2.1.3) comme le suggère Kitamura *et coll.* à propos de SusB (Kitamura *et al.*, 2008).

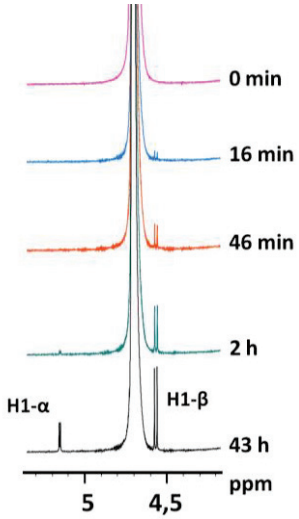


Figure 82 : Spectres RMN 1D centrés sur la zone du signal de l'eau acquis au cours du temps de la réaction enzymatique. H1-α : signal du proton du carbone anomérique de l'α-D-glucose, H1-β : signal du proton du carbone anomérique du β-D-glucose.

V. Étude structurale

A. Cristallogenèse

La stratégie adoptée pour les expériences de cristallogenèse a été de faire un large criblage de conditions en utilisant plusieurs kits commerciaux proposés par MDL, Qiagen ou Hampton Research.

Les premiers essais ont été réalisés en goutte assise selon la méthode de diffusion de vapeur grâce au robot Mosquito® (TTP Labtech) dans des boites Greiner Low profile 96 puits (Greiner Bio One). Cet appareil permet de prélever de très petits volumes (jusqu'à 100 nl) de façon précise et reproductible. Ainsi, près de 200 conditions ont été testées en utilisant la protéine préalablement concentrée à 2,5 mg/ml. Des gouttes de 300 nl (1:1, $v_{protéine}$:$v_{solution}$) ont été réalisées avec 100 µl de solution de cristallisation dans le puits. Toutes les gouttes sont restées claires après deux semaines, démontrant que la protéine n'était potentiellement pas suffisamment concentrée pour atteindre la phase de sursaturation obligatoire pour initier la formation de points de nucléation.

Une deuxième série d'essais de cristallogenèse a été réalisée en concentrant la protéine à 5,5 mg/ml et en répétant les mêmes conditions. Cette fois, des cristaux en forme de baguette sont apparus en 3-4 jours à 16°C dans des conditions contenant du chlorure de lithium (0,2 M) et 20% de PEG 3350. D'autres cristaux en forme de plaquette sont apparus au bout de deux semaines dans une condition totalement différente (0,2 M ammonium phosphate monobasique, 0,1 M Tris pH 8,5 et 50% (v/v) MPD) (Figure 83).

Ces cristaux ont été testés avec la source de rayons X du laboratoire (anode tournante couplée à un détecteur de type « Image Plate »). Il s'est avéré que le cristal en forme de baguette est un cristal de protéine et que celui en forme de plaquette est un cristal de sel.

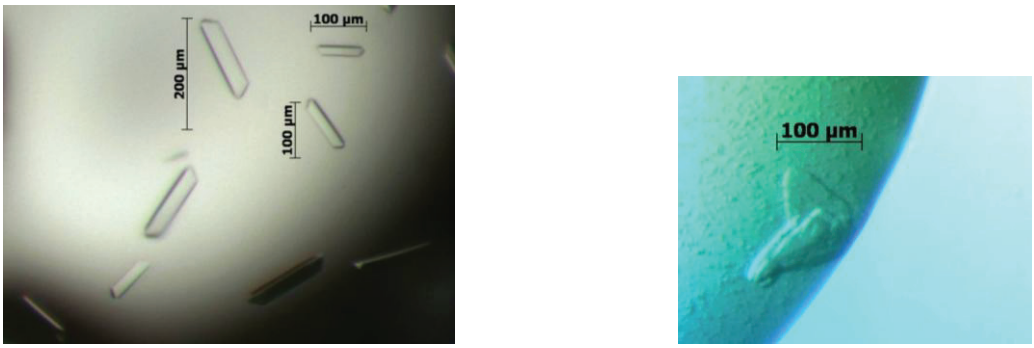


Figure 83 : Cristaux obtenus lors du criblage de conditions de cristallisation de SHWaglu à 5,5 mg/ml. *A gauche : cristaux en forme de baguette, solution de cristallogenèse : LiCl (0,2 M) et 20% PEG 3350. A droite : cristaux en forme de plaquette, solution de cristallogenèse : 0,2 M ammonium phosphate monobasique, 0,1 M Tris pH 8,5 et 50% (v/v) MPD.*

Une optimisation des conditions de cristallisation a été effectuée autour de la condition initiale (0,2 M LiCl, 20% PEG 3350) en faisant varier la concentration des ions et le pourcentage de PEG dans la solution. Des gouttes suspendues de 1 µl (1:1) ont ainsi été réalisées dans 0,18 M à 0,22 M LiCl, 18% à 22% PEG 3350. Des cristaux de même morphologie sont apparus au bout d'une semaine à 16°C dans toute la matrice d'optimisation (Figure 84).

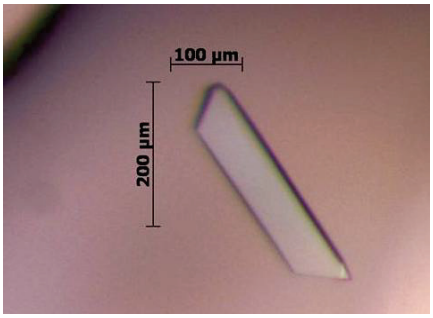


Figure 84 : Cristal de SHWaglu obtenu dans une goutte suspendue lors de l'optimisation des conditions de cristallisation.

B. Cristallographie aux rayons X

1. Résolution de la structure

Un premier jeu de données synchrotron a été enregistré sur la ligne ID14-1 à l'European Synchrotron Radiation Facility (ESRF) de Grenoble (Figure 85) à une résolution de 1,77 Å, à la longueur d'onde 0,9334 Å. Au préalable, le cristal a été cryoprotégé par trempage dans sa solution de cristallisation supplémentée de 20% de glycérol pendant 2h40.

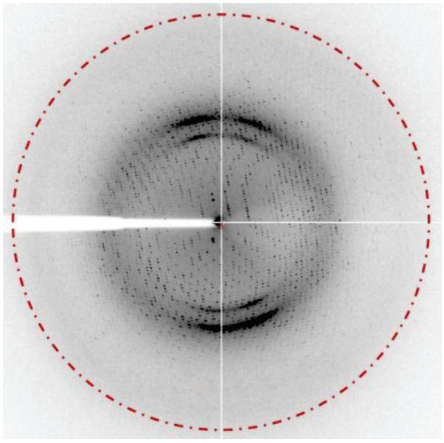


Figure 85 : Cliché de diffraction d'un cristal de SHWaglu native obtenu sur ID14-1. Les pointillés rouges délimitent l'anneau de résolution de 2 Å.

Les données ont été intégrées et moyennées avec les programmes XDS et XSCALE (Kabsch, 2010). Le problème des phases a été résolu par remplacement moléculaire en employant le programme BALBES (Long *et al.*, 2008) et en utilisant, comme modèle guide, la structure de l'α-glucosidase de *Bacteroides thetaiotaomicron* (code PDB 2D73) (Kitamura *et al.*, 2008) avec laquelle elle partage 36% d'identité de séquence.

Cette étape a été suivie d'une reconstruction manuelle dans COOT (Emsley *et al.*, 2010) puis d'un affinement par le logiciel REFMAC (Vagin *et al.*, 2004) jusqu'à des facteurs de confiance finaux R_{factor} et R_{free} de 14,9% et 18,2% respectivement.

La structure cristalline finale de SHWaglu contient 1314 résidus d'acides aminés répartis en deux monomères appelés A et B, 1505 molécules d'eau, trois molécules de glycérol, deux ions calcium et neuf ions chlore. Tous les résidus constituant la protéine ont pu être construits dans la densité électronique à l'exception des 47 premiers des deux monomères, à cause de l'absence totale de densité électronique. La Figure 86 montre un aperçu de la qualité de la carte de densité électronique $2F_o$-F_c du modèle après le dernier cycle d'affinement.

La qualité de la structure 3D, en termes de géométrie et d'erreurs de coordonnées, a été examinée avec les programmes PROCHECK (Laskowski *et al.*, 1993) et WHATCHECK (Hooft *et al.*, 1996). L'ensemble des statistiques d'enregistrement et de traitement des données est résumé dans le Tableau 18 suivant.

	SHWaglu native
Enregistrement de données	
Ligne	ID14-1 (ESRF)
Longueur d'onde (Å)	0,93340
Groupe d'espace	$P2_12_12_1$
Paramètres de maille	
a, b, c (Å)	100,56, 111,77, 128,93
α, β, γ (°)	90 90 90
Vm ($Å^3.Da^{-1}$)	2,22
Pourcentage de solvant (%)	44,6
Résolution (Å)	48,85-1,77 (1,82-1,77)
Nombre total de réflexions	1056421 (51208)
Nombres de réflexions uniques	141567 (6856)
R_{sym} (%)	10,3 (66)
I/σ (I)	19,6 (3,5)
Complétude (%)	100 (100)
Redondance	7,5 (7,5)
Affinement	
R_{factor}/R_{free} (%)	14,8 / 18,2
Nombre d'atomes	
Protéine (non-hydrogène)	10750
Ligands/ions	29
Molécules d'eau	1505
Facteur B moyen ($Å^2$)	14
Déviation r.m.s.	
Longueur de liaison (Å)	0,015
Angle de liaison (°)	1,55

Tableau 18 : Statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de SHWaglu. Les valeurs pour la coquille de plus haute résolution sont indiquées entre parenthèses.

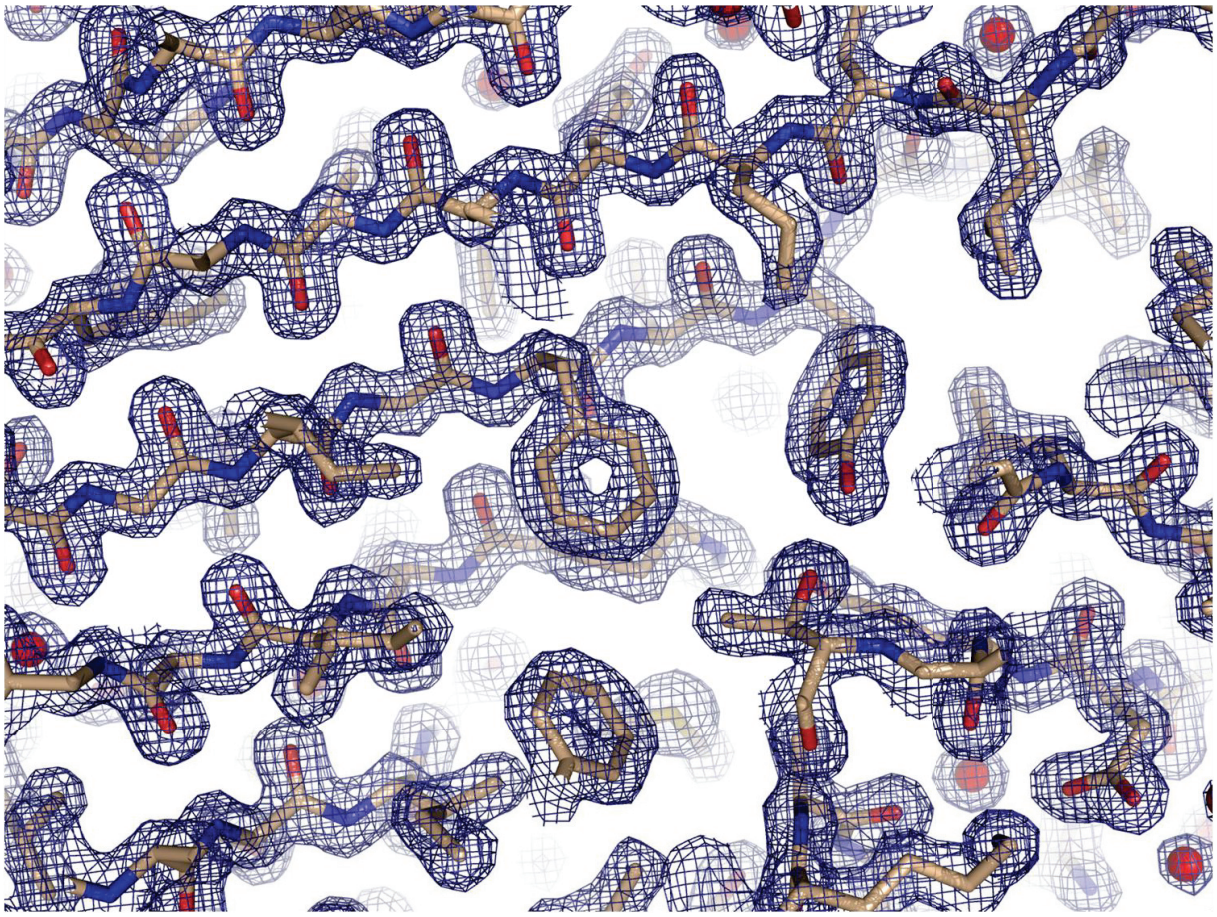


Figure 86 : Apercu de la carte de densité électronique $2F_o$-F_c à 1,77 Å de résolution (contourée à 1,4 σ) et centrée sur la phénylalanine 155. Les atomes de carbone, d'oxygène et d'azote sont respectivement représentés en beige, rouge et bleu. Les molécules d'eau sont représentées par des sphères rouges.

2. Description de la structure de SHWaglu

La structure cristalline de SHWaglu est constituée de deux chaines polypeptidiques similaires, nommées A et B et formant l'unité asymétrique. Comme chez son homologue structural (l'α-glucosidase de *Bacteroides thetaiotaomicron*, SusB), chacun de ces monomères est constitué de trois domaines : le domaine N-terminal (domaine N), le domaine catalytique (domaine A) et le domaine C-terminal (domaine C) (Figure 87).

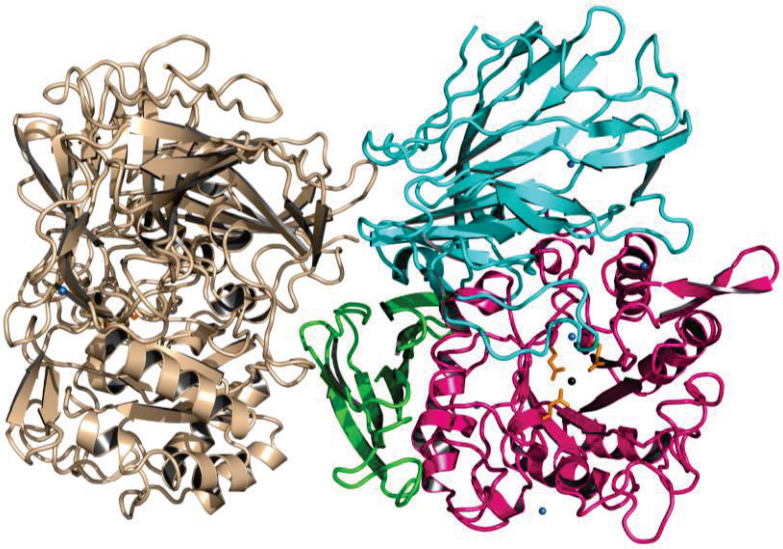


Figure 87 : Représentation en ruban des deux monomères présents dans l'unité asymétrique. Cyan : domaine N ; magenta : domaine A, vert : domaine C, or : deuxième monomère. Les résidus catalytiques sont représentés en bâtons orange avec, à proximité, un ion Ca^{2+} coloré en gris. Les ions Cl^- sont colorés en bleu.

Ces trois domaines organisés sur un même plan sont très étroitement liés et forment ainsi une entité compacte. Il est donc possible de distinguer deux faces chez cette enzyme : la face 1 expose la poche catalytique et la face 2 correspondant à la face diamétralement opposée (Figure 88).

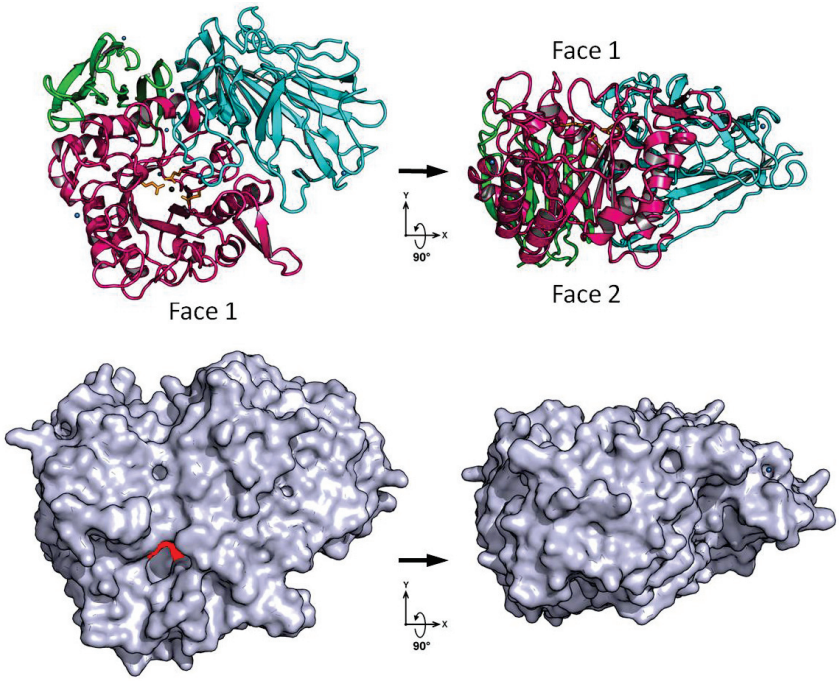


Figure 88 : Représentation en ruban et en surface de la structure 3D du monomère dont les trois domaines sont organisés sur un même plan. Les codes couleurs sont identiques à la Figure 87. Le site actif est coloré en rouge dans la représentation en surface. La face 1 correspond à la face exposant la poche catalytique, la face 2 est la face opposée.

Les monomères A et B peuvent être superposés l'un sur l'autre avec un R.M.S.D de 0,33 Å calculé pour 628 carbones α. Les principales différences sont observables au niveau des résidus 323 à 335 situés dans une boucle présente entre deux hélices α au début du domaine catalytique (RMSD de 0,3 à 0,53 Å entre les paires de carbones α), mais surtout entre les résidus 647 à 653 (RMSD de 0,35 à 3,2 Å entre les paires de carbones α). Cette dernière série de résidus se situe dans une boucle appartenant au domaine C-terminal qui, dans le cas du monomère B, participe à l'interface entre les deux monomères de l'unité asymétrique ou, dans le cas du monomère A, se trouve en surface.

a. Le domaine N

Le domaine N (jusqu'au résidu 298) est un domaine en super sandwich β. Il comporte 19 brins β antiparallèles et 3 hélices α très courtes. Quatre boucles de ce domaine (résidus 117-122, 194-205, 237-242 et 259-264) interagissent avec le domaine A catalytique. De façon très intéressante, trois de ces boucles débutent par une proline (Pro117, Pro194 et Pro259). Ces positions sont d'autant plus remarquables qu'elles sont toutes les trois conservées chez les deux homologues structuraux SusB et *Bt*GH97b. L'une d'entre elles (résidus 117-122) interagit non seulement avec une boucle du domaine A (résidus 570-573) mais aussi avec un brin β du domaine C (résidus 600-603). Cette dernière (résidus 117-122) est également particulière car elle est isolée par rapport aux autres : en effet, il s'agit de la seule sur les quatre qui n'est pas retrouvée sur la face 1 de l'enzyme et donc n'interagit pas du coté du site actif du domaine catalytique (Figure 89). Chez SusB (*B. thetaiotaomicron*), deux boucles forment une partie de la poche du site catalytique (Kitamura *et al.*, 2008). En comparaison, chez l'α-glucosidase de *Shewanella* sp. ANA-3, une seule boucle (résidus 194-205) vient participer au domaine catalytique (Figure 89); certains de ces résidus participant même à la poche du site actif (résidu 202). Autre différence, une grande boucle de 20 résidus chez l'α-glucosidase *B. thetaiotaomicron* (résidus 67-87) n'est pas du tout présente chez SHWaglu (au niveau de la sérine 92). Ce domaine N est relié au domaine catalytique *via* une boucle (résidus 299-311).

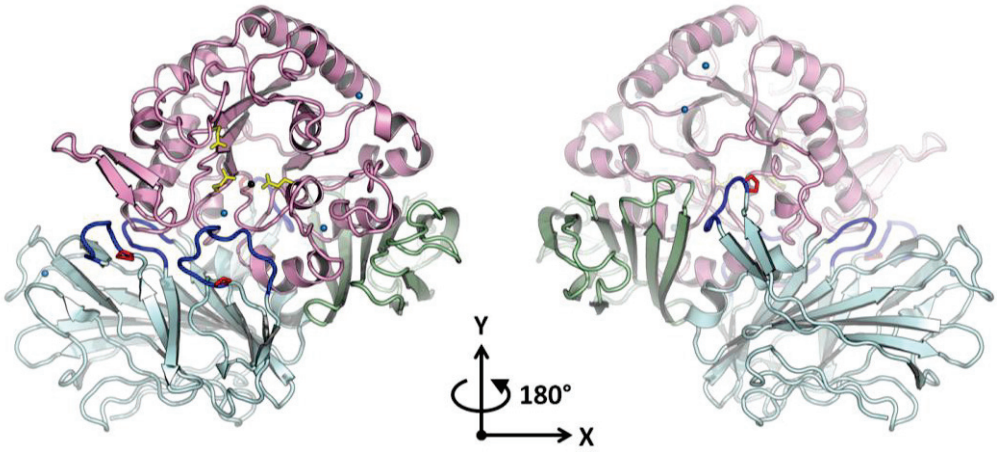


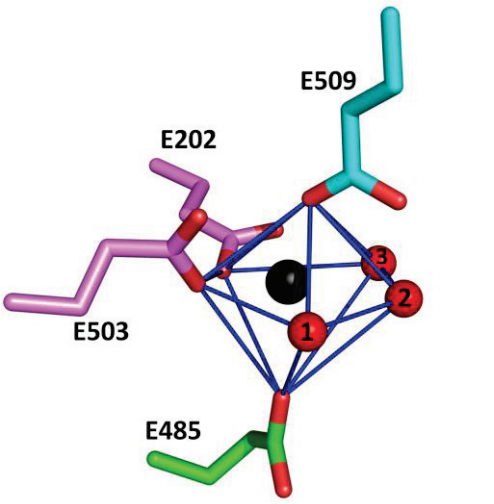
Figure 89 : Représentation en ruban de la structure 3D de SHWaglu et mise en évidence des boucles du domaine N interagissant avec le domaine A. Les domaines N, A et C sont colorés respectivement en cyan, rose et vert. Les boucles impliquées sont colorées en bleu. Les prolines impliquées dans ces boucles et les résidus catalytiques sont représentés en bâtons et colorés en rouge et jaune respectivement. Les ions Ca^{2+} et Cl^{-} sont representés en sphère noire et bleue, respectivement.

b. Le domaine A catalytique

Le domaine catalytique A (résidus 311-599) forme un tonneau $(\beta/\alpha)_8$ avec (comme chez son homologue structural) une longue boucle (résidus 534-547) et 7 brins β parallèles centrés autour d'un axe et encerclés par 8 hélices α dont une très restreinte (α7). Ceci peut s'expliquer par la présence de deux prolines (Pro486 et Pro497) de part et d'autre des résidus normalement impliqués dans cette hélice d'un tonneau $(\beta/\alpha)_8$ canonique. En effet, les prolines sont reconnues pour rigidifier les structures mais aussi pour être généralement des « briseuses » de structures secondaires. Ceci s'explique par leur particularité structurale empêchant la formation d'une liaison hydrogène mais aussi à cause de leur rigidité conformationelle et est renforcé par leur faible propension à former des hélices α (Pace & Scholtz, 1998). Alors que la structure du tonneau $(\beta/\alpha)_8$ typique est composée de 8 brins β, la région correspondant au brin β7 d'un $(\beta/\alpha)_8$ typique forme une boucle chez l'α-glucosidase de *Shewanella* sp. ANA-3, comme chez son homologue. Les boucles du domaine N interagissant avec ce domaine catalytique sont localisées justement au niveau de cette région, stabilisant très probablement l'ensemble en compensant les structures secondaires manquantes.

Le site actif, comme chez l'homologue structural (Kitamura *et al.*, 2008) et d'autres protéines en tonneau $(\beta/\alpha)_8$ (Nagano *et al.*, 2002; Ochoa-Leyva *et al.*, 2009), est situé du coté C-terminal de ce tonneau. Un ion calcium est localisé dans cette poche et détermine des liaisons de coordination avec les résidus catalytiques E406, E485 et E509 mais aussi avec les résidus E202 et E503. Dans la structure de la protéine native, cet ion détermine

également des liaisons avec des molécules d'eau pour, au final, acquérir une géométrie de type bipyramide pentagonale. Le O-E2 du E485 et le O-E2 de E509 sont positionnés au sommet de cette bipyramide alors que le O-E2 de E202, le O-E1 de E503 et trois molécules d'eau forment la base plan de celle-ci. Les distances de coordination entre l'ion calcium et les atomes mentionnés ci-dessus (monomère A) sont comprises entre 2,1 et 3,5 Å (Figure 90). Vu qu'aucun réactif contenant des ions calcium n'a été utilisé dans les étapes de purification et de cristallisation, ces observations indiquent que l'ion provient de la cellule hôte *E. coli* et qu'il étroitement est lié au centre actif de l'enzyme, ce qui est également le cas chez l'homologue structural (Kitamura *et al.*, 2008). Les essais de caractérisation biochimique montrent d'ailleurs que cet ion joue un rôle clé dans l'activité catalytique. L'ion « recrute » les groupements fonctionnels et le substrat et sert d'acide de Lewis en polarisant le carbone anomérique et le rendant meilleur électrophile.



Atome coordonné avec Ca^{2+}	Distance (Å)
O-E2 de E202	2,9
O-E2 du E485	2,9
O-E1 de E503	2,2
O-E2 de E509	2,1
Molécule d'eau 1	2,3
Molécule d'eau 2	2,8
Molécule d'eau 3	3,5

Figure 90 : *A gauche*, vue rapprochée du site actif centrée sur l'ion calcium. Les résidus et molécules d'eau impliqués dans la coordination sont représentés en bâton et en sphères, respectivement. La représentation de la bipyramide pentagonale est colorée en bleu. ***A droite*, tableau récapitulant les atomes de l'enzyme impliqués dans la coordination de l'ion calcium et leurs distances de coordination (en Å).**

A l'entrée de la poche catalytique, on retrouve une « couronne » de résidus aromatiques qui sont tous conservés ou similaires chez SusB à l'exception de l'histidine 322 (Ile chez SusB) (Figure 91).

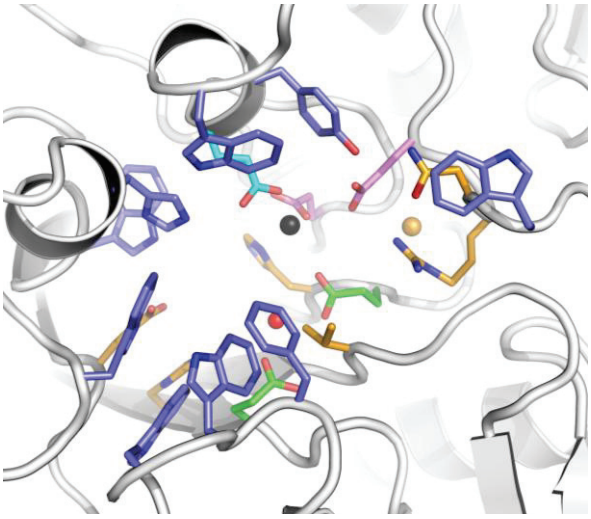


Figure 91 : Vue rapprochée de l'entrée du site actif de SHWaglu et visualisation de la « couronne » de résidus aromatiques. Les ions Ca^{2+} et Cl^- sont représentés en sphères grise et jaune, respectivement.

c. Le domaine C

Le domaine C (résidus 600-703) est un sandwich β formé de 8 brins β antiparallèles. Un des feuillets est composé de 6 brins (Cβ1, Cβ2, Cβ3, Cβ5, Cβ6, Cβ8) alors que le second, beaucoup plus petit, n'est composé que de 2 brins (Cβ4, de 636 à 644 et Cβ7, de 684 à 691). Chez SHWaglu, l'extrémité C-terminale est plus courte que chez SusB et se termine juste à la fin du dernier brin β. Il est à noter que le facteur d'agitation thermique B moyen diffère entre les deux monomères seulement pour ce domaine-ci (Figure 92). Ceci peut être expliqué par l'organisation des deux molécules au sein de l'unité asymétrique. En effet, de par leur orientation respective, le domaine C du monomère A crée plus de contacts avec le monomère B que l'inverse.

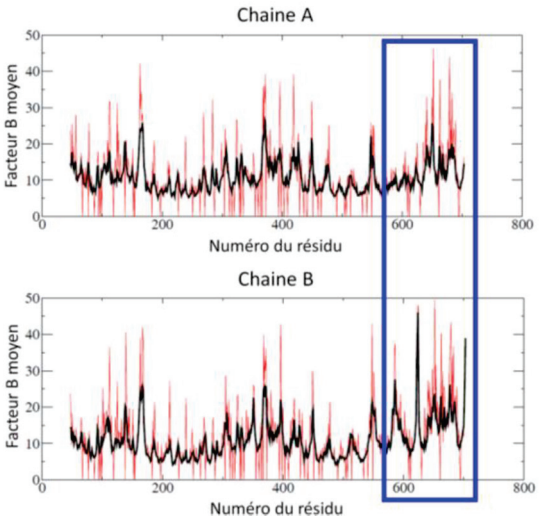


Figure 92: Représentation de la valeur du facteur d'agitation thermique B par résidu. Le cadre bleu montre la différence majeure en facteur d'agitation thermique B moyen entre les deux monomères de l'unité asymétrique au niveau du domaine C.

d. Un « motif » tridimensionnel conservé dans la famille GH97

Face à la proline 497, on retrouve la proline 117 ainsi que les prolines 311, 535 et 573. Le tout est centré autour d'un tryptophane (Trp118) rendant l'environnement local hydrophobe, ceci étant conforté par la présence de valine et leucine (Val306, Val309 et Leu500) tendant à stabiliser l'ensemble (Figure 93). D'après l'alignement multiple de séquences et la superposition structurale réalisés entre SHWaglu, SusB et *Bt*GH97b, ces prolines sont conservées et se structurent en formant un patch autour d'un résidu hydrophobe.

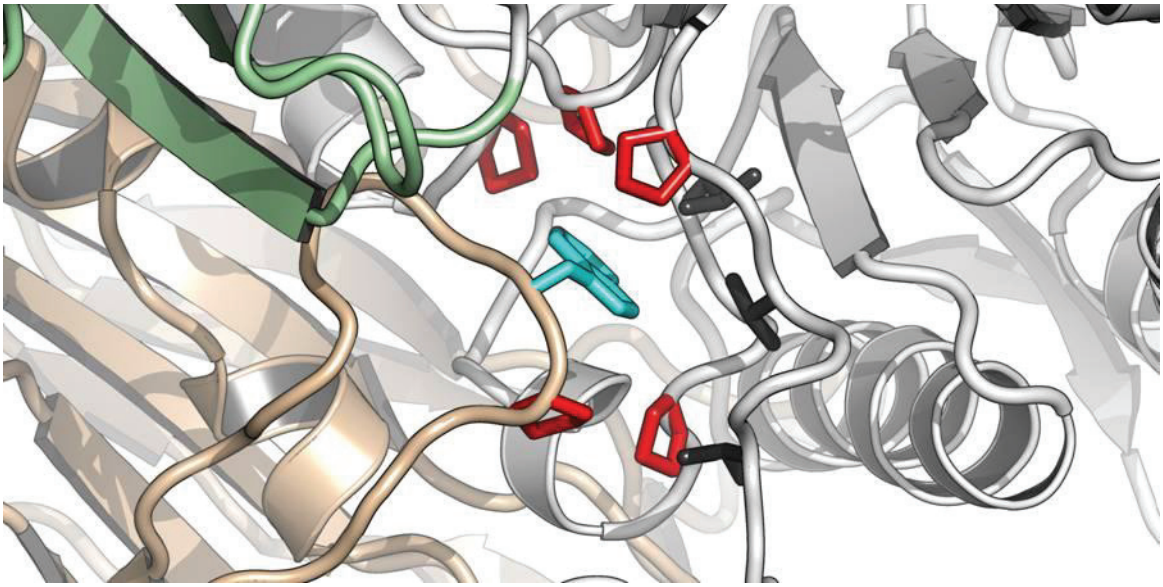


Figure 93 : Vue rapprochée de SHWaglu focalisée sur le patch de prolines. En blanc : domaine A, en or : domaine N, en vert : domaine C. En rouge : les prolines 117, 311, 497, 535 et 573 ; en cyan : le tryptophane 118, en noir : les valines 306, 309 et leucine 500.

De manière plus générale, un alignement multiple a été réalisé sur un plus grand nombre de séquences d'homologues « inverting » (43 séquences) ou « retaining » (52 séquences) de la famille GH97 référencées dans la base de données CAZy. Il s'avère qu'un motif de séquence est retrouvé pour chacune des prolines. Ainsi, nous avons pu établir pour la sous-famille des « inverting » les motifs suivants : WXX(P/V)WGE, (W/Y)(V/I/F)XPXK(Y/F)(V/I/M)G(V/I)WW, RT(Y/W)PN, GPMD(F/Y)TPG et YSPXQMA(A/S)D pour les prolines 117, 311, 497, 535 et 573 respectivement, avec des variations possibles sur certaines positions pour quelques séquences protéiques (Figure 94). La valine 306 est très peu conservée contrairement à la valine 309 et la leucine 500 qui correspondent toujours à des résidus hydrophobes. Dans la sous-famille des « retaining », les motifs sont plus dégénérés. Toutefois, la proline 497 est entièrement conservée et le motif dans lequel est retrouvée la proline 311 est W(V/I)XPG.

Ces constatations tendent à souligner l'importance, chez les GH97, de la présence de ce patch hydrophobe pour la stabilisation de cette partie du domaine catalytique grâce au domaine N.

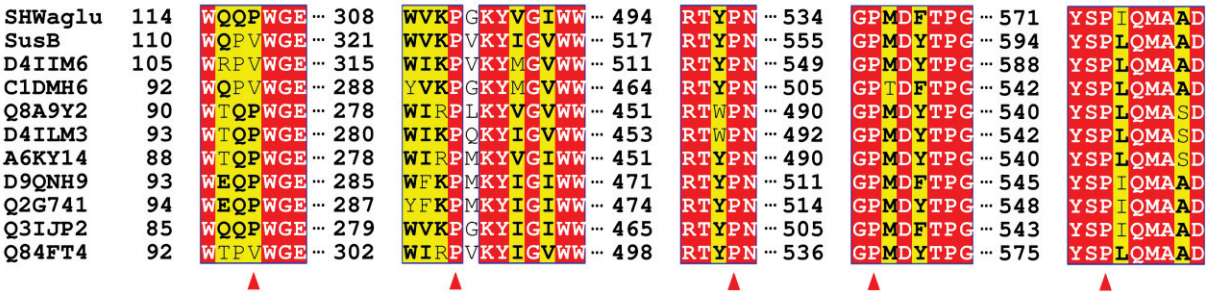


Figure 94 : Alignement multiple de séquences focalisé sur les motifs conservés dans lesquels sont retrouvés les différentes prolines impliquées dans le « patch » hydrophobe. Les prolines impliquées sont repérées par les triangles rouges. Les codes Uniprot des enzymes de la sous-famille « inverting » de GH97 sont montrés à gauche de l'alignement. D4IIM6 et D4ILM3 proviennent de *Alistipes shahii*, C1DMH6 de *Azotobacter vinelandii*, Q8A9Y2 de *Bacteroides tethaiotaomicron*, A6KY14 de *Bacteroides vulgatus*, D9QNH9 de *Brevundimonas subvibrioides*, Q2G741 de *Novosphingobium aromaticivorans*, Q3IJP2 de *Pseudoalteromonas haloplanktis* et Q84FT4 de *Bacteroides forsythus*.

Dans cet alignement multiple, il est à noter que l'enzyme « Q3IJP2 » provient de l'organisme psychrophile *Pseudomonas haloplanktis*. Un alignement de séquences simple entre SHWaglu et « Q3IJP2 » met en avant une très forte homologie avec 61,4% d'identité de séquence.

e. Les caractères psychrophiles de l'enzyme

Il est généralement accepté que la thermophilie d'une enzyme est corrélée à la rigidité de la protéine et, qu'à l'inverse, les enzymes psychrophiles ont une structure relativement plus flexible (Aghajari *et al.*, 1998; Siddiqui & Cavicchioli, 2006). Dans ce cadre, il a été rapporté que comparativement à leur homologues mésophiles, les enzymes psychrophiles possédaient souvent en proportion un nombre supérieur de glycines (Aghajari *et al.*, 2003; Georlette *et al.*, 2004; Goldstein, 2007). Il apparait, effectivement, que les domaines N et A de SHWaglu possèdent proportionnellement plus de glycines (6,1% et 10%, respectivement) que ces mêmes domaines chez SusB (4,3% et 7,7%, respectivement). Il a également été noté que le nombre d'interactions (liaisons hydrogène et ponts salins) entre les domaines étaient souvent réduites chez les enzymes psychrophiles comparativement à leurs homologues mésophiles (Smalås *et al.*, 2000; Cavicchioli *et al.*, 2002; Aghajari *et al.*, 2003). Dans notre cas, il apparait que les liaisons hydrogène et les ponts salins sont en nombre moins importants chez SHWaglu que chez SusB. Ceci est d'autant plus flagrant à l'examen des interactions intermoléculaires du dimère chez SusB (Tableau 19). En

considérant l'état oligomérique de ces deux protéines en solution, il s'avère qu'il y a un écart d'énergie de stabilisation de 10,5 kcal/mol entre SHWaglu et SusB démontrant une stabilité plus importante pour cette dernière (Tableau 19). Cette différence d'énergie est similaire à ce qui a parfois été constaté entre les enzymes psychrophiles et mésophiles (Smalås *et al.*, 2000).

	SHWaglu			SusB		
Interdomaine	Liaisons hydrogène	Ponts salins	Energie de stabilisation (ΔG kcal/mol)	Liaisons hydrogène	Ponts salins	Energie de stabilisation (ΔG kcal/mol)
N/A	32	6	-29,8	40	10	-25,6
A/C	19	2	-24,6	16	3	-22,4
N/C	10	1	-3	13	3	-3,8
Intermoléculaire	-			32	8	-12.5

Tableau 19 : Récapitulatif du nombre de liaisons hydrogène, de ponts salins et de l'énergie de stabilisation (en kcal/mol) chez SHWaglu et SusB entre 1) les domaines N, A et C au sein d'un monomère et 2) les monomères. Calculés avec le logiciel en ligne PISA (Krissinel & Henrick, 2007).

Fréquemment, les calculs des potentiels électrostatiques des enzymes psychrophiles révèlent un excès de charges négatives à la surface de ces protéines comparativement à leurs homologues mésophiles ou thermophiles (Siddiqui & Cavicchioli, 2006; Struvay & Feller, 2012). Ainsi, même si la face exposant le site actif (face 1) de SHWaglu est chargée négativement de façon similaire à celle de SusB avec un maximum au niveau du site actif (Figure 95 A et D), la face opposée (face 2) de SHWaglu montre une surface beaucoup plus négative que chez son homologue (Figure 95 B et E). La différence majeure est retrouvée au niveau du domaine C, légèrement chargé négativement chez SHWaglu alors qu'il est chargé positivement chez SusB (Figure 95 C et F).

Dans plusieurs études, il apparait que le site actif chez les enzymes psychrophiles est plus « ouvert » que chez leurs homologues mésophiles (Smalås *et al.*, 2000; Georlette *et al.*, 2004; Struvay & Feller, 2012) avec, par exemple, des délétions au niveau de boucles bordant ledit site (Smalås *et al.*, 2000; Georlette *et al.*, 2004). L'examen de l'entrée du site actif de SHWaglu montre une poche plus ouverte que celle de SusB (Figure 95 A et D) du fait de l'absence de la boucle coiffant le site chez SusB (Figure 95 G et H). Cette caractéristique permettrait à ces enzymes de réduire l'énergie d'activation requise pour la formation du complexe enzyme-substrat mais aussi de faciliter la libération des produits (Struvay & Feller, 2012).

D'autres caractéristiques souvent retrouvées chez les psychrophiles, comme un nombre plus important de prolines ou de résidus hydrophobes à la surface de l'enzyme, ont

également été recherchées chez SHWaglu. Or, d'autres différences majeures que celles déjà pré-citées ne sont pas ressorties comparativement à SusB.

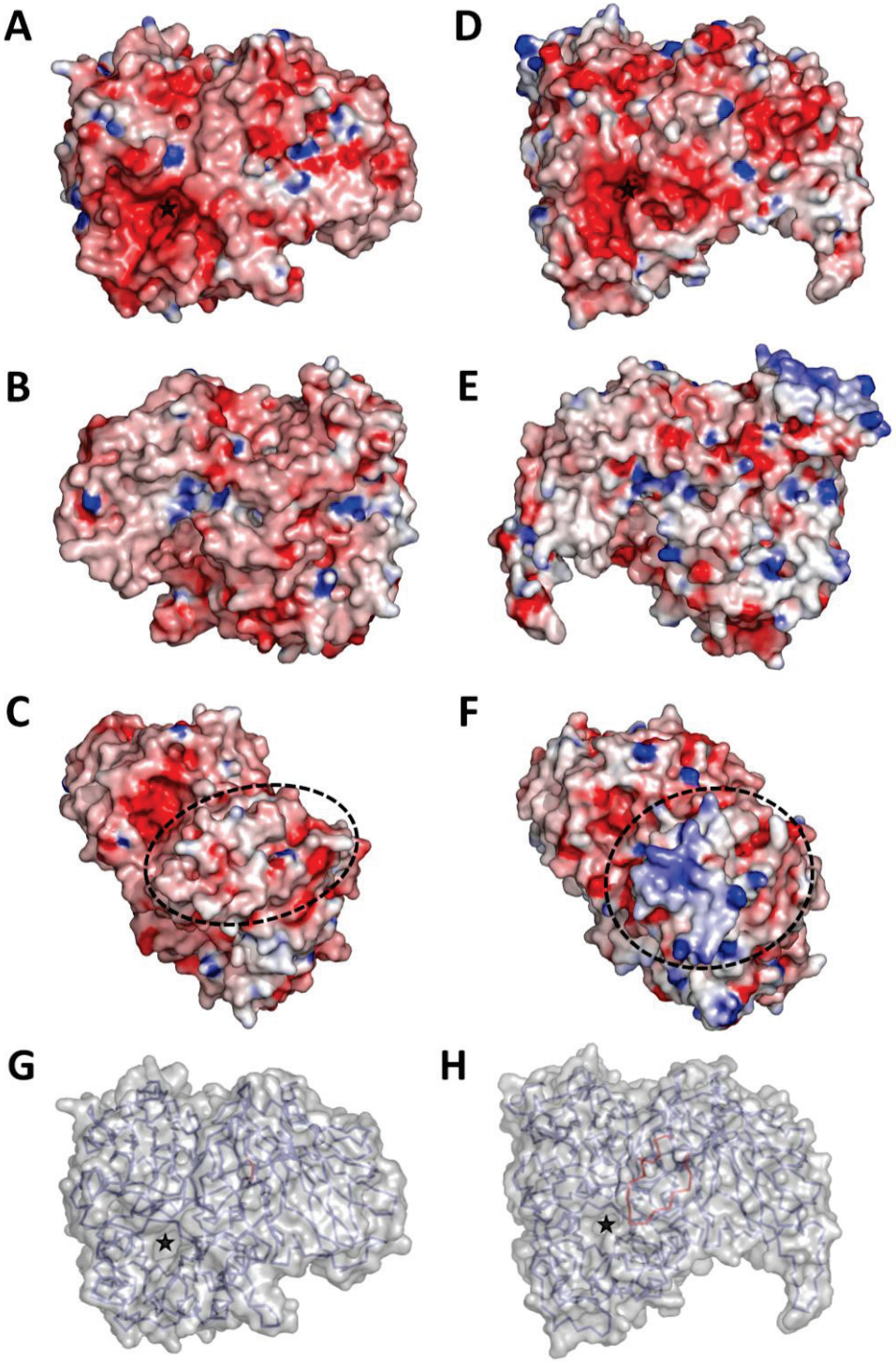


Figure 95 : Projections des potentiels électrostatiques sur les surfaces moléculaires de SHWaglu (A, B, C) et SusB (D, E, F). Les sites actifs sont localisés par une étoile dans les figures A et D. Les figures B et E montrent la face opposée au site actif (face 2) et les figures C et F montrent en premier plan le domaine C (encerclé en pointillés). Les charges positives (bleues) et négatives (rouges) sont distribuées sur la même échelle de valeurs dans toutes les représentations. **Visualisation de la boucle coiffant le site actif (en rouge) chez SusB (H) et son absence chez SHWaglu (G).**

3. Études structurales de SHWaglu native en complexe avec des inhibiteurs

a. Acquisition des données

Des cristaux de la protéine native ont été mis en contact avec les différents ligands d'intérêt par trempage pendant au minimum 2h40 : de l'acarbose (ACA), de la castanospermine (CST) ou de la déoxynojirimycine (DJN) (10 mM, 10 mM, 5 mM respectivement) ont ainsi étés ajoutés dans la solution de cristallisation supplémentée de 20% de glycérol.

Les données synchrotron ont été enregistrées sur la ligne FIP-BM30A aux résolutions de 1,85 Å, 2 Å et 1,9 Å respectivement. Les données ont été intégrées et moyennées avec les programmes XDS et XSCALE (Kabsch, 2010). Le problème des phases a été résolu par remplacement moléculaire en employant le programme MolRep (Vagin & Teplyakov, 2010) et en utilisant, comme modèle guide, la structure de la protéine native affinée. Les modèles de ligands trouvés dans la base de données HIC-Up ont été insérés manuellement dans les cartes de densité électroniques correspondantes. Comme pour la structure native, les 47 premiers résidus ne sont pas visibles dans la densité électronique.

La qualité des structures 3D, en termes de géométrie et d'erreurs de coordonnées, a été examinée avec les programmes PROCHECK (Laskowski *et al.*, 1993) et WHATCHECK (Hooft *et al.*, 1996). L'ensemble des statistiques d'enregistrement et de traitement des données est présenté dans le Tableau 20 suivant.

	SHWaglu-ACA	SHWaglu-CST	SHWaglu-DJN
Enregistrement de données			
Ligne	BM30 (ESRF)	BM30 (ESRF)	BM30 (ESRF)
Longueur d'onde (Å)	0,979631	0,979772	0,979777
Groupe d'espace	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Paramètres de maille			
a, b, c (Å)	100,78, 111,86, 128,82	101,69 110,74, 132,01	101,06 111,91 129,96
α, β, γ (°)	90 90 90	90 90 90	90 90 90
Vm ($Å^3.Da^{-1}$)	2,22	2,28	2,25
Pourcentage de solvant (%)	44,7	46	45,4
Résolution (Å)	48,83-1,85 (1,9-1,85)	49,52-2 (2,1-2)	50-1,9 (2-1,9)
Nombre de réflexions totales	807897 (54584)	692202 (66825)	807117 (80055)
Nombres de réflexions uniques	125160 (8959)	98256 (11214)	116345 (16390)
R_{sym} (%)	13,4 (65,2)	14,2 (54,3)	11 (55,8)
I/σ (I)	14,5 (3,2)	17,4 (5,3)	15,9 (3,6)
Complétude (%)	99,8 (97,6)	97,2 (82,5)	99,9 (99,8)
Redondance	6,5 (6,1)	7 (5,6)	6,9 (4,9)
Affinement			
R_{factor}/R_{free} (%)	13,9 / 17,8	16 / 20,2	14,1 / 17,8
Nombre d'atomes			
Protéine (non-hydrogène)	10771	10679	10595
Ligand/ion	120	33	36
Molécules d'eau	1573	1188	1272
Facteur B moyen ($Å^2$)	15,4	15,6	17,8
Déviation r.m.s.			
Longueur de liaison (Å)	0,012	0,011	0,013
Angle de liaison (°)	1,82	1,58	1,74

Tableau 20 : Statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de SHWaglu-ACA, SHWaglu-CST et SHWaglu-DJN. Les valeurs pour la coquille de plus haute résolution sont indiquées entre parenthèses.

b. En complexe avec l'acarbose

La structure de SHWglu en complexe avec l'acarbose a été résolue à 1,85 Å de résolution (Figure 96) et contient 1314 résidus répartis en deux molécules A et B, 1573 molécules d'eau, quatre molécules de glycérol, deux molécules d'acarbose, deux ions calcium et six ions chlore. L'unité valiénamine (cycle A de l'acarbose) est localisée au plus profond de la poche du site actif (au niveau du sous-site -1). Au niveau de l'ion calcium présent dans le site actif, les molécules d'eau impliquées dans des liaisons de coordination avec cet ion et visibles dans la structure de l'enzyme seule sont remplacées par des atomes de l'acarbose à savoir : O2A du cycle A, N4B et C3B du cycle B (Figure 96). Nous avons pu constater que la présence d'une lysine en position 434 obstrue complètement le fond de la

poche catalytique et exclut de fait l'existence d'un potentiel sous-site -2. Un mutant K434A (décrit plus loin) a donc été envisagé pour comprendre le rôle réel de cette lysine.

Remarquablement, dans la structure de la molécule native seule, des molécules d'eau sont retrouvées exactement aux mêmes positions que les groupements hydroxyles de l'unité valiénamine de l'acarbose (Figure 96). Il est souvent observé que les oxygènes des molécules d'eau miment les oxygènes des sucres dans les structures cristallines de GH. Ceci suggère fortement que les liaisons hydrogène au fond de la poche catalytique sont nombreuses et cruciales pour la reconnaissance du glucose au niveau du sous-site -1. Le constat de l'importance de ces liaisons pour ce sous-site a déjà été dressé chez son homologue structural (Kitamura *et al.*, 2008). *A contrario*, au niveau du sous-site +1, les positions des molécules d'eau présentes dans la structure de l'enzyme seule ne correspondent pas exactement à celles des groupements hydroxyles de l'unité B de l'acarbose. Cela pourrait expliquer la large spécificité de substrat de SHWaglu qui hydrolyse non seulement des liaisons glucosidiques de type α-(1,4), mais aussi α-(1,6), α-(1,2) et α-(1,3). Au niveau du sous-site +2, l'interaction de ce glucose (cycle C de l'acarbose) est renforcé par une interaction hydrophobe de type « stacking » avec le tryptophane 513. Le glucose en sous-site +3 de l'acarbose (cycle D) est situé dans la poche catalytique tout en étant très proche de la surface de la protéine, et donc du solvant. Ainsi, une seule liaison hydrogène avec SHWaglu est répertoriée au niveau de la glycine 514 (Tableau 21). En conséquence, les substrats plus longs qu'un tétra-saccharide doivent vraisemblablement dépasser dans le solvant.

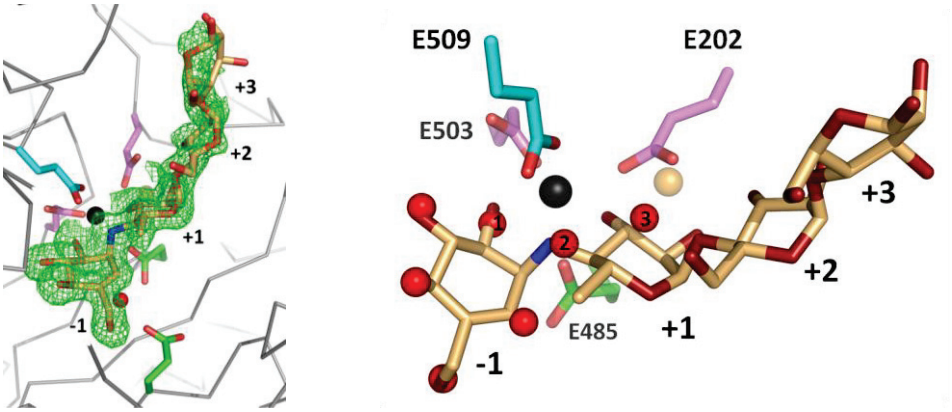


Figure 96 : A gauche, vue rapprochée centrée sur le site actif de SHWaglu en complexe avec l'acarbose. Le squelette carboné de la structure 3D de SHWaglu est représenté en ligne grise. Les résidus catalytiques base (E406 et E485) et acide (E509) sont colorés en vert et cyan respectivement. Les résidus (E202 et E503) déterminant des liaisons de coordination avec l'ion calcium (sphère grise) sont colorés en rose et le ligand est coloré en beige. La densité électronique F_o-F_c à 1,85 Å de résolution correspondant à la présence du ligand est contourée à 3σ. **A droite, vue rapprochée centrée sur le site actif de SHWaglu-ACA superposé à SHWaglu native.** Les molécules d'eau présentes dans la structure de SHW native sont représentées en sphères rouges. L'ion Cl^- est représenté en sphère orange.

Atome ACA	Atome SHWaglu	Distance (Å)
O2A	Ca^{2+}	2,00
O2A	Eau 331	2,95
O3A	Eau 331	2,60
O3A	NZ-K434	2,85
O3A	NE1-W318	3,00
O4A	NE1-W318	3,00
O4A	OE1-E359	2,55
O4A	OE2-E359	3,45
O6A	NE2-H404	3,20
O6A	OE1-E406	2,80
N4B	OE2-E509	2,75
N4B	Ca^{2+}	2,30
O2B	OE1-E202	2,90
O2B	NH1-R200	2,75
O2B	NH2-R200	3,45
O3B	Ca^{2+}	2,00
O3B	OE2-E202	2,50
O2C	ND2-N199	3,40
O3C	OD1-N199	2,50
O3C	ND2-N199	3,35
O3C	OE1-E202	3,4
O3C	OH-Y510	2,85
O6D	O-G514	2,85

Tableau 21 : Récapitulatif des interactions et de leurs distances entre l'acarbose et SHWaglu.

c. En complexe avec la castanospermine

La structure SHWaglu en complexe avec la castanospermine a été résolue à 2 Å de résolution et contient 1314 résidus répartis en deux molécules A et B, 1188 molécules d'eau, deux molécules de castanospermine, deux ions calcium et cinq ions chlore. La conformation adoptée par la castanospermine a été décrite comme un potentiel précurseur de l'état de transition (Ravaud *et al.*, 2007). La carte de densité électronique $2F_o$-F_c a permis de placer la molécule de CST dans chaque monomère de l'unité asymétrique sans aucune ambiguïté (Figure 97).

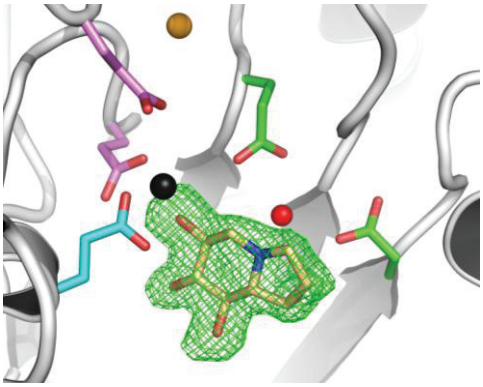


Figure 97 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec la castanospermine. SHWaglu est représentée en ruban gris. Les codes couleurs sont identiques à la Figure 96. La densité électronique F_o-F_c à 2 Å de résolution correspondant à la présence du ligand est contourée à 3,2σ.

Atome CST	Atome SHWaglu	Distance (Å)
N1	Eau 852	2,70
O2	Ca^{2+}	1,90
O2	Eau 272	3,00
O3	NE1-W318	3,05
O3	NZ-K434	2,70
O3	ND1-H484	3,50
O3	Eau 272	2,70
O4	NE1-W318	3,05
O4	NE2-H322	2,90
O4	OE1-E359	2,55
O4	OE2-E359	3,15
O6	OE1-E406	3,10
O6	NE2-H404	3,30
O6	OE1-E359	3,50
O6	Eau 852	3,50

Tableau 22 : Récapitulatif des interactions et de leurs distances entre la castanospermine et SHWaglu.

d. En complexe avec la déoxynojirimycine

La structure de SHWaglu en complexe avec la déoxynojirimycine a été résolue à 1,95 Å de résolution et contient 1314 résidus répartis en deux molécules A et B, 1272 molécules d'eau, une molécule de glycérol, deux ions calcium et six ions chlore. La molécule de déoxynojirimycine mime l'état de transition (Ravaud *et al.*, 2007). La carte de densité électronique $2F_o$-F_c a permis de placer la molécule de DJN dans chaque monomère de l'unité asymétrique sans aucune ambiguïté (Figure 98).

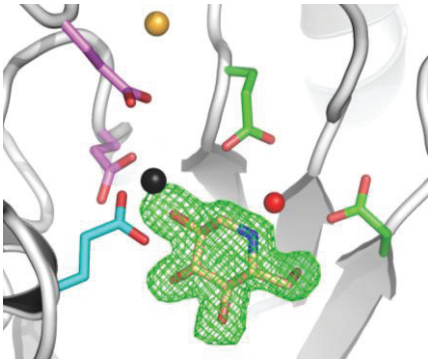


Figure 98 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec la déoxynojirimycine. Les codes couleurs sont identiques à la Figure 96. La densité électronique F_o-F_c à 1,95 Å de résolution correspondant à la présence du ligand est contourée à 4σ.

Atome DJN	Atome SHWaglu	Distance (Å)
N5	Eau 852	2,70
N5	Eau 1665	2,80
O2	Ca^{2+}	1,90
O2	Eau 272	3,15
O3	NE1-W318	3,00
O3	NZ-K434	2,80
O3	Eau 272	2,75
O4	NE1-W318	3,00
O4	NE2-H322	2,85
O4	OE1-E359	2,50
O4	OE2-E359	3,25
O6	OE1-E406	2,65
O6	NE2-H404	3,35
O6	Eau 852	3,20

Tableau 23 : Récapitulatif des interactions et de leurs distances entre la déoxynojirimycine et SHWaglu.

4. Études structurales de SHWaglu native en complexe avec des substrats naturels

Juste avant d'être congelés dans l'azote liquide et après 2h40 de trempage dans la solution de cryoprotection, des cristaux de la protéine native ont été mis en contact rapide avec des substrats naturels : du nigérose (NIG) et du kojibiose (KJB). Des trempages succincts ont également été réalisés avec de l'isomaltose ou du panose, mais sans succès. En effet, aucune densité électronique correspondant à un di- ou trisaccharide n'a été observée après traitement des données.

Les données synchrotron ont été enregistrées sur les lignes ID29 et ID23-1 de l'ESRF à une même résolution de 1,9 Å. Les données ont été intégrées et moyennées avec les programmes XDS et XSCALE (Kabsch, 2010). Le problème des phases a été résolu par remplacement moléculaire en employant le programme MolRep (Vagin & Teplyakov, 2010) et en utilisant, comme modèle guide, la structure de l'enzyme native affinée. Les modèles de ligands trouvés dans la base de données HIC-Up ont été insérés manuellement dans les cartes de densité électronique correspondantes. Comme pour la structure native, les 47 premiers résidus ne sont pas visibles en densité.

La qualité des structures 3D, en termes de géométrie et d'erreurs de coordonnées, a été examinée avec les programmes PROCHECK (Laskowski *et al.*, 1993) et WHATCHECK (Hooft *et al.*, 1996). L'ensemble des statistiques d'enregistrement et de traitement des données est présenté dans le Tableau 24 suivant.

	SHWaglu-NIG	**SHWaglu-KJB**
Enregistrement de données		
Ligne	ID29 (ESRF)	ID23-1 (ESRF)
Longueur d'onde (Å)	0,97625	1,28287
Groupe d'espace	$P2_12_12_1$	$P2_12_12_1$
Paramètres de maille		
a, b, c (Å)	100,81, 111,47, 129,16	101,12, 111,99, 129,61
α, β, γ (°)	90 90 90	90 90 90
Vm ($Å^3.Da^{-1}$)	2,22	2,25
Pourcentage de solvant (%)	44,7	45,3
Résolution (Å)	48,78-1,9 (1,94-1,9)	36,05-1,9 (1,95-1,9)
Nombre de réflexions totales	829653 (55248)	707180 (42473)
Nombres de réflexions uniques	114899 (8216)	116047 (7970)
R_{sym} (%)	8,5 (82,2)	9,7 (60,2)
I/σ (I)	17,72 (2,73)	18,5 (4,9)
Complétude (%)	99,2 (96,9)	99,2 (92,9)
Redondance	7,2 (6,7)	6,1 (5,3)
Affinement		
R_{factor}/R_{free} (%)	15,4 / 19,5	15,4 / 19,3
Nombre d'atomes		
Protéine (non-hydrogène)	10646	10643
Ligand/ion	46	86
Molécules d'eau	1046	968
Facteur B moyen ($Å^2$)	24,7	22,5
Déviation r.m.s.		
Longueur de liaison (Å)	0,013	0,011
Angle de liaison (°)	1,85	1,75

Tableau 24 : Statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de SHWaglu-NIG et SHWaglu-KJB. Les valeurs pour la coquille de plus haute résolution sont indiquées entre parenthèses.

a. Trempage avec le nigerose

La structure cristallographique de SHWaglu dont le cristal a été trempé dans une solution de nigérose contient 1314 résidus répartis en deux molécules A et B, 1046 molécules d'eau, deux molécules de glycérol, deux ions calcium et huit ions chlore. Sa résolution à 1,9 Å révèle la présence d'une molécule de glucose dans chacun des sites actifs des deux molécules de l'unité asymétrique. La carte de densité électronique $2F_o$-F_c a permis de placer sans ambiguïté ces molécules de glucose (Figure 99). L'enzyme cristallisée a hydrolysé son substrat, prouvant ainsi qu'elle est complètement active sous forme cristalline. Le constat de la présence du produit d'hydrolyse dans le site actif confirme bien l'efficacité catalytique élevée de SHWglu pour ce substrat naturel.

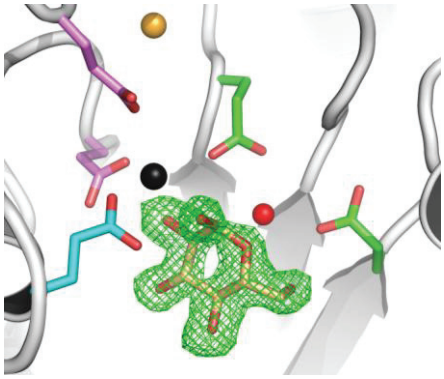


Figure 99 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec le glucose. Les codes couleurs sont identiques à la Figure 96. La densité électronique F_o-F_c à 1,9 Å de résolution correspondant à la présence du ligand est contourée à 3,2σ.

Atome GLC	Atome SHWaglu	Distance (Å)
O1	Ca^{2+}	2,90
O1	Eau 15	2,85
O1	OE1-E509	2,85
O1	OE2-E509	3,55
O2	OE1-E503	3,05
O2	OE2-E503	3,35
O2	Eau 625	3,15
O2	OE2-E485	3,25
O2	OE2-E509	3 ,35
O2	ND1-H484	2,70
O2	Ca^{2+}	2,30
O3	Eau 625	2,60
O3	NE1-W318	2,95
O3	ND1-H484	3,60
O3	NZ-K434	2,80
O4	NE1-W318	3,00
O4	NE2-H322	2,75
O4	OE1-E359	2,50
O4	OE2-E359	3,30
O5	Eau 297	2,50
O6	Eau 297	3,30
O6	OE1-E406	2,75
O6	NE2-H404	3,20

Tableau 25 : Récapitulatif des interactions et de leurs distances entre le glucose et SHWaglu.

b. En complexe avec le kojibiose

La structure cristallographique de SHWaglu dont le cristal a été trempé dans une solution de kojibiose contient 1314 résidus répartis en deux molécules A et B, 968 molécules d'eau, une molécule de glycérol, deux molécules de glucose, deux ions calcium et huit ions chlore.

Sa résolution à 1,9 Å révèle la présence d'une densité électronique correspondant à un disaccharide dans chacun des sites actifs des deux molécules de l'unité asymétrique. Cette carte de densité électronique $2F_o$-F_c a permis de placer sans ambiguïté ces molécules de kojibiose.

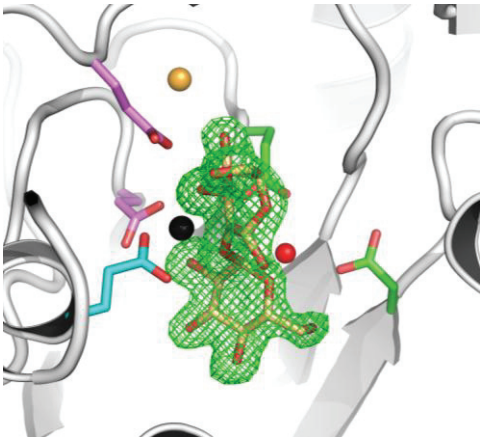


Figure 100 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec le kojibiose. Les codes couleurs sont identiques à la Figure 96. La densité électronique F_o-F_c à 1,9 Å de résolution correspondant à la présence du ligand est contourée à 4σ.

Atome KJB	Atome SHWaglu	Distance (Å)
O1A	Ca^{2+}	2,30
O2A	Ca^{2+}	2,05
O2A	Eau 1	3,00
O2A	OE1-E503	2,90
O2A	OE2-E503	2,95
O3A	NE1-W318	2,90
O3A	NZ-K434	2,85
O3A	Eau 1	2,70
O4A	NE2-H322	2,80
O4A	OE1-E359	2,55
O4A	OE2-E359	3,30
O4A	NE1-W318	2,95
O5A	Eau 159	2,95
O6A	NE2-H404	3,10
O6A	OE1-E406	3,00
O3B	OE1-E503	3,05
O3B	OE2-E485	2,60
O3B	OE2-E202	2,40
O3B	Ca^{2+}	2,05
O4B	OE1-E202	2,80
O4B	OE2-E202	3,30
O4B	ND2-N199	3,50
O4B	NH2-R200	2,65
O5B	Eau 1298	3,35
O6B	Eau 1298	3,45

Tableau 26 : Récapitulatif des interactions et de leurs distances entre le kojibiose et SHWaglu.

C. Comparaison de structures

Dans toutes les structures résolues, nous retrouvons la présence de la molécule d'eau supposée catalytique entre Glu406 et Glu485 ainsi que l'ion Ca^{2+} qui détermine les liaisons de coordinations décrites plus haut. En comparant ces différentes structures, nous avons constaté que tous les résidus de la poche catalytique se superposent parfaitement, à l'exception de 3 résidus à l'entrée de cette poche (Trp369, Phe370 et His371). La différence la plus remarquable se situe au niveau de la phénylalanine qui change complètement d'orientation en fonction du ligand présent dans le site actif. En effet, deux configurations possibles sont retrouvées : une « ouverte » et une autre « fermée ». Dans cette deuxième configuration, la Phe370 interagit avec le Trp369 par « π-stacking » avec une distance entre les deux carbones les plus proches de 3,4 Å, qui est la valeur minimale interatomique entre deux cycles aromatiques (McGaughey *et al.*, 1998). La présence d'un clamp aromatique à l'entrée du site actif chez des tréhalulose synthases, appartenant à la GH13, ont déjà été rapportées (Ravaud *et al.*, 2007; Ravaud *et al.*, 2009; Lipski *et al.*, 2013). Tout comme dans ces travaux de thèse, le clamp aromatique peut adopter deux conformations possibles, indiquant un mouvement concerté possible lors de la réaction enzymatique. Il a été supposé que ce clamp contrôle partiellement la spécificité de réaction et l'entrée-relargage des substrats/produits (Ravaud *et al.*, 2007).

La superposition de nos structures cristallines avec celles de SusB montre que chez l'homologue, la phénylalanine correspondante est orientée selon une conformation intermédiaire qui ne varie pas entre la protéine native et les différents complexes (Figure 101).

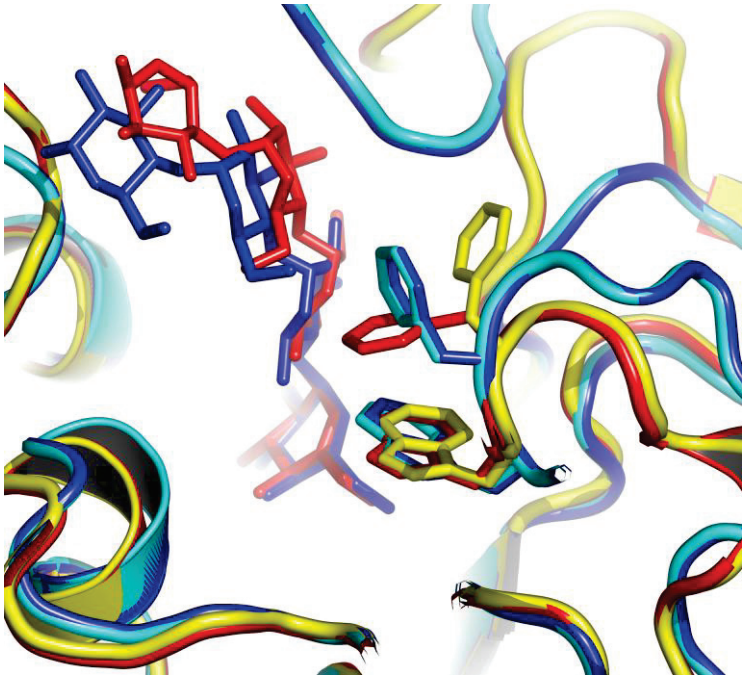


Figure 101 : Superposition des structures représentées en ruban de SHWaglu native (jaune), SHWaglu en complexe avec l'acarbose (rouge), SusB native (cyan) et SusB en complexe avec l'acarbose (bleu). Le tryptophane 369 (400 chez SusB), la phénylalanine 370 (401 chez SusB) et l'acarbose sont représentés en bâtons.

Les structures SHWaglu-ACA et SHWaglu-KJB ont également été superposées afin de comparer les interactions des deux ligands dans le site actif (Figure 102). Comme attendu, les cycles (osidique ou le mimant) en sous-site -1 sont parfaitement superposés entre le kojibiose et l'acarbose. Ceci confirme la très grande spécificité d'interaction déterminant la reconnaissance du glucose, régie par de nombreuses liaisons hydrogène.

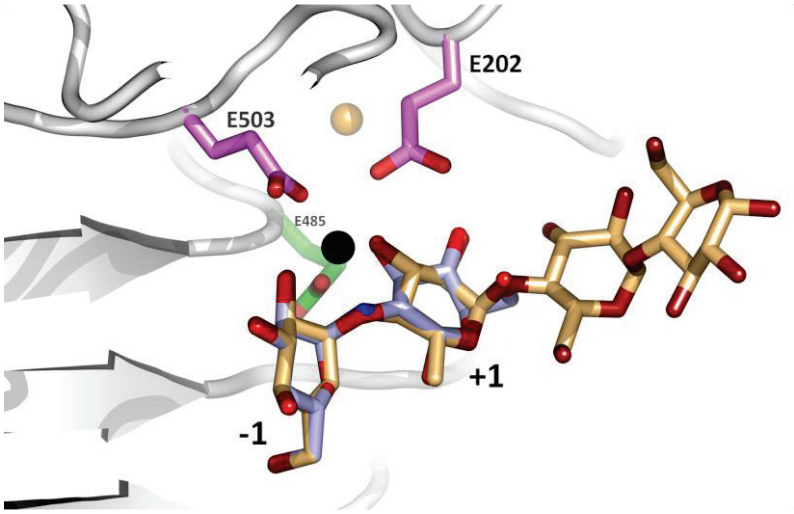


Figure 102 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec l'acarbose superposé à SHWaglu en complexe avec le kojibiose. L'acarbose et le kojibiose sont colorés en beige et en lilas, respectivement. Les ions Ca^{2+} et Cl^- sont représentés en sphère noire et jaune, respectivement.

De façon plus remarquable, bien que la liaison osidique engagée ne soit pas la même (α-(1,4) ou α-(1,2)), tous les oxygènes du cycle de l'acarbose et du kojibiose en sous-site +1 sont également parfaitement superposés (Figure 102). Ceci traduirait qu'une orientation particulière de ce cycle est privilégiée même si, comparativement au sous-site -1, moins d'interactions avec l'enzyme sont dénombrées. La seule différence observée est la conformation du cycle en +1 : celui de l'acarbose est en conformation chaise de type $^4C_1$ alors que celui du kojibiose est en conformation enveloppe de type ^4E (Figure 103). Les multiples liaisons hydrogène avec les O3 (résidus E202 et E503) et O4 (résidus N199, R200 et E202) du kojibiose « forcent » cette conformation ^4E du glucopyranose qui est énergétiquement moins favorable que celle en $^4C_1$. Pour rappel, la constante catalytique k_{cat} de SHWaglu pour le kojibiose est environ 20 fois supérieure à celle mesurée pour le maltose alors que le K_m n'est pas modifié. Cette différence apparait donc être d'une importance majeure : il semblerait que la conformation enveloppe du cycle pyrane, induite par les interactions des hydroxyls, soit l'état de transition. Cette hypothèse a déjà été soulevée dans une étude sur une hyaluronidase qui a démontré une réduction drastique du k_{cat} (de 2000 fois) lorsque l'acide aspartique provoquant la formation de la conformation ^4E du substrat était muté en alanine. Ceci présume que ce résidu chez cette enzyme joue un rôle dans la formation de l'état de transition (Rao *et al.*, 2006).

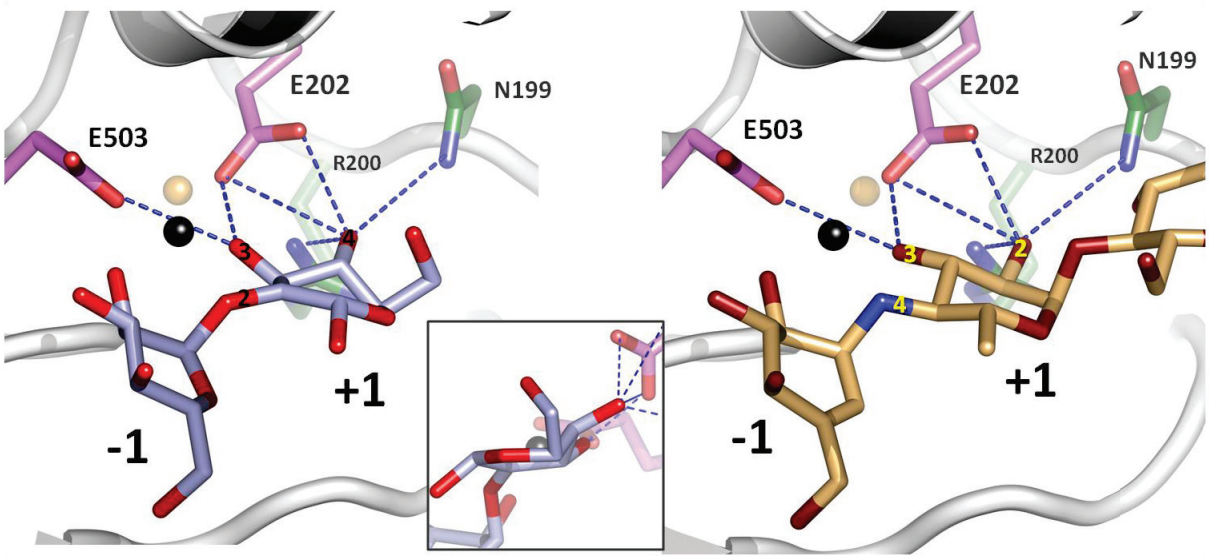


Figure 103 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec le kojibiose (à gauche) et avec l'acarbose (à droite). Les codes couleurs sont conservés par rapport à la Figure 102. Le cycle dans le sous-site +1 n'adopte pas la même conformation : ^4E pour le kojibiose et $^4C_1$ pour l'acarbose. L'encadré est centré sur l'unité glucosidique en conformation ^4E du kojibiose et observée selon une autre orientation.

Nous nous sommes également intéressés aux ions chlore présents dans les différentes structures : trois sont strictement conservés parmi toutes les structures résolues. De plus, l'un d'eux occupe une position particulièrement remarquable. En effet, ce dernier se situe à proximité du site actif, très près de l'ion calcium et du substrat, et est coordoné par trois liaisons par l'intermédiaire de l'Arg 200 et la Tyr201 (Figure 104). Il est à noter que cet ion n'est pas retouvé chez son homologue SusB. Ceci tend à indiquer que cet ion joue un rôle majeur dans l'hydrolyse chez l'enzyme psychrotolérante SHWaglu, ce qui est en accord avec les études enzymatiques.

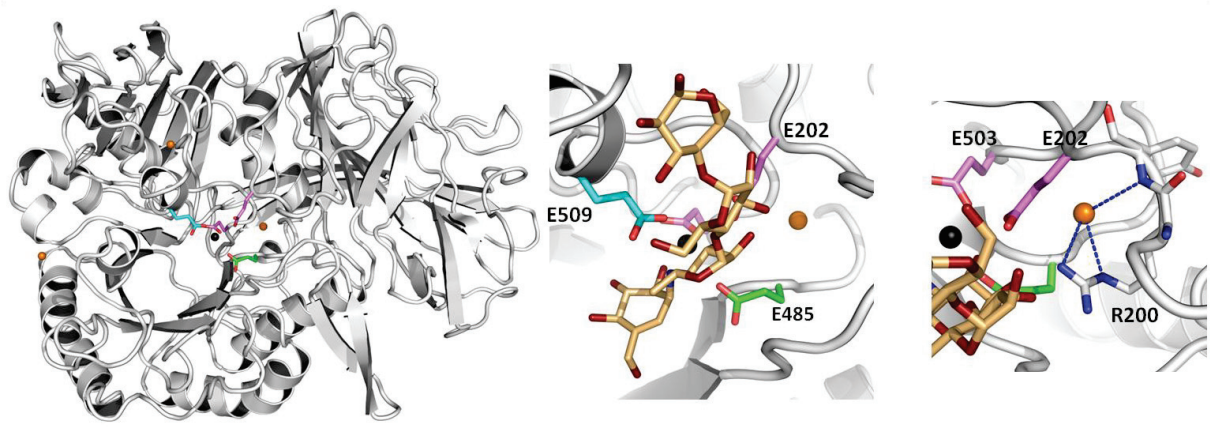


Figure 104 : A gauche, représentation en ruban de la structure de SHWaglu et mise en évidence des ions Cl^- strictement conservés dans toutes les structures résolues. *Les ions Cl^- sont représentés en sphère orange.* **Au centre, vue rapprochée centrée sur le site actif. A droite, vue rapprochée centrée sur l'ion Cl^- proche du site actif.**

D. Docking avec le panose

Des trempages de cristaux de SHWaglu native ont été réalisés en présence de panose, malheureusement sans succès. En effet, après traitement des données, aucune densité électronique correspondant au trisaccharide n'a été observée. Au vu de notre étude enzymatique montrant que la meilleure efficacité catalytique est obtenue pour le panose, nous avons procédé à des expériences de docking en utilisant comme ligand ce substrat pour essayer de définir le mode d'interaction de ce sucre avec l'enzyme.

Le docking a été réalisé en ciblant la recherche au niveau de la poche du site actif. Il en est ressorti neuf solutions de plus fortes probabilités. Parmi celles-ci, deux prédisent le panose avec son extrémité réductrice placée dans le sous-site -1 et ont donc été exclues puisqu'elles ne répondent pas à la convention des GH. En effet, depuis l'adoption de cette nomenclature, les sous-sites sont définis de *-n* à *+n* où *-n* représente l'extrémité non-réductrice et *+n* l'extrémité réductrice (Davies *et al.*, 1997). Une troisième a également été éliminée car le cycle glucose de l'extrémité non-réductrice du panose (GLC1), certes situé dans le sous-site -1, apparaît perpendiculaire à tous les ligands observés dans les structures

cristallines précédemment résolues. Or, il s'avère que la reconnaissance du glucose au niveau de ce sous-site semble être régie par de nombreuses liaisons hydrogène conférant une très grande spécificité d'interaction. Parmi les solutions restantes, trois montrent le ligand à l'entrée du site actif et non pas au niveau du site catalytique. Les trois dernières (solutions numérotées 316, 384 et 501) exposent le panose avec le glucose de l'extrémité non-réductrice dans le sous-site -1. Ce cycle osidique présente une orientation très similaire entre ces trois solutions mais aussi aux ligands présents dans les structures résolues et présentées précédemment. Les énergies libres de liaison estimées par AutoDOCK de ces trois solutions sont de -8, -7 et -6,5 kcal/mol, respectivement, avec une constante de dissociation de l'ordre du micromolaire. Les différences majeures entre ces solutions sont surtout retrouvées pour la fixation des glucoses au niveau des sous-sites +1 et +2 (GLC2 et GLC3). Le glucose GLC2 de la solution 384 présente une interaction hydrophobe avec le Trp513 et le GLC3 forme des liaisons hydrogène avec les Asn199, Arg200, Glu202 et Asp440. La solution 501 ne présente aucune interaction hydrophobe et ne forme que des liaisons hydrogène avec les résidus Asn199, Glu202 et Tyr510 *via* le GLC2, et le Trp369 *via* le GLC3. Pour la solution 316, le glucose GLC2 interagit par des liaisons hydrogène avec les Tyr510 et Glu202, et le GLC3 avec la His322. De plus, ce dernier forme une interaction hydrophobe avec le Trp513 et se superpose avec le glycérol retrouvé dans la structure native. Il a déjà été rapporté que le glycérol pouvait mimer une partie du glucose cyclique (Burmeister *et al.*, 1997; Lariviere *et al.*, 2003; Vester-Christensen *et al.*, 2010; Xia *et al.*, 2012; Lipski *et al.*, 2013). Enfin, la liaison glucosidique entre GLC1 et GLC2 de cette solution 316 est orientée d'une façon très similaire à la liaison *N*-glycosidique de l'acarbose de la structure du complexe SHWaglu-ACA. Toutes ces observations tendent à présenter cette dernière solution comme la plus probable (Figure 105 et Figure 106).

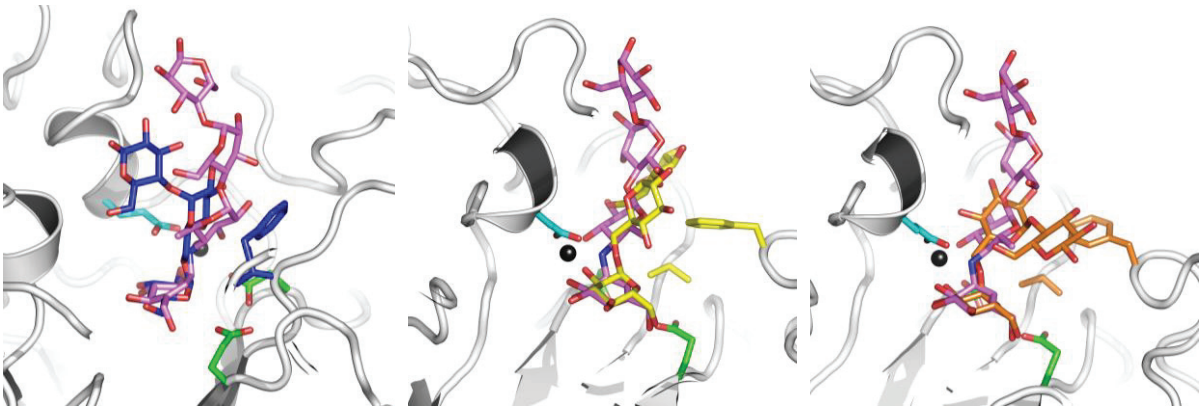


Figure 105 : Superposition de la structure 3D de SHWaglu en complexe avec l'acarbose et des trois solutions de docking retenues. Les squelettes carbonés et l'ion calcium sont représentés en ruban gris et sous forme de sphère noire respectivement. Les résidus 406 et 485 sont colorés en vert. Le résidu 509 et l'acarbose sont colorés en cyan et en rose respectivement. Les trois solutions 316, 384 et 501 représentées en bâton sont colorées en bleu, jaune et orange, respectivement.

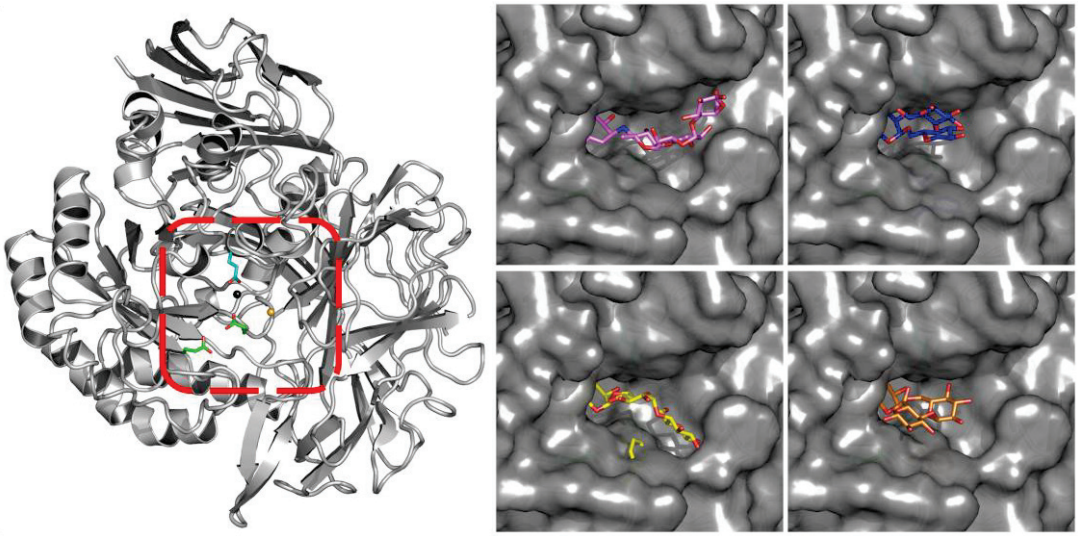


Figure 106 : Représentation et comparaison des solutions retenues issues du docking du panose avec l'acarbose en complexe avec SHWaglu. SHWaglu est visualisée sous forme de surface moléculaire grise centrée sur la poche catalytique et selon la même orientation. Les codes couleurs des ligands et des résidus catalytiques sont identiques à la Figure 105. L'ion Cl^- est représenté en sphère orange.

IV. Études structurales de mutants

A. Biologie moléculaire, expression et purification

Dans un premier temps, il a fallu apporter les mutations dans le gène d'intérêt *SHWaglu*. Des mutagenèses dirigées à partir du plasmide pET15b contenant le gène *SHWaglu* ont été réalisées comme décrit dans le chapitre « Matériel et Méthodes » en utilisant des amorces spécifiques. Ainsi, dix mutants différents ont été obtenus et transformés avec succès dans une souche d'expression *E. coli* BL21 (DE3) : E202A, E202Q, E359A, E359Q, E406A, E406Q, K434A, E485Q, E503Q et E509Q. Ces mutants ont été construits dans le but de comprendre le rôle réel de ces résidus situés dans le site actif (Figure 107). Toutes les constructions ont été validées par séquençage. Nous ne détaillerons ici que les mutants K434A et E485Q (Tableau 27). Les mêmes protocoles d'expression et de purification que ceux utilisés pour SHWaglu ont été réalisés : une chromatographie d'affinité au nickel suivie d'une échangeuse d'anions.

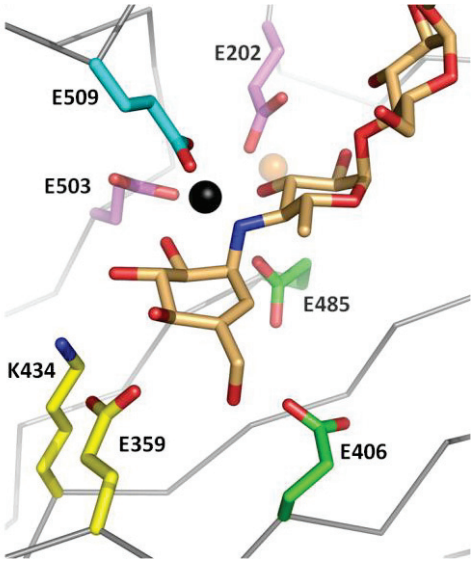


Figure 107 : Vue rapprochée centrée sur le site actif de SHWaglu en complexe avec l'acarbose. Les résidus à muter sont représentés en bâton. Les ions Ca^{2+} et Cl^- sont représentés en sphères grise et orange, respectivement.

NOM	SEQUENCE DE L'AMORCE 5'→ 3'
OL-K434A-Direct	GGCGTGACCCAAGTC**GC**AACCGG
OL-K434A-Reverse	GCCACATAACCGGTT**GC**GACTTGGG
OL-E485Q-Direct	CATCAACAGCCAT**C**AGCCGATTAAAG
OL- E485Q -Reverse	CCCAGTATCTTTAATCGGCT**G**ATGGC

Tableau 27 : Amorces utilisées pour la mutagenèse dirigée par PCR de *SHWaglu*. Les mutations apportées sont surlignées en gris.

B. Études cristallographiques des mutants E485Q et K434A

Après leur expression et purification respectives, $SHWaglu_{K434A}$ et $SHWaglu_{E485Q}$ ont été concentrées jusqu'à 5,5 mg/ml. Une première série d'essais de cristallogenèse a été réalisée en utilisant les mêmes conditions de cristallisation que SHWaglu native (0,18 M à 0,22 M LiCl, 18% à 22% PEG 3350). Aucun cristal n'est apparu. Une deuxième série a été effectuée dans ces mêmes conditions, mais cette fois, en utilisant la technique de micro-ensemencement. A l'aide d'une moustache de chat, les gouttes pré-équilibrées ont été ensemencées par des points de nucléation obtenus par broyage d'un cristal de SHWaglu native. Très rapidement (au bout de 24h), des cristaux de morphologie similaire à ceux de SHWaglu native sont apparus.

Les données synchrotron ont été enregistrées sur les lignes Proxima-1 de SOLEIL et ID23-2 de l'ESRF à des résolutions de 2,25 Å et 2,7 Å, respectivement. Elles ont été intégrées et moyennées avec les programmes XDS et XSCALE (Kabsch, 2010). Le problème des phases a été résolu par remplacement moléculaire en employant le programme MolRep (Vagin &

Teplyakov, 2010) et en utilisant, comme modèle guide, la structure de la molécule native affinée. Comme pour la structure native, les 47 premiers résidus ne sont pas visibles en densité électronique.

La qualité de la structure 3D, en termes de géométrie et d'erreurs de coordonnées, a été examinée avec les programmes PROCHECK (Laskowski *et al.*, 1993) et WHATCHECK (Hooft *et al.*, 1996). L'ensemble des statistiques d'enregistrement et de traitement des données est présenté dans le Tableau 28 suivant et la qualité de la carte de densité électronique finale $2F_o$-F_c de $SHWaglu_{E485Q}$ est visible dans la Figure 108.

	$SHWaglu_{E485Q}$	**$SHWaglu_{K434A}$**
Enregistrement de données		
Ligne	Proxima-1 (SOLEIL)	ID23-2 (ESRF)
Longueur d'onde (Å)	0,97911	0,87260
Groupe d'espace	$P2_12_12_1$	$P2_12_12_1$
Paramètres de maille		
a, b, c (Å)	100,79, 112,01, 129,39	100,51, 112,34, 129,85
α, β, γ (°)	90 90 90	90 90 90
Vm ($Å^3.Da^{-1}$)	2,24	2,25
Pourcentage de solvant (%)	45	45,3
Résolution (Å)	48,97-2,25 (2,31-2,25)	46,87-2,7 (2,75-2,7)
Nombre de réflexions totales	419809 (29241)	306817 (16342)
Nombres de réflexions uniques	69900 (4940)	41060 (2158)
R_{sym} (%)	19,6 (74,9)	18,9 (50,7)
I/σ (I)	9 (2,7)	12,9 (5,1)
Complétude (%)	99,3 (96)	100 (100)
Redondance	6 (5,9)	7,5 (7,6)
Affinement		
R_{factor}/R_{free} (%)	15,2 / 21,8	16,2 [a] / 22,7 [a]
Nombre d'atomes		
Protéine (non-hydrogène)	10566	- [a]
Ligand/ion	16	- [a]
Molécules d'eau	1068	- [a]
Facteur B moyen ($Å^2$)	20,3	- [a]
Déviation r.m.s.		
Longueur de liaison (Å)	0,015	0,013 [a]
Angle de liaison (°)	1,29	1,48 [a]

Tableau 28 : Statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de $SHWaglu_{E485Q}$ et $SHWaglu_{K434A}$. Les valeurs pour la coquille de plus haute résolution sont indiquées entre parenthèses. [a] : la structure est en cours d'affinement.

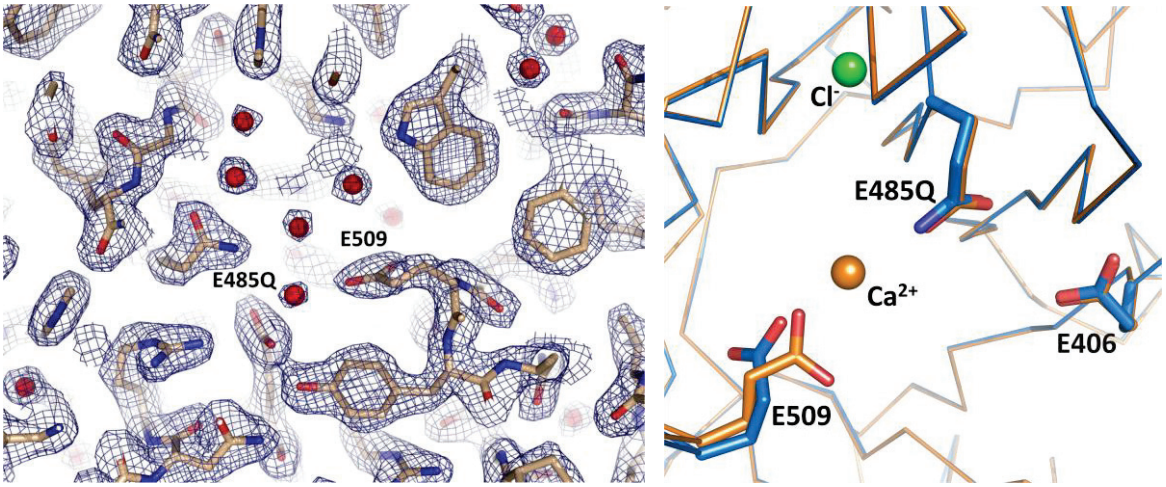


Figure 108 : *A gauche* : carte de densité électronique $2F_o$-F_c à 2,25 Å de résolution (contourée à 1,2σ) de SHWaglu-E485Q centrée sur le site actif. Les atomes de carbone, d'oxygène et d'azote sont respectivement représentés en beige, rouge et bleu. Les molécules d'eau sont indiquées en sphères rouges. ***A droite* : superposition de SHWaglu native (orange) et de SHWaglu-E485Q (bleu) centrée sur le site actif.** Les squelettes carbonés des structures 3D sont représentés en ligne et les résidus catalytiques (résidus 406, 485 et 509) en bâtons. L'ion Cl^- est représenté en sphère verte.

La structure cristallographique de $SHWaglu_{E485Q}$ à une résolution de 2,25 Å révèle la présence de 1314 résidus d'acides aminés répartis en deux monomères A et B, 1068 molécules d'eau, une molécule de glycérol et dix ions chlore. Nous noterons l'absence des ions calcium retrouvés dans le site actif jusqu'à présent (Figure 108). La superposition de la structure de l'enzyme native avec celle du mutant E485Q montre un changement d'orientation du résidu catalytique acide E509 (Figure 108). Ainsi, ces constatations révèlent que le résidu E485 est indispensable à la coordination de l'ion calcium dans le site actif et que ce dernier est essentiel pour orienter correctement le résidu catalytique acide. Ceci confirme que la présence de cet ion est obligatoire pour la réaction enzymatique de SHWaglu, ceci étant en accord avec nos études fonctionnelles de l'enzyme.

La structure cristallographique de $SHWaglu_{K434A}$, résolue à 2,7 Å, contient, elle, 1314 résidus d'acides aminés répartis en deux monomères A et B et deux ions calcium. Elle révèle la réduction de l'encombrement de la poche catalytique avec la création d'un espace au niveau du sous-site -1 (Figure 109). Ainsi, ce seul et unique résidu est responsable de l'obstruction du fond de la poche. Sans ce dernier, le tonneau $(\beta/\alpha)_8$ est quasi-traversant, ouvrant la perspective de création d'éventuels nouveaux sous-sites -2 et -3.

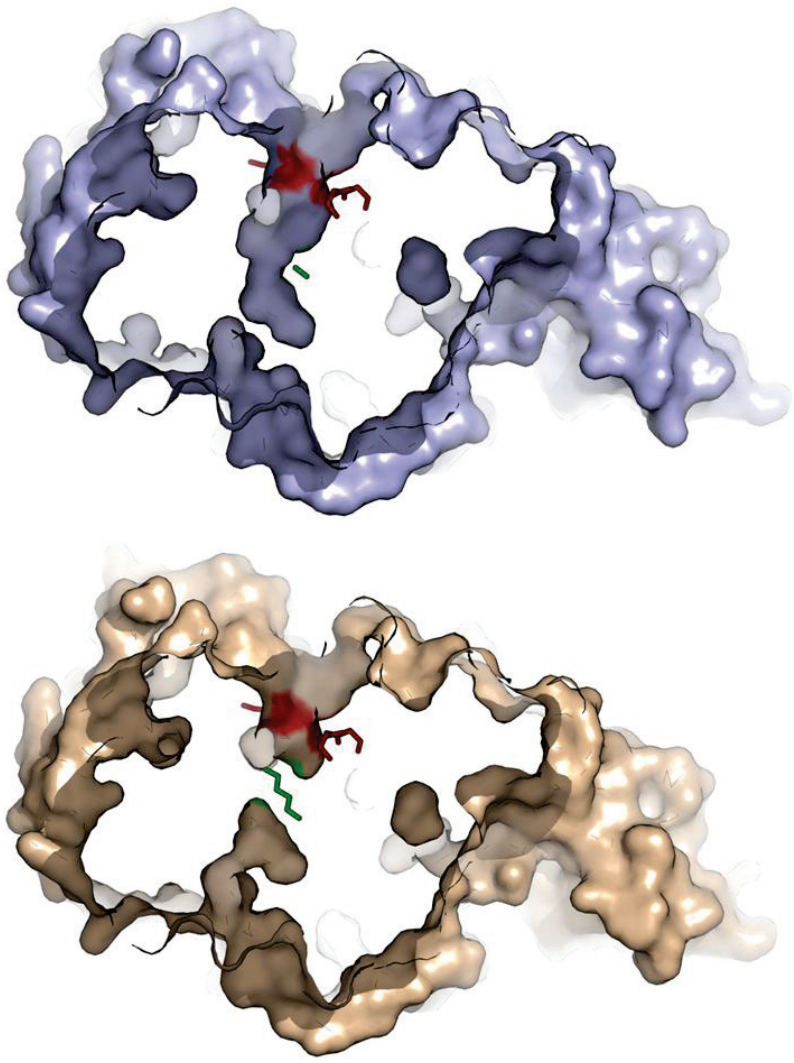


Figure 109 : Coupes transversales des surfaces moléculaires de SHWaglu$_{K434A}$ et de SHWaglu colorés en bleu et beige, respectivement. Les résidus catalytiques (E406, E485 et E509) et le résidu 434 sont représentés en bâton et colorés en rouge et vert, respectivement. Les deux molécules sont présentées selon la même orientation. Un prolongement de la poche catalytique dans le cas de SHWaglu$_{K434A}$ est bien observé.

Conclusions & Perspectives

Au cours de ces travaux de thèse, nous nous sommes intéressés à l'étude structurale et fonctionnelle de trois α-glucosidases bactériennes dans le but d'approfondir les connaissances sur cette classe d'enzymes. Ces études s'inscrivent dans le contexte des pro- et prébiotiques et participent à une meilleure compréhension du fonctionnement de ces enzymes au niveau fondamental. Au-delà de ces aspects, des applications potentielles (notamment agro-alimentaires) pourraient en découler par l'optimisation de leurs propriétés grâce à des travaux d'ingénierie protéique dirigée. Les trois études menées en parallèle, dans l'équipe de Biocristallographie et Biologie Structurale des Cibles Thérapeutiques, ont toutes été initiées durant cette thèse. Elles ont cependant des degrés d'avancement différents.

- α-Glucosidase de *Lactobacillus bulgaricus* ATCC11842

Les études bioinformatiques de 11842aglu ont permis d'attribuer cette enzyme à la famille GH31, mais également de révéler la présence d'un module « Carbohydrate Binding Module » appartenant à la famille CBM32 à l'extrémité C-terminale. Cette famille, jusqu'à présent l'une des plus diversifiées (Abbott *et al.*, 2008), est connue pour reconnaitre et interagir avec des ligands possédant au moins une unité de galactose (Ficko-Blean & Boraston, 2006; Abbott *et al.*, 2008; Ficko-Blean *et al.*, 2012). Le rôle de ce CBM, en tant qu' « attracteur » de l'enzyme vers son substrat préférentiel ou interactant avec les mucines de la muqueuse gastro-intestinale de l'hôte, n'a pas encore été confirmé pour 11842aglu. Pour ce faire, mais aussi dans le but de caractériser l'enzyme entière, il sera nécessaire d'exprimer et de purifier en grande quantité la protéine grâce aux protocoles établis lors de cette thèse. La combinaison des études fonctionnelles (comme la détermination de ses pH et température optimaux d'activité ou encore de ses substrats naturels) avec des études structurales (par cristallographie aux rayons X) permettra d'établir les relations structure/fonction/activité de cette enzyme appartenant à la famille GH31. Ainsi, la mise en évidence de déterminants structuraux et fonctionnels pourrait contribuer à la compréhension du rôle probiotique de *Lactobacillus bulgaricus*. Dans ce cadre, il serait également intéressant de réaliser une construction plasmidique dans le but de produire, d'extraire et de purifier le module CBM32 de 11842aglu seul afin de déterminer les interactants mais aussi de réaliser des essais de cristallogenèse. A terme, en se basant sur la (ou les) structure(s) obtenue(s), il serait envisageable d'élaborer de nouvelles constructions où les résidus impliqués dans l'interaction avec les ligands naturels du CBM (et de l'enzyme entière) seraient mutés dans le but d'observer des altérations d'interactions éventuelles.

- α-Glucosidase de *Lactococcus lactis* IL1403

Le clonage du gène *1403aglu* dans le vecteur d'expression pET15b a été réalisé avec succès. Les protocoles de surexpression et de purification de l'enzyme correspondante ont également pu être établis. Ces étapes fructueuses ont conduits aux études biochimiques et structurales de 1403aglu.

La caractérisation préliminaire de 1403aglu a permis de déterminer les conditions optimales de réaction de l'enzyme (à pH 7 et 30°C, sans apport supplémentaire de cations) ainsi que sa spécificité de substrat. Cependant, des études cinétiques plus approfondies pour en déterminer les constantes catalytiques ont démontré de très faibles constantes d'affinité et efficacités catalytiques. Ceci pourrait suggérer un éventuel problème de repliement de l'enzyme ou bien que cette GH13 nécessiterait un partenaire protéique pour être pleinement active. Il a toutefois été possible de définir que seules les liaisons interglucosidiques de type α-(1,4) étaient hydrolysées par 1403aglu et, plus particulièrement, les maltooligosaccharides dont le degré de polymérisation est faible.

De plus, au niveau génomique, le gène *1403aglu* est localisé à proximité de plusieurs gènes codant des enzymes et transporteurs impliqués dans la voie d'utilisation du maltose. Ces enzymes « génomiquement voisines » et potentiellement exprimées de manière concomitante chez *L. lactis* IL1403 permettraient l'hydrolyse des liaisons interglucosidiques de type α-(1,6) mais aussi α-(1,4) de polyosides ayant des degrés de polymérisation importants et/ou étant ramifiés. De fait, elles peuvent « prendre en charge » une grande partie des sucres impliqués dans cette voie métabolique. *L. lactis* IL1403 aurait donc évolué de sorte à ce que 1403aglu devienne très spécifique.

Un très vaste criblage de conditions de cristallogenèse a été mené en utilisant la protéine dans des tampons distincts, à plusieurs concentrations, mais aussi en présence de différents ligands potentiels et à des températures variées. Plusieurs cristaux de morphologies différentes ont ainsi été obtenus et différentes solutions de cryoprotection ont été testées. Malheureusement, aucun cristal n'a jusqu'à présent diffracté suffisamment les rayons X pour rendre possible la résolution de la structure de 1403aglu. Ce problème récurrent de diffraction pourrait être dû à un empilement cristallin insuffisamment ordonné ou à un pourcentage de solvant trop élevé dans le cristal, ou bien encore à un mauvais repliement de la protéine. Cette dernière hypothèse apparait la plus vraisemblable au vu des résultats de dichroïsme circulaire obtenus qui montrent des divergences dans les

proportions de structures secondaires calculées expérimetalement comparativement à celles théoriquement attendues. Toutefois, il serait intéressant de tester les cristaux en capillaire à 20°C pour affirmer sans ambiguïte que la mauvaise diffraction n'est pas due à un problème de cryoprotection. Dernièrement, des mesures de diffusion des rayons X aux petits angles (SAXS) sur la ligne BM29 à l'ESRF ont été réalisées afin d'obtenir des informations structurales à basse résolution. Ces données sont en cours de traitement.

- α-Glucosidase de *Shewanella* sp. ANA-3

L'étude de l'α-glucosidase de *Shewanella* sp. ANA-3 est la plus aboutie de toutes. Le clonage du gène *SHWaglu* dans le vecteur d'expression pET15b a été réalisé avec succès et les protocoles de surexpression et de purification ont été établis. Il a donc été possible de réaliser une étude enzymatique et structurale approfondie de cette enzyme appartenant à la famille GH97, unique car elle rassemble tant des enzymes agissant selon un mécanisme d'inversion que d'autres procédant par un mécanisme de rétention de la configuration du carbone anomérique.

Les influences du pH, de la température et des ions divalents ont été examinées. Les paramètres optimaux ont ainsi pu être déterminés : l'enzyme sous sa forme libre est pleinement active à pH 6,5 et à 50°C en présence d'ions calcium. Cependant, ses propriétés les plus notables sont ses besoins essentiels en ions chlorure (voir « Le rôle des ions chlorure » p. 155) et la conservation de son activité aux basses températures (50% à 4°C, voir « Détermination de la température optimale » p. 153). Cette étude a donc permis de mettre en évidence le caractère psychrotolérant de l'enzyme, ce qui est en accord avec l'environnement naturel de la bactérie. Une étude comparative réalisée avec l'enzyme immobilisée dans des billes d'alginate démontre une amélioration considérable des profils d'activité relative en fonction du pH ou de la température mais aussi de la stabilité lorsque l'enzyme est dans ces conditions expérimentales.

La détermination des constantes cinétiques a été réalisée pour chacun des substrats naturels caractérisés de l'enzyme mais également des constantes d'inhibition pour plusieurs ligands compétitifs (acarbose, castanospermine, déoxynojirimycine et glucose). Ainsi, SHWaglu est active sur un grand nombre de glucooligosaccharides avec une meilleure efficacité catalytique à basse température confirmant la psychrotolérance de l'enzyme (voir « Détermination des constantes catalytiques » p. 161). Cette large spécificité de l'enzyme semble cohérente avec la localisation génomique de *SHWaglu*. En effet, ce gène est co-

localisé à proximité immédiate des gènes d'une glucoamylase (définie comme spécifique des glucopolysaccharides) et d'un transporteur de maltose/maltooligosaccharides, deux protéines potentiellement exprimées simultanément. *Shewanella* sp. ANA-3 aurait donc évolué de sorte que SHWaglu conserve une large spécificité de substrat.

Le mécanisme d'hydrolyse par inversion de la configuration du carbone anomérique de SHWaglu déterminé par RMN soulève une interrogation quant à la classification IUBMB actuelle de l'enzyme (EC 3.2.1.20). En effet, les enzymes EC 3.2.1.20 libèrent leur produit avec une conformation α. Or, il a été établi que SHWaglu libère son produit en conformation β. Ainsi, SHWaglu aurait les caractéristiques de libération de produit d'une glucoamylase (EC 3.2.1.3).

Uniquement deux autres structures d'enzymes issues de la famille GH97 ont été résolues à ce jour (Gloster *et al.*, 2008; Kitamura *et al.*, 2008; Okuyama *et al.*, 2009). L'analyse de la structure de SHWaglu native résolue à 1,77 Å par cristallographie aux rayons X a permis d'approfondir les connaissances de cette famille et de mettre en évidence un motif 3D conservé au sein de la sous-famille des enzymes agissant selon un mécanisme d'inversion. Nous avons proposé l'importance de ce motif dans la stabilisation de la structure du domaine catalytique (voir « Un motif tridimensionnel dans la famille GH97 » p. 174). D'après un alignement multiple de séquences, ce motif semble dégénéré dans la sous-famille des enzymes agissant selon un mécanisme de rétention, ce qui permet d'avancer l'hypothèse que l'ancêtre commun serait une enzyme présentant un mécanisme d'hydrolyse avec inversion de la configuration du carbone anomérique.

Une analyse approfondie de la structure a permis de mettre en lumière certaines caractéristiques retrouvées généralement chez les enzymes psychrophiles. En effet, comparée à son homologue mésophile SusB, SHWaglu apparait plus flexible avec une énergie de stabilisation plus faible et un nombre de glycines proportionnellement plus important. De plus, elle révèle un excès de charges négatives à sa surface et possède une poche catalytique plus ouverte. Ces caractéristiques souvent retrouvées chez les psychrophiles permettent d'une part de conserver une certaine flexibilité aux basses températures pour compenser les vitesses réduites retrouvées à ces températures (Feller & Gerday, 2003; Siddiqui & Cavicchioli, 2006), mais aussi de réduire l'énergie d'activation requise pour la formation du complexe enzyme-substrat tout en facilitant la libération des produits formés (Struvay & Feller, 2012).

Toutes les structures tridimensionnelles de SHWaglu en complexe avec les différents ligands (acarbose, déoxynojirimycine, castanospermine, glucose et kojibiose) ont démontré la spécificité d'interaction du glucose au niveau du sous-site -1, ce qui est cohérent avec les études enzymatiques réalisées. De plus, la structure présentant une molécule de glucose dans chaque site actif provient d'un cristal trempé dans une solution contenant du nigérose. Ceci traduit donc le fait que l'enzyme est toujours active dans sa forme cristalline.

La superposition structurale de SHWaglu-ACA avec SHWaglu-KJB a démontré que les mêmes résidus interagissaient de façon semblable avec l'acarbose ou le kojibiose au niveau du sous-site +1, et cela malgré une liaison interglucosidique distincte. Ainsi, cette comparaison a mis en avant la conformation très particulière (de type ^4E) adoptée par le glucose dans le kojibiose. Cette conformation pourrait être l'état de transition du kojibiose durant l'hydrolyse, ce qui expliquerait la grande amélioration de l'efficacité catalytique de SHWaglu sur ce substrat.

Toutes ces études ont également souligné le caractère indispensable et essentiel de deux ions : Ca^{2+} et Cl^-, tous deux situés dans ou à proximité du site actif. Il a également été mis en évidence le caractère allostérique de l'enzyme en présence de ces deux ions. Ainsi, il serait intéressant d'approfondir cette analyse sur les propriétés ion-métallique dépendant de l'enzyme en déterminant le mécanisme d'interaction de ces ions avec l'enzyme.

La détermination des structures des mutants $SHWaglu_{E485Q}$ et $SHWaglu_{K434A}$ a permis de souligner l'importance toute particulière des résidus E485 et K434.

En effet, même s'il n'est pas un résidu catalytique à proprement parler, l'acide glutamique 485, de par sa coordination avec l'ion calcium, permet d'orienter correctement le résidu E509 qui, lui, est le résidu catalytique acide. Ceci explique donc son rôle crucial pour l'hydrolyse du substrat.

La lysine 434, quant à elle, est responsable à elle seule de l'obstruction du fond de la poche catalytique de SHWaglu. Sa mutation en alanine a permis d'obtenir un domaine catalytique avec un tonneau $(\beta/\alpha)_8$ quasi-traversant et la création d'éventuels nouveaux sous-sites -2 et -3. Il serait donc intéressant de réaliser à nouveau un criblage étendu des substrats et d'en déterminer les nouvelles constantes enzymatiques. Mais, plus encore, la création potentielle de ces deux nouveaux sous-sites permettrait à l'enzyme d'accepter des substrats ayant des degrés de polymérisation plus importants. Cette modification de spécificité de substrat pourrait donc notablement étendre le spectre des sucres reconnus et, par conséquent, celle des produits réactionnels. Cette mutation apparait d'emblée comme une piste applicative à explorer. Dans le même but et par extension, il faudrait envisager, sur la base des structures résolues dans cette étude, toute mutation qui autoriserait un changement de topologie de la poche catalytique.

- Perspectives générales

De futures études d'ingénierie protéique s'appuyant sur l'ensemble des données obtenues pourraient conduire à l'obtention d'enzymes optimisées, voire même à l'attribution de nouvelles activités et/ou de spécificités de substrat modifiées (voir ci-dessus). Ainsi, elles pourraient constituer un outil de choix pour la synthèse de glycosides à fort potentiel. En effet, l'emploi des glycosidases dans les procédés industriels devient de plus en plus courant. Ceci s'explique par leur utilisation selon des conditions « vertes », mais aussi par la facilité avec laquelle il est possible de produire ces enzymes en grande quantité et à moindre coût comparativement à la synthèse de sucres par voie chimique. L'une des stratégies possibles et prometteuses consiste en la création de glycosynthases. Ces dernières constituent une classe d'enzymes créée par l'introduction de mutations dirigées au niveau du résidu catalytique nucléophile. Elles autorisent ainsi la synthèse de nouvelles liaisons glycosidiques tout en empêchant l'hydrolyse des liaisons néoformées.

Ainsi, en mutant ce résidu nucléophile, généralement en alanine (ou glycine ou encore serine), il serait possible de transférer un sucre à partir d'un substrat donneur glycosyl-fluoré sur un saccharide accepteur d'intérêt (Mayer *et al.*, 2000; Honda & Kitaoka, 2006; Ohnuma *et al.*, 2012). Cette technique permettrait, par exemple, de libérer un produit valorisable, comme un sucre prébiotique. Jusqu'à récemment, seules les enzymes agissant selon un mécanisme de rétention et transformées en glycosynthases avaient été rapportées dans la littérature (Fairweather *et al.*, 2002; Okuyama *et al.*, 2002; Perugino *et al.*, 2003). Depuis lors, deux études ont montré qu'il était également possible de générer cette classe d'enzyme avec des enzymes « *inverting* » (Honda & Kitaoka, 2006; Ohnuma *et al.*, 2012). Avec ces découvertes récentes, il serait donc envisageable de synthétiser de nouveaux sucres dérivés d'intérêt prébiotique en s'appuyant sur les spécificités de substrats de SHWaglu.

Enfin, les informations obtenues sur SHWaglu et les relations structure-fonction-activité qui en ont découlé pourraient servir de base pour envisager des stratégies de bio-ingénierie pour les deux autres α-glucosidases étudiées durant cette thèse. Ainsi, il serait possible d'envisager une application agro-alimentaire. En modifiant ou en optimisant leurs activités enzymatiques par bio-ingénierie, les enzymes d'intérêt pourraient être, par exemple, utilisées dans le cadre de la production de produits laitiers à haute valeur ajoutée nécessitant des conditions particulières et exigentes lors de la fabrication.

Références bibliographiques

[A]

Abbott, D.W., Eirin-Lopez, J.M. and Boraston, A.B. (2008) Insight into ligand diversity and novel biological roles for family 32 carbohydrate-binding modules. *Mol Biol Evol*, **25**, 155-167.

Aghajari, N., Feller, G., Gerday, C. and Haser, R. (1998) Structures of the psychrophilic *Alteromonas haloplanctis* alpha-amylase give insights into cold adaptation at a molecular level. *Structure*, **6**, 1503-1516.

Aghajari, N., Feller, G., Gerday, C. and Haser, R. (2002) Structural basis of alpha-amylase activation by chloride. *Protein Sci*, **11**, 1435-1441.

Aghajari, N., Van Petegem, F., Villeret, V., Chessa, J.P., Gerday, C., Haser, R. and Van Beeumen, J. (2003) Crystal structures of a psychrophilic metalloprotease reveal new insights into catalysis by cold-adapted proteases. *Proteins*, **50**, 636-647.

Alarico, S., da Costa, M.S. and Empadinhas, N. (2008) Molecular and physiological role of the trehalose-hydrolyzing alpha-glucosidase from *Thermus thermophilus* HB27. *J Bacteriol*, **190**, 2298-2305.

Aleshin, A.E., Feng, P.H., Honzatko, R.B. and Reilly, P.J. (2003) Crystal structure and evolution of a prokaryotic glucoamylase. *J Mol Biol*, **327**, 61-73.

Almeida, C.C., Lorena, S.L., Pavan, C.R., Akasaka, H.M. and Mesquita, M.A. (2012) Beneficial effects of long-term consumption of a probiotic combination of *Lactobacillus casei Shirota* and *Bifidobacterium breve* Yakult may persist after suspension of therapy in lactose-intolerant patients. *Nutr Clin Pract*, **27**, 247-251.

Andersen, J.M., Barrangou, R., Hachem, M.A., Lahtinen, S.J., Goh, Y.J., Svensson, B. and Klaenhammer, T.R. (2012) Transcriptional analysis of prebiotic uptake and catabolism by *Lactobacillus acidophilus* NCFM. *PLoS One*, **7**, e44409.

Angelov, A., Putyrski, M. and Liebl, W. (2006) Molecular and biochemical characterization of alpha-glucosidase and alpha-mannosidase and their clustered genes from the thermoacidophilic archaeon *Picrophilus torridus*. *J Bacteriol*, **188**, 7123-7131.

[B]

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. and Gordon, J.I. (2005) Host-bacterial mutualism in the human intestine. *Science*, **307**, 1915-1920.

Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D. and Waley, S.G. (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature*, **255**, 609-614.

Barnes, J.L., Hartmann, B., Holst, J.J. and Tappenden, K.A. (2012) Intestinal adaptation is stimulated by partial enteral nutrition supplemented with the prebiotic short-chain fructooligosaccharide in a neonatal intestinal failure piglet model. *J Parenter Enteral Nutr*, **36**, 524-537.

Bassaganya-Riera, J., Viladomiu, M., Pedragosa, M., De Simone, C. and Hontecillas, R. (2012) Immunoregulatory mechanisms underlying prevention of colitis-associated colorectal cancer by probiotic bacteria. *PLoS One*, **7**, e34676.

Bercik, P., Verdu, E.F., Foster, J.A., Macri, J., Potter, M., Huang, X., Malinowski, P., Jackson, W., Blennerhassett, P., Neufeld, K.A., Lu, J., Khan, W.I., Corthesy-Theulaz, I., Cherbut,

C., Bergonzelli, G.E. and Collins, S.M. (2010) Chronic gastrointestinal inflammation induces anxiety-like behavior and alters central nervous system biochemistry in mice. *Gastroenterology*, **139**, 2102-2112 e2101.

Blanchard, S., Cottaz, S., Coutinho, P.M., Patkar, S., Vind, J., Boer, H., Koivula, A., Driguez, H. and Armand, S. (2007) Mutation of fungal endoglucanases into glycosynthases and characterization of their acceptor substrate specificity. *Journal of Molecular Catalysis B: Enzymatic*, **44**, 106-116.

Bocci, V. (1992) The neglected organ: bacterial flora has a crucial immunostimulatory role. *Perspect Biol Med*, **35**, 251-260.

Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001) The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res*, **11**, 731-753.

Boraston, A.B., Bolam, D.N., Gilbert, H.J. and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*, **382**, 769-781.

Boraston, A.B., Ficko-Blean, E. and Healey, M. (2007) Carbohydrate recognition by a large sialidase toxin from *Clostridium perfringens*. *Biochemistry*, **46**, 11352-11360.

Boraston, A.B., McLean, B.W., Kormos, J.M., Alam, M., Gilkes, N.R., Haynes, C.A., Tomme, P., Kilburn, D.G. and Warren, R.A.J. (1999) *Recent Advances in Carbohydrate Bioengineering*, Cambridge.

Borges de Melo, E., da Silveira Gomes, A. and Carvalho, I. (2006) alpha-and beta-Glucosidase inhibitors: chemical structure and biological activity. *Tetrahedron*, **62**, 10277-10302.

Brady, L.J., Gallaher, D.D. and Busta, F.F. (2000) The role of probiotic cultures in the prevention of colon cancer. *J Nutr*, **130**, 410S-414S.

Bravo, J.A., Forsythe, P., Chew, M.V., Escaravage, E., Savignac, H.M., Dinan, T.G., Bienenstock, J. and Cryan, J.F. (2011) Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc Natl Acad Sci U S A*, **108**, 16050-16055.

Bravo-Torres, J.C., Villagomez-Castro, J.C., Calvo-Mendez, C., Flores-Carreon, A. and Lopez-Romero, E. (2004) Purification and biochemical characterisation of a membrane-bound alpha-glucosidase from the parasite *Entamoeba histolytica*. *Int J Parasitol*, **34**, 455-462.

Breyer, W.A. and Matthews, B.W. (2001) A structural basis for processivity. *Protein Sci*, **10**, 1699-1711.

Brizova, K., Kralova, B., Demnerova, K. and Vins, I. (1992) Isolation and characterization of alpha-glucosidase from *Aspergillus niger*. *J Chromatogr*, **593**, 125-131.

Brunger, A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472-475.

Burmeister, W.P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. and Henrissat, B. (1997) The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure*, **5**, 663-675.

Burmeister, W.P., Cottaz, S., Rollin, P., Vasella, A. and Henrissat, B. (2000) High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J Biol Chem*, **275**, 39385-39393.

Burstein, C. and Kepes, A. (1971) The alpha-galactosidase from *Escherichia coli* K12. *Biochim Biophys Acta*, **230**, 52-63.

[C]

Candela, M., Perna, F., Carnevali, P., Vitali, B., Ciati, R., Gionchetti, P., Rizzello, F., Campieri, M. and Brigidi, P. (2008) Interaction of probiotic *Lactobacillus* and *Bifidobacterium* strains with human intestinal epithelial cells: adhesion properties, competition against enteropathogens and modulation of IL-8 production. *Int J Food Microbiol*, **125**, 286-292.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*, **37**, D233-238.

Cavicchioli, R., Siddiqui, K.S., Andrews, D. and Sowers, K.R. (2002) Low-temperature extremophiles and their applications. *Curr Opin Biotechnol*, **13**, 253-261.

Chapman, C.M., Gibson, G.R. and Rowland, I. (2011) Health benefits of probiotics: are mixtures more effective than single strains? *Eur J Nutr*, **50**, 1-17.

Chapman, C.M., Gibson, G.R. and Rowland, I. (2012) In vitro evaluation of single- and multi-strain probiotics: Inter-species inhibition between probiotic strains, and inhibition of pathogens. *Anaerobe*, **18**, 405-413.

Christensen, H.R., Frokiaer, H. and Pestka, J.J. (2002) *Lactobacilli* differentially modulate expression of cytokines and maturation surface markers in murine dendritic cells. *J Immunol*, **168**, 171-178.

Christiansen, C., Abou Hachem, M., Janecek, S., Vikso-Nielsen, A., Blennow, A. and Svensson, B. (2009) The carbohydrate-binding module family 20--diversity, structure, and function. *FEBS J*, **276**, 5006-5029.

Clarke, G., Grenham, S., Scully, P., Fitzgerald, P., Moloney, R.D., Shanahan, F., Dinan, T.G. and Cryan, J.F. (2012) The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner. *Mol Psychiatry*.

Claros, M.G. and von Heijne, G. (1994) TopPred II: an improved software for membrane protein structure predictions. *Comput Appl Biosci*, **10**, 685-686.

Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **50**, 760-763.

Comfort, D.A., Chou, C.J., Conners, S.B., VanFossen, A.L. and Kelly, R.M. (2008) Functional-genomics-based identification and characterization of open reading frames encoding alpha-glucoside-processing enzymes in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Appl Environ Microbiol*, **74**, 1281-1283.

Costantino, H.R., Brown, S.H. and Kelly, R.M. (1990) Purification and characterization of an alpha-glucosidase from a hyperthermophilic archaebacterium, *Pyrococcus furiosus*, exhibiting a temperature optimum of 105 to 115 degrees C. *J Bacteriol*, **172**, 3654-3660.

Costello, E.K., Stagaman, K., Dethlefsen, L., Bohannan, B.J. and Relman, D.A. (2012) The application of ecological theory toward an understanding of the human microbiome. *Science*, **336**, 1255-1262.

Coutinho, P.M., Deleury, E., Davies, G.J. and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol*, **328**, 307-317.

Creagh, A.L., Ong, E., Jervis, E., Kilburn, D.G. and Haynes, C.A. (1996) Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. *Proc Natl Acad Sci U S A*, **93**, 12229-12234.

Cross, M.L., Mortensen, R.R., Kudsk, J. and Gill, H.S. (2002) Dietary intake of *Lactobacillus rhamnosus* HNOO1 enhances production of both Th1 and Th2 cytokines in antigen-primed mice. *Med Microbiol Immunol*, **191**, 49-53.

Cryan, J.F. and O'Mahony, S.M. (2011) The microbiome-gut-brain axis: from bowel to behavior. *Neurogastroenterol Motil*, **23**, 187-192.

Cserzo, M., Wallin, E., Simon, I., von Heijne, G. and Elofsson, A. (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng*, **10**, 673-676.

[D]

Daigo, K., Inamori, Y. and Takemoto, T. (1986) Studies on the Constituents of the Water Extract of the Root of Mulberry Tree (*Morus bombycis* KOIDZ). *Chemical & Pharmaceutical Bulletin*, **34**, 2243-2246.

Dairaku, K. and Spiro, R.G. (1997) Phylogenetic survey of endomannosidase indicates late evolutionary appearance of this N-linked oligosaccharide processing enzyme. *Glycobiology*, **7**, 579-586.

Davies, G. and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure*, **3**, 853-859.

Davies, G.J., Wilson, K.S. and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*, **321 (Pt 2)**, 557-559.

de Vos, W.M. (1996) Metabolic engineering of sugar catabolism in lactic acid bacteria. *Antonie Van Leeuwenhoek*, **70**, 223-242.

de Vrese, M., Stegelmann, A., Richter, B., Fenselau, S., Laue, C. and Schrezenmeir, J. (2001) Probiotics--compensation for lactase insufficiency. *Am J Clin Nutr*, **73**, 421S-429S.

Dhanawansa, R., Faridmoayer, A., van der Merwe, G., Li, Y.X. and Scaman, C.H. (2002) Overexpression, purification, and partial characterization of *Saccharomyces cerevisiae* processing alpha glucosidase I. *Glycobiology*, **12**, 229-234.

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Bjorkholm, B., Samuelsson, A., Hibberd, M.L., Forssberg, H. and Pettersson, S. (2011) Normal gut microbiota modulates brain development and behavior. *Proc Natl Acad Sci U S A*, **108**, 3047-3052.

Dobson, A., Cotter, P.D., Ross, R.P. and Hill, C. (2012) Bacteriocin production: a probiotic trait? *Appl Environ Microbiol*, **78**, 1-6.

Dubey, A.P., Rajeshwari, K., Chakravarty, A. and Famularo, G. (2008) Use of VSL[sharp]3 in the treatment of rotavirus diarrhea in children: preliminary results. *J Clin Gastroenterol*, **42 Suppl 3 Pt 1**, S126-129.

[E]

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486-501.

Enomoto, M., Noguchi, S., Hattori, M., Sugiyama, H., Suzuki, Y., Hanaoka, A., Okada, S. and Yoshida, T. (2009) Oral administration of *Lactobacillus plantarum* NRIC0380 suppresses IgE production and induces CD4(+)CD25(+)Foxp3(+) cells in vivo. *Biosci Biotechnol Biochem*, **73**, 457-460.

Ernst, H.A., Lo Leggio, L., Willemoes, M., Leonard, G., Blum, P. and Larsen, S. (2006) Structure of the *Sulfolobus solfataricus* alpha-glucosidase: implications for domain conservation and substrate recognition in GH31. *J Mol Biol*, **358**, 1106-1124.

Fairweather, J.K., Faijes, M., Driguez, H. and Planas, A. (2002) Specificity studies of *Bacillus* 1,3-1,4-beta-glucanases and application to glycosynthase-catalyzed transglycosylation. *Chembiochem*, **3**, 866-873.

FAO/WHO. (2002) Working group report on drafting guidelines for the evaluation of probiotics in food. FAO/WHO, London, Ontario, Canada.

[F]

Fell, J.M. (2005) Neonatal inflammatory intestinal diseases: necrotising enterocolitis and allergic colitis. *Early Hum Dev*, **81**, 117-122.

Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: hot topics in cold adaptation. *Nat Rev Microbiol*, **1**, 200-208.

Feller, G., Narinx, E., Arpigny, J.L., Aittaleb, M., Baise, E., Genicot, S. and Gerday, C. (1996) Enzymes from psychrophilic organisms. *FEMS Microbiology Reviews*, **18**, 189-202.

Feller, G., Payan, F., Theys, F., Qian, M., Haser, R. and Gerday, C. (1994) Stability and structural analysis of alpha-amylase from the antarctic psychrophile *Alteromonas haloplanctis* A23. *Eur J Biochem*, **222**, 441-447.

Feng, H.Y., Drone, J., Hoffmann, L., Tran, V., Tellier, C., Rabiller, C. and Dion, M. (2005) Converting a beta-glycosidase into a beta-transglycosidase by directed evolution. *J Biol Chem*, **280**, 37088-37097.

Ferrer, M., Golyshina, O.V., Plou, F.J., Timmis, K.N. and Golyshin, P.N. (2005) A novel alpha-glucosidase from the acidophilic archaeon *Ferroplasma acidiphilum* strain Y with high transglycosylation activity and an unusual catalytic nucleophile. *Biochem J*, **391**, 269-276.

Ficko-Blean, E. and Boraston, A.B. (2006) The interaction of a carbohydrate-binding module from a *Clostridium perfringens* N-acetyl-beta-hexosaminidase with its carbohydrate receptor. *J Biol Chem*, **281**, 37748-37757.

Ficko-Blean, E., Gregg, K.J., Adams, J.J., Hehemann, J.H., Czjzek, M., Smith, S.P. and Boraston, A.B. (2009) Portrait of an enzyme, a complete structural analysis of a multimodular beta-N-acetylglucosaminidase from *Clostridium perfringens*. *J Biol Chem*, **284**, 9876-9884.

Ficko-Blean, E., Stuart, C.P., Suits, M.D., Cid, M., Tessier, M., Woods, R.J. and Boraston, A.B. (2012) Carbohydrate recognition by an architecturally complex alpha-N-acetylglucosaminidase from *Clostridium perfringens*. *PLoS One*, **7**, e33524.

Fort, J., de la Ballina, L.R., Burghardt, H.E., Ferrer-Costa, C., Turnay, J., Ferrer-Orta, C., Uson, I., Zorzano, A., Fernandez-Recio, J., Orozco, M., Lizarbe, M.A., Fita, I. and Palacin, M. (2007) The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane. *J Biol Chem*, **282**, 31444-31452.

Frandsen, T.P. and Svensson, B. (1998) Plant alpha-glucosidases of the glycoside hydrolase family 31. Molecular properties, substrate specificity, reaction mechanism, and comparison with family members of different origin. *Plant Mol Biol*, **37**, 1-13.

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., Tobe, T., Clarke, J.M., Topping, D.L., Suzuki, T., Taylor, T.D., Itoh, K., Kikuchi, J., Morita, H., Hattori, M. and

Ohno, H. (2011) Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature*, **469**, 543-547.
Fuller, R. (1989) Probiotics in man and animals. *J Appl Bacteriol*, **66**, 365-378.

[G]

Gabrisko, M. and Janecek, S. (2009) Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 alpha-amylase family. *FEBS J*, **276**, 7265-7278.
Galdeano, C.M. and Perdigon, G. (2004) Role of viability of probiotic strains in their persistence in the gut and in mucosal immune stimulation. *J Appl Microbiol*, **97**, 673-681.
Gareau, M.G., Sherman, P.M. and Walker, W.A. (2010) Probiotics and the gut microbiota in intestinal health and disease. *Nat Rev Gastroenterol Hepatol*, **7**, 503-514.
Garman, E. (1999) Cool data: quantity AND quality. *Acta Crystallogr D Biol Crystallogr*, **55**, 1641-1653.
Garman, E. and Schneider, T.R. (1997) Macromolecular cryocrystallography. *J. Appl. Cryst.*, **30**, 211-237.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In Walker, J.M. (ed.), *The Proteomics Protocols Handbook*. Humana Press, Totowa, pp. 571-607.
Gebler, J., Gilkes, N.R., Claeyssens, M., Wilson, D.B., Beguin, P., Wakarchuk, W.W., Kilburn, D.G., Miller, R.C., Jr., Warren, R.A. and Withers, S.G. (1992) Stereoselective hydrolysis catalyzed by related beta-1,4-glucanases and beta-1,4-xylanases. *J Biol Chem*, **267**, 12559-12561.
Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., Marx, J.C., Sonan, G., Feller, G. and Gerday, C. (2004) Some like it cold: biocatalysis at low temperatures. *FEMS Microbiol Rev*, **28**, 25-42.
Gerritsen, J., Smidt, H., Rijkers, G.T. and de Vos, W.M. (2011) Intestinal microbiota in human health and disease: the impact of probiotics. *Genes Nutr*, **6**, 209-240.
Giacopuzzi, E., Bresciani, R., Schauer, R., Monti, E. and Borsani, G. (2012) New insights on the sialidase protein family revealed by a phylogenetic analysis in metazoa. *PLoS One*, **7**, e44193.
Giannesi, G.C., Polizeli, M.d.L., Terenzi, H.F. and Jorge, J.A. (2006) A novel alpha-glucosidase from *Chaetomium thermophilum* var. *coprophilum* that converts maltose into trehalose: Purification and partial characterisation of the enzyme. *Process Biochem.*, **41**, 1729-1735.
Gibson, G.R., Probert, H.M., Loo, J.V., Rastall, R.A. and Roberfroid, M.B. (2004) Dietary modulation of the human colonic microbiota: updating the concept of prebiotics. *Nutr Res Rev*, **17**, 259-275.
Gibson, G.R. and Roberfroid, M.B. (1995) Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *J Nutr*, **125**, 1401-1412.
Gilkes, N.R., Warren, R.A., Miller, R.C., Jr. and Kilburn, D.G. (1988) Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *J Biol Chem*, **263**, 10401-10407.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M. and Nelson, K.E. (2006) Metagenomic analysis of the human distal gut microbiome. *Science*, **312**, 1355-1359.

Gilliland, S.E., Nelson, C.R. and Maxwell, C. (1985) Assimilation of cholesterol by *Lactobacillus acidophilus*. *Appl Environ Microbiol*, **49**, 377-381.

Giuliano, M., Schiraldi, C., Marotta, M.R., Hugenholtz, J. and De Rosa, M. (2004) Expression of *Sulfolobus solfataricus* alpha-glucosidase in *Lactococcus lactis*. *Appl Microbiol Biotechnol*, **64**, 829-832.

Gloster, T.M. and Davies, G.J. (2010) Glycosidase inhibition: assessing mimicry of the transition state. *Org Biomol Chem*, **8**, 305-320.

Gloster, T.M., Turkenburg, J.P., Potts, J.R., Henrissat, B. and Davies, G.J. (2008) Divergence of catalytic mechanism within a glycosidase family provides insight into evolution of carbohydrate metabolism by human gut flora. *Chem Biol*, **15**, 1058-1067.

Goehler, L.E., Gaykema, R.P., Opitz, N., Reddaway, R., Badr, N. and Lyte, M. (2005) Activation in vagal afferents and central autonomic pathways: early responses to intestinal infection with *Campylobacter jejuni*. *Brain Behav Immun*, **19**, 334-344.

Goldin, B.R. (1998) Health benefits of probiotics. *Br J Nutr*, **80**, S203-207.

Goldstein, R.A. (2007) Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: insights from the quasi-chemical approximation. *Protein Sci*, **16**, 1887-1895.

Gros, D.F., Antony, M.M., McCabe, R.E. and Swinson, R.P. (2009) Frequency and severity of the symptoms of irritable bowel syndrome across the anxiety disorders and depression. *J Anxiety Disord*, **23**, 290-296.

Gustaw, W., Kordowska-Wiater, M. and Koziol, J. (2011) The influence of selected prebiotics on the growth of lactic acid bacteria for bio-yoghurt production. *Acta Sci Pol Technol Aliment*, **10**, 455-466.

[H]

Haberer, P., du Toit, M., Dicks, L.M., Ahrens, F. and Holzapfel, W.H. (2003) Effect of potentially probiotic lactobacilli on faecal enzyme activity in minipigs on a high-fat, high-cholesterol diet-a preliminary in vivo trial. *Int J Food Microbiol*, **87**, 287-291.

Hancock, S.M., Vaughan, M.D. and Withers, S.G. (2006) Engineering of glycosidases and glycosyltransferases. *Curr Opin Chem Biol*, **10**, 509-519.

Hanes, C.S. (1932) Studies on plant amylases: The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. *Biochem J*, **26**, 1406-1421.

Harker, D. (1956) The determination of the phases of the structure factors of non-centrosymmetric crystals by the method of double isomorphous replacement. *Acta Cryst.*, **9**, 1-9.

Hata, Y., Yamamoto, M., Ohni, M., Nakajima, K., Nakamura, Y. and Takano, T. (1996) A placebo-controlled study of the effect of sour milk on blood pressure in hypertensive subjects. *Am J Clin Nutr*, **64**, 767-771.

Hau, H.H. and Gralnick, J.A. (2007) Ecology and biotechnology of the genus *Shewanella*. *Annu Rev Microbiol*, **61**, 237-258.

Heller, K.J. (2001) Probiotic bacteria in fermented foods: product characteristics and starter organisms. *Am J Clin Nutr*, **73**, 374S-379S.

Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, **280 (Pt 2)**, 309-316.

Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, **293 (Pt 3)**, 781-788.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem J*, **316 (Pt 2)**, 695-696.

Heras, B., Edeling, M.A., Byriel, K.A., Jones, A., Raina, S. and Martin, J.L. (2003) Dehydration converts DsbG crystal diffraction from low to high resolution. *Structure*, **11**, 139-145.

Hermans, M.M., Kroos, M.A., van Beeumen, J., Oostra, B.A. and Reuser, A.J. (1991) Human lysosomal alpha-glucosidase. Characterization of the catalytic site. *J Biol Chem*, **266**, 13507-13512.

Hernandez-Ledesma, B., Miguel, M., Amigo, L., Aleixandre, M.A. and Recio, I. (2007) Effect of simulated gastrointestinal digestion on the antihypertensive properties of synthetic beta-lactoglobulin peptide sequences. *J Dairy Res*, **74**, 336-339.

Herscovics, A. (1999) Processing glycosidases of *Saccharomyces cerevisiae*. *Biochim Biophys Acta*, **1426**, 275-285.

Hickson, M., D'Souza, A.L., Muthu, N., Rogers, T.R., Want, S., Rajkumar, C. and Bulpitt, C.J. (2007) Use of probiotic *Lactobacillus* preparation to prevent diarrhoea associated with antibiotics: randomised double blind placebo controlled trial. *BMJ*, **335**, 80.

Hofle, M.G. and Brettar, I. (1996) Genotyping of heterotrophic bacteria from the central baltic sea by use of low-molecular-weight RNA profiles. *Appl Environ Microbiol*, **62**, 1383-1390.

Hofmann, K. and Stoffel, W. (1993) TMbase - A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler*, **374**, 166.

Hofstee, B.H. (1959) Non-inverted versus inverted plots in enzyme kinetics. *Nature*, **184**, 1296-1298.

Honda, Y. and Kitaoka, M. (2006) The first glycosynthase derived from an inverting glycoside hydrolase. *J Biol Chem*, **281**, 1426-1431.

Hondoh, H., Saburi, W., Mori, H., Okuyama, M., Nakada, T., Matsuura, Y. and Kimura, A. (2008) Substrate recognition mechanism of alpha-1,6-glucosidic linkage hydrolyzing enzyme, dextran glucosidase from *Streptococcus mutans*. *J Mol Biol*, **378**, 913-922.

Hooft, R.W., Vriend, G., Sander, C. and Abola, E.E. (1996) Errors in protein structures. *Nature*, **381**, 272.

Hooper, L.V. and Gordon, J.I. (2001) Commensal host-bacterial relationships in the gut. *Science*, **292**, 1115-1118.

Hooper, L.V., Littman, D.R. and Macpherson, A.J. (2012) Interactions between the microbiota and the immune system. *Science*, **336**, 1268-1273.

Hooper, L.V., Wong, M.H., Thelin, A., Hansson, L., Falk, P.G. and Gordon, J.I. (2001) Molecular analysis of commensal host-microbial relationships in the intestine. *Science*, **291**, 881-884.

Human Microbiome Project, C. (2012) Structure, function and diversity of the healthy human microbiome. *Nature*, **486**, 207-214.

Hummel, S., Veltman, K., Cichon, C., Sonnenborn, U. and Schmidt, M.A. (2012) Differential targeting of the E-Cadherin/beta-Catenin complex by gram-positive probiotic lactobacilli improves epithelial barrier function. *Appl Environ Microbiol*, **78**, 1140-1147.

Hung, V.S., Hatada, Y., Goda, S., Lu, J., Hidaka, Y., Li, Z., Akita, M., Ohta, Y., Watanabe, K., Matsui, H., Ito, S. and Horikoshi, K. (2005) alpha-Glucosidase from a strain of deep-sea *Geobacillus*: a potential enzyme for the biosynthesis of complex carbohydrates. *Appl Microbiol Biotechnol*, **68**, 757-765.

[I]

Iannitti, T. and Palmieri, B. (2010) Therapeutical use of probiotic formulations in clinical practice. *Clin Nutr*, **29**, 701-725.

Inouye, S., Tsuruoka, T., Ito, T. and Niida, T. (1968) Structure and synthesis of nojirimycin. *Tetrahedron*, **24**, 2125-2144.

Inouye, S., Tsuruoka, T. and Nida, T. (1966) The structure of nojirimycin, a piperidinose sugar antibiotic. *J Antibiot (Tokyo)*, **19**, 288-292.

Isoda, Y., Shimizu, Y., Hashimoto, A., Fujiwara, H., Nitta, Y. and Kagemoto, A. (1992) Mechanism of hydrolyses of phenyl alpha-maltosides catalyzed by taka-amylase A. *J Biochem*, **111**, 204-209.

Iwanami, S., Matsui, H., Kimura, A., Ito, H., Mori, H., Honma, M. and Chiba, S. (1995) Chemical modification and amino acid sequence of active site in sugar beet alpha-glucosidase. *Biosci Biotechnol Biochem*, **59**, 459-463.

Iyer, P.V. and Ananthanarayan, L. (2008) Enzyme stability and stabilization - Aqueous and non-aqueous environment. *Process Biochemistry*, **43**, 1019–1032.

[J]

Janecek, S. and Sevcik, J. (1999) The evolution of starch-binding domain. *FEBS Lett*, **456**, 119-125.

Janecek, S., Svensson, B. and Henrissat, B. (1997) Domain evolution in the alpha-amylase family. *J Mol Evol*, **45**, 322-331.

Jernberg, C., Lofmark, S., Edlund, C. and Jansson, J.K. (2011) Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology*, **156**, 3216-3223.

Jie, Z., Bang-Yao, L., Ming-Jie, X., Hai-Wei, L., Zu-Kang, Z., Ting-Song, W. and Craig, S.A. (2000) Studies on the effects of polydextrose intake on physiologic functions in Chinese people. *Am J Clin Nutr*, **72**, 1503-1509.

Jiménez, E., Marín, M.L., Martín, R., Odriozola, J.M., Olivares, M., Xaus, J., Fernández, L. and Rodríguez, J.M. (2008) Is meconium from healthy newborns actually sterile? *Research in Microbiology*, **159**, 187-193.

Johnson-Henry, K.C., Donato, K.A., Shen-Tu, G., Gordanpour, M. and Sherman, P.M. (2008) *Lactobacillus rhamnosus* strain GG prevents enterohemorrhagic *Escherichia coli* O157:H7-induced changes in epithelial barrier function. *Infect Immun*, **76**, 1340-1348.

Johnston, B.C., Goldenberg, J.Z., Vandvik, P.O., Sun, X. and Guyatt, G.H. (2011) Probiotics for the prevention of pediatric antibiotic-associated diarrhea. *Cochrane Database Syst Rev*, CD004827.

Junge, B., Heiker, F.R., Kurz, J., Müller, L., Schmidt, D.D. and Wünsche, C. (1984) Untersuchungen zur struktur des alpha-D-glucosidase inhibitors acarbose. *Carbohydrate Research*, **128**, 235-268.

[K]

Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, **26**, 795-800.

Kabsch, W. (2010) XDS. *Acta Crystallogr D Biol Crystallogr*, **66**, 125-132.

Kang, M.S., Okuyama, M., Mori, H. and Kimura, A. (2009) The first alpha-1,3-glucosidase from bacterial origin belonging to glycoside hydrolase family 31. *Biochimie*, **91**, 1434-1442.

Karpinets, T.V., Obraztsova, A.Y., Wang, Y., Schmoyer, D.D., Kora, G.H., Park, B.H., Serres, M.H., Romine, M.F., Land, M.L., Kothe, T.B., Fredrickson, J.K., Nealson, K.H. and Uberbacher, E.C. (2010) Conserved synteny at the protein family level reveals genes underlying *Shewanella* species cold tolerance and predicts their novel phenotypes. *Funct Integr Genomics*, **10**, 97-110.

Kashiwabara, S., Azuma, S., Tsuduki, M. and Suzuki, Y. (2000) The primary structure of the subunit in *Bacillus thermoamyloliquefaciens* KP1071 molecular weight 540,000 homohexameric alpha-glucosidase II belonging to the glycosyl hydrolase family 31. *Biosci Biotechnol Biochem*, **64**, 1379-1393.

Kato, N., Suyama, S., Shirokane, M., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2002) Novel alpha-glucosidase from *Aspergillus nidulans* with strong transglycosylation activity. *Appl Environ Microbiol*, **68**, 1250-1256.

Kawase, M., Hashimoto, H., Hosoda, M., Morita, H. and Hosono, A. (2000) Effect of administration of fermented milk containing whey protein concentrate to rats and healthy men on serum lipids and blood pressure. *J Dairy Sci*, **83**, 255-263.

Kim, Y.W., Lovering, A.L., Chen, H., Kantner, T., McIntosh, L.P., Strynadka, N.C. and Withers, S.G. (2006) Expanding the thioglycoligase strategy to the synthesis of alpha-linked thioglycosides allows structural investigation of the parent enzyme/substrate complex. *J Am Chem Soc*, **128**, 2202-2203.

Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N., Kimura, A., Tanaka, I. and Yao, M. (2008) Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*. *J Biol Chem*, **283**, 36328-36337.

Kondepudi, K.K., Ambalam, P., Nilsson, I., Wadstrom, T. and Ljungh, A. (2012) Prebiotic-non-digestible oligosaccharides preference of probiotic bifidobacteria and antimicrobial activity against *Clostridium difficile*. *Anaerobe*, **18**, 489-497.

Koshland, D.E. (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews*, **28**, 416-436.

Kren, V. and Martinkova, L. (2001) Glycosides in medicine: "The role of glycosidic residue in biological activity". *Curr Med Chem*, **8**, 1303-1328.

Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-797.

Kumar, S., Dwevedi, A. and Kayastha, A.M. (2009) Immobilization of soybean (Glycine max) urease on alginate and chitosan beads showing improved stability: Analytical applications. *Journal of Molecular Catalysis B: Enzymatic*, **58**, 138-145.

Kurakata, Y., Uechi, A., Yoshida, H., Kamitori, S., Sakano, Y., Nishikawa, A. and Tonozuka, T. (2008) Structural insights into the substrate specificity and function of *Escherichia coli*

K12 YgjK, a glucosidase belonging to the glycoside hydrolase family 63. *J Mol Biol*, **381**, 116-128.

Kurokawa, K., Itoh, T., Kuwahara, T., Oshima, K., Toh, H., Toyoda, A., Takami, H., Morita, H., Sharma, V.K., Srivastava, T.P., Taylor, T.D., Noguchi, H., Mori, H., Ogura, Y., Ehrlich, D.S., Itoh, K., Takagi, T., Sakaki, Y., Hayashi, T. and Hattori, M. (2007) Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res*, **14**, 169-181.

[L]

Lariviere, L., Gueguen-Chaignon, V. and Morera, S. (2003) Crystal structures of the T4 phage beta-glucosyltransferase and the D100A mutant in complex with UDP-glucose: glucose binding and identification of the catalytic base for a direct displacement mechanism. *J Mol Biol*, **330**, 1077-1086.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, **26**, 283-291.

Lathrop, S.K., Bloom, S.M., Rao, S.M., Nutsch, K., Lio, C.W., Santacruz, N., Peterson, D.A., Stappenbeck, T.S. and Hsieh, C.S. (2011) Peripheral education of the immune system by colonic commensal microbiota. *Nature*, **478**, 250-254.

Lee, S.S., He, S. and Withers, S.G. (2001) Identification of the catalytic nucleophile of the Family 31 alpha-glucosidase from *Aspergillus niger* via trapping of a 5-fluoroglycosyl-enzyme intermediate. *Biochem J*, **359**, 381-386.

Lee, Y.K. and Mazmanian, S.K. (2010) Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science*, **330**, 1768-1773.

Le-Tien, C., Millette, M., Lacroix, M. and Mateescu, M.A. (2004) Modified alginate matrices for the immobilization of bioactive agents. *Biotechnol Appl Biochem*, **39**, 189-198.

Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R. and Gordon, J.I. (2008) Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol*, **6**, 776-788.

Li, K.B. and Chan, K.Y. (1983) Production and properties of alpha-glucosidase from Lactobacillus acidophilus. *Appl Environ Microbiol*, **46**, 1380-1387.

Lilly, D.M. and Stillwell, R.H. (1965) Probiotics: Growth-promoting factors produced by microorganisms. *Science*, **147**, 747-748.

Lineweaver, H. and Burk, D. (1934) The Determination of Enzyme Dissociation Constants. *J. Am. Chem. Soc*, **56**, 658-666.

Lipski, A., Watzlawick, H., Ravaud, S., Robert, X., Rhimi, M., Haser, R., Mattes, R. and Aghajari, N. (2013) Mutations inducing an active-site aperture in *Rhizobium* sp. sucrose isomerase confer hydrolytic activity. *Acta Crystallogr D Biol Crystallogr*, **69**, 298-307.

Lodge, J.A., Maier, T., Liebl, W., Hoffmann, V. and Strater, N. (2003) Crystal structure of *Thermotoga maritima* alpha-glucosidase AglA defines a new clan of NAD^+-dependent glycosidases. *J Biol Chem*, **278**, 19151-19158.

Long, F., Vagin, A.A., Young, P. and Murshudov, G.N. (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr*, **64**, 125-132.

Lovering, A.L., Lee, S.S., Kim, Y.W., Withers, S.G. and Strynadka, N.C. (2005) Mechanistic and structural analysis of a family 31 alpha-glycosidase and its glycosyl-enzyme intermediate. *J Biol Chem*, **280**, 2105-2115.

Lyte, M., Li, W., Opitz, N., Gaykema, R.P. and Goehler, L.E. (2006) Induction of anxiety-like behavior in mice during the initial stages of infection with the agent of murine colonic hyperplasia *Citrobacter rodentium*. *Physiol Behav*, **89**, 350-357.

[M]

Macfarlane, G.T., Steed, H. and Macfarlane, S. (2008) Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *J Appl Microbiol*, **104**, 305-344.

MacGregor, E.A., Janecek, S. and Svensson, B. (2001) Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. *Biochim Biophys Acta*, **1546**, 1-20.

Machovic, M. and Janecek, S. (2006) Starch-binding domains in the post-genome era. *Cell Mol Life Sci*, **63**, 2710-2724.

Makelainen, H., Hasselwander, O., Rautonen, N. and Ouwehand, A.C. (2009) Panose, a new prebiotic candidate. *Lett Appl Microbiol*, **49**, 666-672.

Mark, B.L., Vocadlo, D.J., Knapp, S., Triggs-Raine, B.L., Withers, S.G. and James, M.N. (2001) Crystallographic evidence for substrate-assisted catalysis in a bacterial beta-hexosaminidase. *J Biol Chem*, **276**, 10330-10337.

Marteau, P., Cuillerier, E., Meance, S., Gerhardt, M.F., Myara, A., Bouvier, M., Bouley, C., Tondu, F., Bommelaer, G. and Grimaud, J.C. (2002) *Bifidobacterium animalis* strain DN-173 010 shortens the colonic transit time in healthy women: a double-blind, randomized, controlled study. *Aliment Pharmacol Ther*, **16**, 587-593.

Maslowski, K.M., Vieira, A.T., Ng, A., Kranich, J., Sierro, F., Yu, D., Schilter, H.C., Rolph, M.S., Mackay, F., Artis, D., Xavier, R.J., Teixeira, M.M. and Mackay, C.R. (2009) Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature*, **461**, 1282-1286.

Matsuura, Y., Kusunoki, M., Harada, W. and Kakudo, M. (1984) Structure and possible catalytic residues of Taka-amylase A. *J Biochem*, **95**, 697-702.

Mayer, C., Zechel, D.L., Reid, S.P., Warren, R.A. and Withers, S.G. (2000) The E358S mutant of *Agrobacterium* sp. beta-glucosidase is a greatly improved glycosynthase. *FEBS Lett*, **466**, 40-44.

McCarter, J.D. and Withers, S.G. (1994) Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol*, **4**, 885-892.

McGaughey, G.B., Gagne, M. and Rappe, A.K. (1998) pi-Stacking interactions. Alive and well in proteins. *J Biol Chem*, **273**, 15458-15463.

McVie-Wylie, A.J., Lee, K.L., Qiu, H., Jin, X., Do, H., Gotschall, R., Thurberg, B.L., Rogers, C., Raben, N., O'Callaghan, M., Canfield, W., Andrews, L., McPherson, J.M. and Mattaliano, R.J. (2008) Biochemical and pharmacological characterization of different recombinant acid alpha-glucosidase preparations evaluated for the treatment of Pompe disease. *Mol Genet Metab*, **94**, 448-455.

Mercenier, A., Pavan, S. and Pot, B. (2003) Probiotics as biotherapeutic agents: present knowledge and future prospects. *Curr Pharm Des*, **9**, 175-191.

Metchnikoff, E. (1907) *The Prolongation of Life*. Heinemann, London.

Mitaku, S., Hirokawa, T. and Tsuji, T. (2002) Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. *Bioinformatics*, **18**, 608-616.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. and Olson, A.J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, **30**, 2785-2791.

Mountzouris, K.C., McCartney, A.L. and Gibson, G.R. (2002) Intestinal microflora of human infants and current trends for its nutritional modulation. *Br J Nutr*, **87**, 405-420.

Mountzouris, K.C., Tsirtsikos, P., Kalamara, E., Nitsch, S., Schatzmayr, G. and Fegeros, K. (2007) Evaluation of the efficacy of a probiotic containing *Lactobacillus*, *Bifidobacterium*, *Enterococcus*, and *Pediococcus* strains in promoting broiler performance and modulating cecal microflora composition and metabolic activities. *Poult Sci*, **86**, 309-317.

Munjal, N. and Sawhney, S.K. (2002) Stability and properties of mushroom tyrosinase entrapped in alginate, polyacrylamide and gelatin gels. *Enzyme and Microbial Technology*, **30**, 613-619.

Murata, T. and Usui, T. (2006) Enzymatic synthesis of oligosaccharides and neoglycoconjugates. *Biosci Biotechnol Biochem*, **70**, 1049-1059.

Murgas Torrazza, R. and Neu, J. (2011) The developing intestinal microbiome and its relationship to health and disease in the neonate. *J Perinatol*, **31 Suppl 1**, S29-34.

Musso, G., Gambino, R. and Cassader, M. (2011) Interactions between gut microbiota and host metabolism predisposing to obesity and diabetes. *Annu Rev Med*, **62**, 361-380.

[N]

Nagano, N., Orengo, C.A. and Thornton, J.M. (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J Mol Biol*, **321**, 741-765.

Nakao, M., Nakayama, T., Harada, M., Kakudo, A., Ikemoto, H., Kobayashi, S. and Shibano, Y. (1994a) Purification and characterization of a *Bacillus* sp. SAM1606 thermostable alpha-glucosidase with transglucosylation activity. *Appl Microbiol Biotechnol*, **41**, 337-343.

Nakao, M., Nakayama, T., Kakudo, A., Inohara, M., Harada, M., Omura, F. and Shibano, Y. (1994b) Structure and expression of a gene coding for thermostable alpha-glucosidase with a broad substrate specificity from *Bacillus* sp. SAM1606. *Eur J Biochem*, **220**, 293-300.

Narayanappa, D. (2008) Randomized double blinded controlled trial to evaluate the efficacy and safety of Bifilac in patients with acute viral diarrhea. *Indian J Pediatr*, **75**, 709-713.

Nazzaro, F., Fratianni, F., Nicolaus, B., Poli, A. and Orlando, P. (2012) The prebiotic source influences the growth, biochemical features and survival under simulated gastrointestinal conditions of the probiotic *Lactobacillus acidophilus*. *Anaerobe*, **18**, 280-285.

Neufeld, K.M., Kang, N., Bienenstock, J. and Foster, J.A. (2011) Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterol Motil*, **23**, 255-264, e119.

Ng, S.C., Hart, A.L., Kamm, M.A., Stagg, A.J. and Knight, S.C. (2009) Mechanisms of action of probiotics: recent advances. *Inflamm Bowel Dis*, **15**, 300-310.

Nguyen, T.D., Kang, J.H. and Lee, M.S. (2007) Characterization of *Lactobacillus plantarum* PH04, a potential probiotic bacterium with cholesterol-lowering effects. *Int J Food Microbiol*, **113**, 358-361.

Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W. and Pettersson, S. (2012) Host-gut microbiota metabolic interactions. *Science*, **336**, 1262-1267.

Nonaka, Y., Izumo, T., Izumi, F., Maekawa, T., Shibata, H., Nakano, A., Kishi, A., Akatani, K. and Kiso, Y. (2008) Antiallergic effects of *Lactobacillus pentosus* strain S-PT84 mediated by modulation of Th1/Th2 immunobalance and induction of IL-10 production. *Int Arch Allergy Immunol*, **145**, 249-257.

[O]

Ochoa-Leyva, A., Soberon, X., Sanchez, F., Arguello, M., Montero-Moran, G. and Saab-Rincon, G. (2009) Protein design through systematic catalytic loop exchange in the $(beta/alpha)_8$ fold. *J Mol Biol*, **387**, 949-964.

Ohnuma, T., Fukuda, T., Dozen, S., Honda, Y., Kitaoka, M. and Fukamizo, T. (2012) A glycosynthase derived from an inverting GH19 chitinase from the moss *Bryum coronatum*. *Biochem J*, **444**, 437-443.

Okuyama, M., Kitamura, M., Hondoh, H., Kang, M.S., Mori, H., Kimura, A., Tanaka, I. and Yao, M. (2009) Catalytic mechanism of retaining alpha-galactosidase belonging to glycoside hydrolase family 97. *J Mol Biol*, **392**, 1232-1241.

Okuyama, M., Mori, H., Watanabe, K., Kimura, A. and Chiba, S. (2002) Alpha-glucosidase mutant catalyzes "alpha-glycosynthase"-type reaction. *Biosci Biotechnol Biochem*, **66**, 928-933.

[P]

Pace, C.N. and Scholtz, J.M. (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J*, **75**, 422-427.

Page, I.H. (1928) Further observations on the chemical composition of Woods Hole sea water - The chlorine content and salt analysis. *The Biological Bulletin*, **55**, 449-452.

Palcic, M.M., Scaman, C.H., Otter, A., Szpacenko, A., Romaniouk, A., Li, Y.X. and Vijay, I.K. (1999) Processing alpha-glucosidase I is an inverting glycosidase. *Glycoconj J*, **16**, 351-355.

Panke, S., Held, M. and Wubbolts, M. (2004) Trends and innovations in industrial biocatalysis for the production of fine chemicals. *Curr Opin Biotechnol*, **15**, 272-279.

Panke, S. and Wubbolts, M. (2005) Advances in biocatalytic synthesis of pharmaceutical intermediates. *Curr Opin Chem Biol*, **9**, 188-194.

Panoff, J.M., Legrand, S., Thammavongs, B. and Boutibonnes, P. (1994) The cold shock response in *Lactococcus lactis* subsp. *lactis*. *Current Microbiology*, **29**, 213-216.

Panoff, J.M., Thammavongs, B., Laplace, J.M., Hartke, A., Boutibonnes, P. and Auffray, Y. (1995) Cryotolerance and Cold Adaptation in *Lactococcus lactis* Subsp. *lactis* IL1403. *Cryobiology*, **32**, 513-520.

Parker, R.B. (1974) Probiotics, the other half of the antibiotic story. *Animal Nutrition and Health*, **29**, 4-8.

Parsiegla, G., Juy, M., Reverbel-Leroy, C., Tardif, C., Belaich, J.P., Driguez, H. and Haser, R. (1998) The crystal structure of the processive endocellulase CelF of *Clostridium*

cellulolyticum in complex with a thiooligosaccharide inhibitor at 2.0 A resolution. *EMBO J*, **17**, 5551-5562.

Patel, S. and Goyal, A. (2012) The current trends and future perspectives of prebiotics research: a review. *3 Biotech*, **2**, 115-125.

Paturi, G., Phillips, M., Jones, M. and Kailasapathy, K. (2007) Immune enhancing effects of *Lactobacillus acidophilus* LAFTI L10 and *Lactobacillus paracasei* LAFTI L26 in mice. *Int J Food Microbiol*, **115**, 115-118.

Pauling, L. (1946) Molecular architecture and biological reactions. *Chemical & Engineering News*, **24**, 1375-1377.

Pauling, L. (1948) Nature of forces between large molecules of biological interest. *Nature*, **161**, 707-709.

Paulsen, H., Sangster, I. and Heyns, K. (1967) Monosaccharide mit stickstoffhaltigem Ring, XIII. Synthese und Reaktionen von Keto-piperidinosen. *Chemische Berichte*, **100**, 802-815.

Perez-Burgos, A., Wang, B., Mao, Y.K., Mistry, B., McVey Neufeld, K.A., Bienenstock, J. and Kunze, W.A. (2012) Psychoactive bacteria *Lactobacillus rhamnosus* (JB-1) elicits rapid frequency facilitation in vagal afferents. *Am J Physiol Gastrointest Liver Physiol*.

Perpetuo, E.A., Juliano, L. and Lebrun, I. (2003) Biochemical and pharmacological aspects of two bradykinin-potentiating peptides obtained from tryptic hydrolysis of casein. *J Protein Chem*, **22**, 601-606.

Perugino, G., Trincone, A., Giordano, A., van der Oost, J., Kaper, T., Rossi, M. and Moracci, M. (2003) Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic pH. *Biochemistry*, **42**, 8484-8493.

Perutz, M.F. (1956) Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Cryst.*, **9**, 867-673.

Phizicky, E.M. and Fields, S. (1995) Protein-protein interactions: methods for detection and analysis. *Microbiol Rev*, **59**, 94-123.

Pikis, A., Immel, S., Robrish, S.A. and Thompson, J. (2002) Metabolism of sucrose and its five isomers by *Fusobacterium mortiferum*. *Microbiology*, **148**, 843-852.

Pool-Zobel, B.L., Neudecker, C., Domizlaff, I., Ji, S., Schillinger, U., Rumney, C., Moretti, M., Vilarini, I., Scassellati-Sforzolini, R. and Rowland, I. (1996) *Lactobacillus*- and *Bifidobacterium*-mediated antigenotoxicity in the colon of rats. *Nutr Cancer*, **26**, 365-380.

Preidis, G.A. and Versalovic, J. (2009) Targeting the human microbiome with antibiotics, probiotics, and prebiotics: gastroenterology enters the metagenomics era. *Gastroenterology*, **136**, 2015-2031.

Puri, M., Marwaha, S.S. and Kothari, R.M. (1996) Studies on the applicability of alginate-entrapped naringinase for the debittering of kinnow juice. *Enzyme and Microbial Technology*, **18**, 281-285.

[Q]

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Jian, M., Zhou, Y., Li, Y., Zhang, X., Qin, N., Yang, H., Wang,

J., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Meta, H.I.T.C., Bork, P. and Ehrlich, S.D. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, **464**, 59-65.

[R]

Raasch, C., Streit, W., Schanzer, J., Bibel, M., Gosslar, U. and Liebl, W. (2000) *Thermotoga maritima* AglA, an extremely thermostable NAD^+-, Mn^{2+}-, and thiol-dependent alpha-glucosidase. *Extremophiles*, **4**, 189-200.

Rajan, S.S., Yang, X., Collart, F., Yip, V.L., Withers, S.G., Varrot, A., Thompson, J., Davies, G.J. and Anderson, W.F. (2004) Novel catalytic mechanism of glycoside hydrolysis based on the structure of an NAD^+/Mn^{2+} -dependent phospho-alpha-glucosidase from *Bacillus subtilis*. *Structure*, **12**, 1619-1629.

Rajilic-Stojanovic, M., Smidt, H. and de Vos, W.M. (2007) Diversity of the human gastrointestinal tract microbiota revisited. *Environ Microbiol*, **9**, 2125-2136.

Rao, A.V., Bested, A.C., Beaulne, T.M., Katzman, M.A., Iorio, C., Berardi, J.M. and Logan, A.C. (2009) A randomized, double-blind, placebo-controlled pilot study of a probiotic in emotional symptoms of chronic fatigue syndrome. *Gut Pathog*, **1**, 6.

Rao, F.V., Dorfmueller, H.C., Villa, F., Allwood, M., Eggleston, I.M. and van Aalten, D.M. (2006) Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. *EMBO J*, **25**, 1569-1578.

Rao, S.T. and Rossmann, M.G. (1973) Comparison of super-secondary structures in proteins. *J Mol Biol*, **76**, 241-256.

Rastall, R.A. (2004) Bacteria in the gut: friends and foes and how to alter the balance. *J Nutr*, **134**, 2022S-2026S.

Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R. and Aghajari, N. (2007) Trehalulose synthase native and carbohydrate complexed structures provide insights into sucrose isomerization. *J Biol Chem*, **282**, 28126-28136.

Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R. and Aghajari, N. (2009) Structural determinants of product specificity of sucrose isomerases. *FEBS Lett*, **583**, 1964-1968.

Reid, T.M., Leiper, J., Inglis, J.R. and Solomon, J.B. (1973) Phytohaemagglutinin-induced transformation of thymus and spleen cells from foetal guinea-pigs cultured in vitro or in diffusion chambers. *Immunology*, **24**, 203-210.

Rhimi, M., Boisson, A., Dejob, M., Boudebouze, S., Maguin, E., Haser, R. and Aghajari, N. (2010) Efficient bioconversion of lactose in milk and whey: immobilization and biochemical characterization of a beta-galactosidase from the dairy *Streptococcus thermophilus* LMD9 strain. *Res. Microbiol.*, **161**, 515-525.

Roberfroid, M. (2007) Prebiotics: the concept revisited. *J Nutr*, **137**, 830S-837S.

Robert, X., Haser, R., Gottschalk, T.E., Ratajczak, F., Driguez, H., Svensson, B. and Aghajari, N. (2003) The structure of barley alpha-amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs. *Structure*, **11**, 973-984.

Rodionov, D.A., Yang, C., Li, X., Rodionova, I.A., Wang, Y., Obraztsova, A.Y., Zagnitko, O.P., Overbeek, R., Romine, M.F., Reed, S., Fredrickson, J.K., Nealson, K.H. and Osterman, A.L. (2010) Genomic encyclopedia of sugar utilization pathways in the *Shewanella* genus. *BMC Genomics*, **11**, 494.

Rodriguez-Sanoja, R., Oviedo, N. and Sanchez, S. (2005) Microbial starch-binding domain. *Curr Opin Microbiol*, **8**, 260-267.

Rolfsmeier, M. and Blum, P. (1995) Purification and characterization of a maltase from the extremely thermophilic crenarchaeote *Sulfolobus solfataricus*. *J Bacteriol*, **177**, 482-485.

Rolfsmeier, M., Haseltine, C., Bini, E., Clark, A. and Blum, P. (1998) Molecular characterization of the alpha-glucosidase gene (*malA*) from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol*, **180**, 1287-1295.

Rossmann, M.G. and Blow, D.M. (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst.*, **15**, 24-31.

Round, J.L. and Mazmanian, S.K. (2010) Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc Natl Acad Sci U S A*, **107**, 12204-12209.

Rowe, G.E. and Margaritis, A. (2004) Enzyme kinetic properties of alpha-glucosidase in *Bacillus thuringiensis*. *Biochemical Engineering Journal*, **17**, 121-128.

Ruszczynski, M., Radzikowski, A. and Szajewska, H. (2008) Clinical trial: effectiveness of *Lactobacillus rhamnosus* (strains E/N, Oxy and Pen) in the prevention of antibiotic-associated diarrhoea in children. *Aliment Pharmacol Ther*, **28**, 154-161.

Rydberg, E.H., Li, C., Maurus, R., Overall, C.M., Brayer, G.D. and Withers, S.G. (2002) Mechanistic analyses of catalysis in human pancreatic alpha-amylase: detailed kinetic and structural studies of mutants of three conserved carboxylic acids. *Biochemistry*, **41**, 4492-4502.

Rye, C.S. and Withers, S.G. (2000) Glycosidase mechanisms. *Curr Opin Chem Biol*, **4**, 573-580.

[S]

Saltikov, C.W., Cifuentes, A., Venkateswaran, K. and Newman, D.K. (2003) The ars detoxification system is advantageous but not required for As(V) respiration by the genetically tractable *Shewanella* species strain ANA-3. *Appl Environ Microbiol*, **69**, 2800-2809.

Sanchez-Weatherby, J., Bowler, M.W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R.B. and Cipriani, F. (2009) Improving diffraction by humidity control: a novel device compatible with X-ray beamlines. *Acta Crystallogr D Biol Crystallogr*, **65**, 1237-1246.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, **74**, 5463-5467.

Savage, D.C. (1977) Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol*, **31**, 107-133.

Savage, D.C., Siegel, J.E., Snellen, J.E. and Whitt, D.D. (1981) Transit time of epithelial cells in the small intestines of germfree mice and ex-germfree mice associated with indigenous microorganisms. *Appl Environ Microbiol*, **42**, 996-1001.

Schomburg, I., Chang, A., Hofmann, O., Ebeling, C., Ehrentreich, F. and Schomburg, D. (2002) BRENDA: a resource for enzyme data and metabolic information. *Trends in Biochemical Sciences*, **27**, 54-56.

Schrezenmeir, J. and de Vrese, M. (2001) Probiotics, prebiotics, and synbiotics--approaching a definition. *Am J Clin Nutr*, **73**, 361S-364S.

Seo, S.H., Choi, K.H., Hwang, S., Kim, J., Park, C.S., Rho, J.R. and Cha, J. (2011) Characterization of the catalytic and kinetic properties of a thermostable *Thermoplasma acidophilum* alpha-glucosidase and its transglucosylation reaction with arbutin. *Journal of Molecular Catalysis B: Enzymatic*, **72**, 305-312.

Shanahan, F. (2002) The host-microbe interface within the gut. *Best Pract Res Clin Gastroenterol*, **16**, 915-931.

Sharon, I., Morowitz, M.J., Thomas, B.C., Costello, E.K., Relman, D.A. and Banfield, J.F. (2013) Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res*, **23**, 111-120.

Shirai, T., Hung, V.S., Morinaka, K., Kobayashi, T. and Ito, S. (2008) Crystal structure of GH13 alpha-glucosidase GSJ from one of the deepest sea bacteria. *Proteins*, **73**, 126-133.

Siddiqui, K.S. and Cavicchioli, R. (2006) Cold-adapted enzymes. *Annu Rev Biochem*, **75**, 403-433.

Sim, L., Quezada-Calvillo, R., Sterchi, E.E., Nichols, B.L. and Rose, D.R. (2008) Human intestinal maltase-glucoamylase: crystal structure of the N-terminal catalytic subunit and basis of inhibition and substrate specificity. *J Mol Biol*, **375**, 782-792.

Sinnott, M. (1998) *Comprehensive Biological Catalysis - A mechanistic reference*. Academic Press, London.

Smalås, A.O., Leiros, H.K., Os, V. and Willassen, N.P. (2000) Cold adapted enzymes. *Biotechnol Annu Rev*, **6**, 1-57.

Sreerama, N. and Woody, R.W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem*, **287**, 252-260.

Stam, M.R., Danchin, E.G., Rancurel, C., Coutinho, P.M. and Henrissat, B. (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng Des Sel*, **19**, 555-562.

Stowell, J. (ed.). (2007) *Calorie Control and Weight Management, in Sweeteners and Sugar Alternatives in Food Technology*. Blackwell Publishing Ltd, Oxford, UK.

Struvay, C. and Feller, G. (2012) Optimization to low temperature activity in psychrophilic enzymes. *Int J Mol Sci*, **13**, 11643-11665.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*, **185**, 60-89.

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X.N., Kubo, C. and Koga, Y. (2004) Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J Physiol*, **558**, 263-275.

Sugimura, M., Nishimoto, M. and Kitaoka, M. (2006) Characterization of glycosynthase mutants derived from glycoside hydrolase family 10 xylanases. *Biosci Biotechnol Biochem*, **70**, 1210-1217.

Suzuki, R., Katayama, T., Kitaoka, M., Kumagai, H., Wakagi, T., Shoun, H., Ashida, H., Yamamoto, K. and Fushinobu, S. (2009) Crystallographic and mutational analyses of substrate recognition of endo-alpha-N-acetylgalactosaminidase from *Bifidobacterium longum*. *J Biochem*, **146**, 389-398.

Suzuki, Y., Nobiki, M., Matsuda, M. and Sawai, T. (1997) *Bacillus thermoamyloliquefaciens* KP1071 alpha-glucosidase II is a thermostable M(r) 540,000 homohexameric alpha-glucosidase with both exo-alpha-1,4-glucosidase and oligo-1,6-glucosidase activities. *Eur J Biochem*, **245**, 129-136.

Svensson, B., Jespersen, H., Sierks, M.R. and MacGregor, E.A. (1989) Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. *Biochem J*, **264**, 309-311.

Szymanski, H., Armanska, M., Kowalska-Duplaga, K. and Szajewska, H. (2008) *Bifidobacterium longum* PL03, *Lactobacillus rhamnosus* KL53A, and *Lactobacillus plantarum* PL02 in the prevention of antibiotic-associated diarrhea in children: a randomized controlled pilot trial. *Digestion*, **78**, 13-17.

[T]

Tan, K., Tesar, C., Wilton, R., Keigher, L., Babnigg, G. and Joachimiak, A. (2010) Novel alpha-glucosidase from human gut microbiome: substrate specificities and their switch. *FASEB J*, **24**, 3939-3949.

Tanaka, H., Kurosawa, H., Kokufuta, E. and Veliky, I.A. (1984) Preparation of immobilized glucomylase using Ca-alginate gel coated with partially quaterized poly(ethleneimine). *Biotechnol Bioeng*, **26**, 1393-1394.

Tanaka, Y., Aki, T., Hidaka, Y., Furuya, Y., Kawamoto, S., Shigeta, S., Ono, K. and Suzuki, O. (2002) Purification and characterization of a novel fungal alpha-glucosidase from *Mortierella alliacea* with high starch-hydrolytic activity. *Biosci Biotechnol Biochem*, **66**, 2415-2423.

Tanaka, Y., Tao, W., Blanchard, J.S. and Hehre, E.J. (1994) Transition state structures for the hydrolysis of alpha-D-glucopyranosyl fluoride by retaining and inverting reactions of glycosylases. *J Biol Chem*, **269**, 32306-32312.

Terwisscha van Scheltinga, A.C., Armand, S., Kalk, K.H., Isogai, A., Henrissat, B. and Dijkstra, B.W. (1995) Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry*, **34**, 15619-15623.

Timmerman, H.M., Koning, C.J., Mulder, L., Rombouts, F.M. and Beynen, A.C. (2004) Monostrain, multistrain and multispecies probiotics--A comparison of functionality and efficacy. *Int J Food Microbiol*, **96**, 219-233.

Tommasini, R., Endrenyi, L., Taylor, P.A., Mahuran, D.J. and Lowden, J.A. (1985) A statistical comparison of parameter estimation for the Michaelis-Menten kinetics of human placental hexosaminidase. *Can J Biochem Cell Biol*, **63**, 225-230.

Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeyssens, M. (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem*, **170**, 575-581.

Tomme, P., Warren, R.A., Miller, R.C.J., Kilburn, D.G. and Gilkes, N.R. (1995) *Enzymatic Degradation of Insoluble Polysaccharides*, Washington.

Tsai, Y.T., Cheng, P.C. and Pan, T.M. (2012) The immunomodulatory effects of lactic acid bacteria for improving immune functions and benefits. *Appl Microbiol Biotechnol*, **96**, 853-862.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., Egholm, M., Henrissat, B., Heath, A.C., Knight, R. and Gordon, J.I. (2009) A core gut microbiome in obese and lean twins. *Nature*, **457**, 480-484.

Turpin, W., Humblot, C., Noordine, M.L., Thomas, M. and Guyot, J.P. (2012) *Lactobacillaceae* and cell adhesion: genomic and functional screening. *PLoS One*, **7**, e38034.

Tusnady, G.E. and Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol*, **283**, 489-506.

[U-V-W]

Uitdehaag, J.C., Mosi, R., Kalk, K.H., van der Veen, B.A., Dijkhuizen, L., Withers, S.G. and Dijkstra, B.W. (1999) X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the alpha-amylase family. *Nat Struct Biol*, **6**, 432-436.

Vagin, A. and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr*, **66**, 22-25.

Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F. and Murshudov, G.N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr*, **60**, 2184-2195.

van de Guchte, M., Penaud, S., Grimaldi, C., Barbe, V., Bryson, K., Nicolas, P., Robert, C., Oztas, S., Mangenot, S., Couloux, A., Loux, V., Dervyn, R., Bossy, R., Bolotin, A., Batto, J.M., Walunas, T., Gibrat, J.F., Bessieres, P., Weissenbach, J., Ehrlich, S.D. and Maguin, E. (2006) The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. *Proc Natl Acad Sci U S A*, **103**, 9274-9279.

van den Broek, L.A., Struijs, K., Verdoes, J.C., Beldman, G. and Voragen, A.G. (2003) Cloning and characterization of two alpha-glucosidases from *Bifidobacterium adolescentis* DSM20083. *Appl Microbiol Biotechnol*, **61**, 55-60.

Van Tassell, M.L. and Miller, M.J. (2011) *Lactobacillus* adhesion to mucus. *Nutrients*, **3**, 613-636.

Van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhabhai, R. and Pettersson, G. (1986) Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*: Separation of functional domains. *FEBS Lett*, **204**, 223-227.

Vester-Christensen, M.B., Abou Hachem, M., Svensson, B. and Henriksen, A. (2010) Crystal structure of an essential enzyme in seed starch degradation: barley limit dextrinase in complex with cyclodextrins. *J Mol Biol*, **403**, 739-750.

Vieille, C., Epting, K.L., Kelly, R.M. and Zeikus, J.G. (2001) Bivalent cations and amino-acid composition contribute to the thermostability of *Bacillus licheniformis* xylose isomerase. *Eur J Biochem*, **268**, 6291-6301.

Vinderola, G., Perdigon, G., Duarte, J., Farnworth, E. and Matar, C. (2006) Effects of the oral administration of the exopolysaccharide produced by *Lactobacillus kefiranofaciens* on the gut mucosal immunity. *Cytokine*, **36**, 254-260.

Vizoso Pinto, M.G., Rodriguez Gomez, M., Seifert, S., Watzl, B., Holzapfel, W.H. and Franz, C.M. (2009) *Lactobacilli* stimulate the innate immune response and modulate the TLR expression of HT29 intestinal epithelial cells *in vitro*. *Int J Food Microbiol*, **133**, 86-93.

Waller, P.A., Gopal, P.K., Leyer, G.J., Ouwehand, A.C., Reifer, C., Stewart, M.E. and Miller, L.E. (2011) Dose-response effect of *Bifidobacterium lactis* HN019 on whole gut transit time and functional gastrointestinal symptoms in adults. *Scand J Gastroenterol*, **46**, 1057-1064.

Wang, L.X. and Huang, W. (2009) Enzymatic transglycosylation for glycoconjugate synthesis. *Curr Opin Chem Biol*, **13**, 592-600.

Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y. and Suzuki, Y. (1997) The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 A resolution: structural characterization of proline-substitution sites for protein thermostabilization. *J Mol Biol*, **269**, 142-153.

Watson, J.N., Dookhun, V., Borgford, T.J. and Bennet, A.J. (2003) Mutagenesis of the conserved active-site tyrosine changes a retaining sialidase into an inverting sialidase. *Biochemistry*, **42**, 12682-12690.

Wehmeier, U.F. and Piepersberg, W. (2004) Biotechnology and molecular biology of the alpha-glucosidase inhibitor acarbose. *Appl Microbiol Biotechnol*, **63**, 613-625.

Wolfenden, R. (1999) Conformational aspects of inhibitor design: enzyme-substrate interactions in the transition state. *Bioorg Med Chem*, **7**, 647-652.

Wolfenden, R. and Snider, M.J. (2001) The depth of chemical time and the power of enzymes as catalysts. *Acc Chem Res*, **34**, 938-945.

Wostmann, B.S., Larkin, C., Moriarty, A. and Bruckner-Kardoss, E. (1983) Dietary intake, energy metabolism, and excretory losses of adult male germfree Wistar rats. *Lab Anim Sci*, **33**, 46-50.

[X-Y-Z]

Xavier, R.J. and Podolsky, D.K. (2007) Unravelling the pathogenesis of inflammatory bowel disease. *Nature*, **448**, 427-434.

Xia, L., Rajendran, C., Ruppert, M., Panjikar, S., Wang, M. and Stoeckigt, J. (2012) High speed X-ray analysis of plant enzymes at room temperature. *Phytochemistry*.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V. and Gordon, J.I. (2003) A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science*, **299**, 2074-2076.

Xu, J. and Gordon, J.I. (2003) Honor thy symbionts. *Proc Natl Acad Sci U S A*, **100**, 10452-10459.

Yamasaki, Y. and Konno, H. (1987) Purification and properties of wall-bound alpha-glucosidase from suspension-cultured rice cells. *Agric. Biol. Chem.*, **51**, 3239-3244.

Yamasaki, Y., Suzuki, Y. and Ozawa, J. (1976) Purification and properties of alpha-glucosidase from *Penicillium purpurogenum*. *Agric. Biol. Chem.*, **40**, 669-676.

Yang, X., Xie, L., Li, Y. and Wei, C. (2009) More than 9,000,000 unique genes in human gut bacterial community: estimating gene numbers inside a human body. *PLoS One*, **4**, e6074.

Yeo, S.K. and Liong, M.T. (2010) Effect of prebiotics on viability and growth characteristics of probiotics in soymilk. *J Sci Food Agric*, **90**, 267-275.

Yip, V.L., Varrot, A., Davies, G.J., Rajan, S.S., Yang, X., Thompson, J., Anderson, W.F. and Withers, S.G. (2004) An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 beta-glycosidase from *Thermotoga maritima*. *J Am Chem Soc*, **126**, 8354-8355.

Yoshikawa, M., Murakami, T., Shimada, H., Matsuda, H., Yamahara, J., Tanabe, G. and Muraoka, O. (1997) Salacinol, potent antidiabetic principle with unique thiosugar sulfonium sulfate structure from the Ayurvedic traditional medicine *Salacia reticulata* in Sri Lanka and India. *Tetrahedron Letters*, **38**, 8367-8370.

Yun, B.Y., Jun, S.Y., Kim, N.A., Yoon, B.Y., Piao, S., Park, S.H., Jeong, S.H., Lee, H. and Ha, N.C. (2011) Crystal structure and thermostability of a putative alpha-glucosidase from *Thermotoga neapolitana*. *Biochem Biophys Res Commun*, **416**, 92-98.

Zhang, D., Li, N., Lok, S.M., Zhang, L.H. and Swaminathan, K. (2003) Isomaltulose synthase (PalI) of *Klebsiella* sp. LX3. Crystal structure and implication of mechanism. *J Biol Chem*, **278**, 35428-35434.

Zhou, C., Xue, Y., Zhang, Y., Zeng, Y. and Ma, Y. (2009) Recombinant expression and characterization of *Thermoanaerobacter tengcongensis* thermostable alpha-glucosidase with regioselectivity for high yield isomaltooligosaccharides synthesis. *J Microbiol Biotechnol*, **19**, 1547-1556.

Zibaee, A., Bandani, A.R. and Ramzi, S. (2009) Enzymatic properties of alpha- and beta-glucosidases extracted from midgut and salivary glands of rice striped stem borer, *Chilo suppressalis* Walker (*Lepidoptera: Pyralidae*). *C R Biol*, **332**, 633-641.

Zivkovic, A.M., German, J.B., Lebrilla, C.B. and Mills, D.A. (2011) Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc Natl Acad Sci U S A*, **108 Suppl 1**, 4653-4658.

Zoetendal, E.G., Raes, J., van den Bogert, B., Arumugam, M., Booijink, C.C., Troost, F.J., Bork, P., Wels, M., de Vos, W.M. and Kleerebezem, M. (2012) The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J*, **6**, 1415-1426.

Zoetendal, E.G., Rajilic-Stojanovic, M. and de Vos, W.M. (2008) High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut*, **57**, 1605-1615.

Publications

Shewanella sp. ANA-3 GH97 α-glucosidase is active at low temperatures and chloride dependent

Running title: Cold-active *Shewanella* sp. ANA-3 α-glucosidase

Magali Dejob[a], Moez Rhimi[a,b], Rimeh Ilhammami[a], Richard Haser[a] & Nushin Aghajari[a#]

[a]*Laboratory for BioCrystallography and Structural Biology of Therapeutic Targets Molecular and Structural Bases of Infectious Systems, UMR5086 CNRS-University of Lyon1, Institute 7 passage du Vercors, 69367 Lyon Cedex 07, France.*
[b]*Present address: INRA, UMR 1319 Micalis, 78350, Jouy-en-Josas, France.*

[#] Corresponding author
e-mail: nushin.aghajari@ibcp.fr
Phone +33 4 72 72 26 33
Fax: +33 4 72 72 26 16

Keywords: α-glucosidase, cold-active, chloride dependence, immobilization, *Shewanella*, glycoside hydrolase family 97

Abbreviations: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; IPTG, isopropyl-β-D-thiogalactopyranoside; MES, 2-(N-morpholino)ethanesulfonic acid; *p*-NPG, 4-nitrophenyl α-D-glucopyranoside.

ABSTRACT

Carbohydrates constitute an essential source of carbon and energy for a variety of bacteria. The diversity of microbial lifestyles is reflected in the variations of sugar metabolism. *Shewanella* sp. ANA-3 found in brackish water possesses 7 different sugar utilization pathways, including the maltodextrin pathway. This latter contains the hydrolytic SusB enzyme belonging to glycoside hydrolase family GH97. Here, we report the cloning, expression and characterization of *Shw*SusB, which hydrolyses maltooligosaccharides to glucose in *Shewanella* sp. ANA-3. The gene encoding *Shw*SusB was cloned and expressed in *Escherichia coli*. The purified enzyme was a monomer in its active state and remarkably showed maltose-hydrolyzing activity at low temperatures, retaining 50% of its activity at 5 °C. Our work indicated that whereas the enzyme is thermostable at temperatures up to 45 °C, it displays thermolability above this temperature with a half-life of 60 min and 30 min at 50 °C and 55 °C, respectively. Upon whole cell immobilization in alginate beads the thermostability was improved at 50 °C and this α-glucosidase moreover displayed noticeable activity in a broad pH range from pH 5 to 8.5. Activation by calcium ions is a hallmark of GH97 α-glucosidases. Our results demonstrated that also chloride ions are essential for the activity of this cold-active enzyme.

INTRODUCTION

Cold regions have been colonized by a broad diversity of microorganisms (1) defined as being psychrophile or psychrotolerant as e.g. *Shewanellaceae*. Bacteria of the genus *Shewanella* are a versatile group of Gram-negative γ-proteobacteria which thrive in different environments such as aquatic and sedimentary systems (2). They demonstrate significant variability in their metabolic and ecophysiological capabilities including cold- and salt tolerance (3). Enzymes produced by psychrophilic microorganisms are adapted to function at low temperatures (4) and display a low thermal stability as compared to their mesophilic homologues (5). There seems to be a direct correlation between activity and stability, and the thermolability of psychrophilic enzymes can be explained by a higher molecular flexibility (6). Carbohydrates present one of the most important natural carbon- and energy sources for bacteria in which several sugar utilization pathways have been identified. As concerns *Shewanella* genus which are capable of utilizing more or less different sugars as carbon sources depending on the strain, different carbohydrate utilization machineries are observed

(7). Nonetheless essential carbohydrate pathways are poorly studied in *Shewanella* strains which are better known for their remarkable respiratory diversity (3, 8-11). Indeed, *Shewanella* species are recognized for their role in remediation of environments contaminated by toxic elements, radionuclides or inorganic compounds to non-toxic chemical states (2, 12).

Shewanella sp. ANA-3 is a motile, anaerobe and nonsporulating bacterium which despite being mesophile is psychrotolerant with an optimum growth temperature range between 25 °C and 37 °C (13). This strain possesses 7 different sugar utilization pathways comprising a total of 87 protein coding genes involved in sugar utilization (7). Amongst these, the maltodextrin pathway contains several genes encoding hydrolytic enzymes (*amy*, *cga*, *susA*, *susB* and *malZ*) along with sugar transporters which are all encoded in the genome and confirmed by growth phenotype characterization (7). In brackish water which is their natural habitat their presence may be explained by the existence of plant materials as source of maltodextrins. One of the starch utilization genes (sus) encodes for an α-glucosidase, SusB. α-Glucosidases (EC 3.2.1.20; α-D-glucoside glucohydrolase) are enzymes involved in carbohydrate metabolism as eg. the maltodextrin pathway. These exo-glycosidases catalyze the hydrolysis of α-1,4-linked-D-glucose non-reducing end residues with release of α-D-glucose. The physiological role of most exo-type glycosidases as α-glucosidases is to produce monosaccharides serving as carbon- and energy sources. α-Glucosidases reported in the literature so far are mainly active at acidic pH. As concerns the temperature range at which these enzymes are active, so far only one α-glucosidase from *Aspergillus niger* has been reported to retain 50% activity at 5 °C (14).

Here we report the cloning, over-expression, purification and biochemical characterization of the cold active *Shewanella* sp. ANA-3 α-glucosidase (*Shw*SusB) which belongs to glycoside hydrolase family 97. Immobilization of enzymes in various supports as eg. beads has been reported to improve enzyme activity (15, 16). Within these studies we explored the effect of immobilization of the enzyme on its activity as a function of pH and temperature.

MATERIALS AND METHODS

Bacterial strains, media and plasmids.

Escherichia coli BL21 (DE3) was used as host strain. Culture of *E. coli* was done in Luria Bertani (LB) medium supplemented, when necessary, with 100 µM ampicillin and 0.5 mM IPTG (isopropyl β-D-thiogalactopyranoside). The pET-15b plasmid (Novagen) was used according to manufacturer's instructions.

DNA manipulation, PCR and molecular cloning of the *Shewanella* sp. ANA-3 α-glucosidase gene.

Preparation of plasmid DNA, digestion with restriction endonucleases and separation of fragments by agarose gel electrophoresis were performed as described by Sambrook (17). The gene encoding the α-glucosidase (GenBank™ accession number: ABK48162.1) was amplified from *Shewanella* sp. ANA-3 genomic DNA by PCR using the two oligonucleotides: forward primer 5'-gtcgtccatATGTTCAGCAATCCTCTTACCCACGTACG-3' and reverse primer 5'-ggaattcttaTTACAGCGCTTTAAAGCGGATCGCTGTG-3' with *Nde*I and *Eco*RI sites (underlined), respectively. PCR amplification was carried out using Gene Amp® PCR System 9700 (Applied Biosystems). The amplification reaction mixture (100 µl) contained Phusion HF buffer, 20 pmol of each primer, 10 ng of DNA template and 10 units of Phusion High-Fidelity DNA Polymerase (Thermo Scientific). The cycling parameters during 35 cycles were: 94 °C, 30 s; 55 °C, 60 s and 72 °C, 120 s. The amplified gene fragment was purified, digested with *Nde*I and *Eco*RI and ligated into the pET-15b vector linearized with the same restriction enzymes (Novagen, Madison, WI, USA). The ligation product was transformed into *Escherichia coli* BL21 (DE3) competent cells (Invitrogen). A clone, harbouring the *Shewanella* sp. ANA-3 α-glucosidase gene under the control of the T7 promoter, was retained for further studies including sequence analysis.

Primary structure analyses.

Multiple alignment of amino-acid sequences was done using the program Clustal W (18, 19) and rendering of the alignment was performed with the ESPript web server (20). Sequences used for multiple alignment included *Shewanella* sp. ANA-3 SusB (*Shw*SusB: accession number A0KWJ3), *B. thetaiotaomicron* SusB (BtGH97a: accession number P71094) and *B. thetaiotaomicron* α-galactosidase (BtGH97b: accession number Q8A6L0).

Expression and purification.

E. coli BL21 (DE3) transformants were used for inoculation of 5 ml of LB media containing 100 μM ampicillin at 37 °C (under shaking) overnight, and hereafter diluted in 300 ml of LB media containing 100 μM ampicillin and grown at 37 °C until an $OD_{600\ nm}$ of 0.8. The protein production was induced with 0.5 mM IPTG overnight at 25 °C. Cells were harvested by centrifugation and resuspended in 100 mM MES (2-(N-morpholino)ethanesulfonic acid) buffer (pH 6.5) containing 10 mg/ml lysozyme (Sigma) to a final concentration of 0.2 g/ml. The cell resuspension was incubated for 3 h at 4 °C and sonicated for cell lysis at 4°C for 5 min (pulsations of 0.5 s, amplitude 40) using a Branson Sonifier 250 sonicator. The soluble fraction was obtained by centrifugation at 10000 *g* for 20 min.

For purification, the supernatant was applied to a 1 ml IMAC column (GE Healthcare) pre-equilibrated in 100 mM MES buffer (pH 6.5). The protein was eluted with 60 mM imidazole. Next, fractions with α-glucosidase activity were pooled and loaded onto a Mono-Q anion-exchange column (Mono Q™ 5/50GL, GE Healthcare) pre-equilibrated in 100 mM MES buffer (pH 6.5). The protein was eluted using a linear NaCl gradient ranging from 0 to 0.5 M in the same buffer.

Protein quantification and electrophoresis.

The protein concentration was determined using a NanoDrop 2000c spectrophotometer (Thermo Scientific). Protein samples were separated in 10% SDS-PAGE according to the Laemmli method (21). Protein bands were visualized by Page Blue™ Protein Staining (Fermentas) solution.

Cross-linking.

The amine-specific homobifunctional cross-linker bis(sulfosuccinimidyl) suberate (BS^3; Thermo Scientific) was used to elucidate whether the protein is a monomer or a dimer in solution. Purified α-glucosidase was diluted to 25 μM in a reaction buffer containing 100 mM MES pH 6.5 into which 50 mM freshly prepared BS^3 solution was added to 20-, 40- or 80-molar excess. A first series of reaction mixtures was incubated for 45 min and a second series for 60 min, respectively, at room temperature and thereafter quenched with Tris-HCl pH 8.0 to a final concentration of 10 mM.

Preparation of cells and immobilization.

Immobilization of the enzyme was carried out by mixing 10 ml recombinant *E. coli* cells (whole cell immobilization) or 10 ml of soluble fraction obtained after cell lysis (enzyme immobilization) with 10 ml 6.0% (w/v) sodium alginate (Fluka). In order to find the optimal conditions for activity in alginate beads, varying cell concentrations between 20 and 45 g/l were tested. Then, the obtained mixture was dropped slowly using a syringe into a cooled and stirred 200 mM calcium chloride ($CaCl_2$) solution (200 ml). The beads having an average size of approximately 0.7 mm were washed twice in 100 mM MES buffer (pH 6.5).

Enzymatic assay.

α-Glucosidase activity on *p*-NPG (4-nitrophenyl α-D-glucopyranoside; Sigma) was determined by measuring *p*-nitrophenol released under different reaction conditions. The reaction was stopped by adding 1 volume of 1 M sodium carbonate after 1 or 3 min of incubation for the free and immobilized enzyme, respectively. The amount of *p*-nitrophenol released was measured by determining the absorption at 420 nm in a 1-cm cuvette. The reaction mixture contained 100 µl of enzyme preparation, 1.25 mM *p*-NPG and 100 mM buffer at various pH values to a final volume of 1 ml. For the immobilized protein three beads of alginate-cell suspension or alginate-enzyme were used instead of 100 µl of free enzyme. For the free enzyme, one unit of α-glucosidase activity was defined as the formation of 1 µmol *p*-nitrophenol per min per mg protein under optimal conditions as specified above.

Temperature and pH effects.

The pH profile of *Shewanella* sp. ANA-3 α-glucosidase was determined at pH values ranging from 3.5 to 9 (3.5 to 5.0 in sodium acetate buffer; 5 to 6.5 in MES buffer; 6.5 to 8.5 in MOPS buffer and 8.5 to 9 in glycine buffer). The temperature effect on the activity was measured after incubation of the enzyme at temperatures between 4 and 65 °C. For thermal stability studies the enzyme was incubated at 45, 50 and 55 °C in 100 mM MES buffer, pH 6.5. The pH stability measurements were carried out after incubation of the enzyme at pH 5.5, 6 and 6.5 in 100 mM MES buffer and pH 7 in 100 mM MOPS buffer at 45 and 50 °C. Samples were withdrawn, at periodic intervals, for determination of the residual activity.

Effect of ions.

In order to determine the effect of ions on the enzyme activity, α-glucosidase was purified and stored in a 100 mM EDTA and 100 mM MES buffer (pH 6.5) solution at 4 °C for 18 h. After dialysis against 100 mM MES-NaOH buffer (milli-Q, pH 6.5) and appropriate

dilution of the enzyme, the activity was measured in presence of 1 mM of various metal ion compounds ($ZnCl_2$, $CaCl_2$, $CaSO_4$, $Ca(CH_3COO)_2$, $MnCl_2$, $FeSO_4$, $CoCl_2$, $CuCl_2$, $MgSO_4$ and $MgCl_2$) and monitored under optimal enzymatic conditions (100 mM MES-milli-Q buffer pH 6.5 at 50 °C) for 1 min. $CaCl_2$ concentrations required for optimal enzyme activity was determined by incubating the enzyme samples at concentrations ranging from 0 to 2 mM.

Substrate specificity.

The substrate specificity was determined under optimal conditions in a reaction mixture containing 10 µl of 100 mM MES buffer (pH 6.5), 8.8 µg of purified enzyme and 1 µl of 50 mM substrate. After 5 min the reaction was stopped by heating at 95°C for 3 min. The activity of the enzyme on natural substrates was analysed by thin-layer chromatography (TLC) on silica gel plates (TLC Aluminium oxide 60 F_{254} neutral, Merck) using 1-propanol-water-ethyl acetate (6:3:1, V/V/V) as the mobile phase. After the separation process, the reaction products were revealed using aniline-diphenylamine (1% [W/V] diphenylamine and 1% [V/V] aniline in acetone mixed with 0.1 volume of 85% [V/V] phosphoric acid prior to use) followed by heating for 5 min at 160 °C.

RESULTS

Cloning, over-expression, purification and sequence analysis

The 2112-bp gene (Uniprot accession number A0KWJ3) corresponds to the single gene of *Shewanella* sp. ANA-3 encoding an α-glucosidase (EC 3.2.1.20). The PCR-amplified fragment of the *Shw*SusB gene was purified and cloned into the pET-15b vector and transformed in the *E. coli* BL21 (DE3) strain. A high number of recombinant functional clones harbouring the α-glucosidase gene were obtained on LB agar and analyzed using restriction enzymes. A clone carrying the gene under control of the T7 promoter was used for liquid culture and α-glucosidase activity was clearly detected in the crude extract. This latter gene encodes a 703 amino acid residue protein with a calculated molecular mass of 79 kDa. The recombinant protein purified by immobilized metal affinity- and anion-exchange chromatographies migrated as a single band with an estimated molecular mass close to 80 kDa on SDS-PAGE which is in agreement with the calculated molecular mass. Size exclusion chromatography showed a single elution peak with a size corresponding to a molecular mass of 80 kDa and to a monomeric enzyme. This was confirmed by chemical cross-linking studies of the native enzyme followed by SDS-PAGE analysis.

Shewanella sp. ANA-3 α-glucosidase has been classified into GH97 in the CAZy database (22). Primary structure alignment indicated a high level of amino acid similarity (36% identity) to *B. thetaiotaomicron* α-glucosidase (*Bt*GH97a) and a lower sequence identity (24%) with the α-galactosidase from *B. thetaiotaomicron* (*Bt*GH97b). Analysis of the primary structure revealed the presence of a HHET motif along with a second motif K(T/S)GY in the *Shewanella* enzyme (Fig. 1) which both have been reported to be characteristic for the GH97a subfamily of inverting enzymes (23).

pH and temperature effects on free and immobilized enzymes

The recombinant α-glucosidase displayed highest activity at 50 °C (Fig. 2B) and retained 50% of its maximal relative activity between 4 °C and 30 °C. At temperatures above 50 °C, the enzyme was rapidly inactivated (Fig. 2B) with a drastic activity loss from 60 °C. This α-glucosidase displays a maximum activity at pH 6.5 with a rapid decrease at both lower and higher pH values (Fig. 2C).

Several studies have demonstrated *improved* enzyme performance after immobilisation *via* enhanced *stability* over a broad *pH and temperature range* (15, 16, 24-26). As preferred matrix, alginate is by far the most widely used polymer (27). We measured the α-glucosidase activity in alginate beads, using both enzyme- and whole cell immobilization techniques. For this latter, the optimal concentration of cells within alginate beads was explored following *p*-NPG hydrolysis. The results shown in Fig. 2A demonstrate maximal activity at cell concentrations between 30 and 36 g/l. Immobilization did not affect the optimal temperature of the enzyme reaction but increased the relative activity at both low and high temperatures within the temperature range from 4 – 65 °C, with a plateau between 50 and 60 °C. At this latter temperature the enzyme recovered 60% of its relative activity (Fig. 2B). Furthermore, more than 60% of the relative activity is retained between pH 5 and 8.5 when immobilizing cells whereas this range is narrower and comprised between pH 6 and 7 for the free enzyme (Fig. 2C). Immobilization of the enzyme was not successful, in that no activity could be measured. Hence all immobilization studies described hereafter have been performed using whole cells, and at a cell concentration of 31 g/l.

Ion dependency

Initially, the metal ion dependence of the enzyme activity was investigated on the crude extract (final concentration of all metal ion solutions 1 mM). As shown in Fig. 3A, the relative activity is enhanced up to 230% by addition of Ca^{2+}. It appears that Mg^{2+} or Mn^{2+}

ions are mild activators with 147% and 122% of the relative activity, respectively. The activity was totally abolished in the presence of Zn^{2+}. Divalent Cu^{2+}, Co^{2+} and Fe^{2+} ions are poor inhibitors with 74%, 87% and 76% of the relative activity, respectively. In a second step, the purified protein was treated with EDTA. The metal depleted enzyme lost most of its activity. As expected, addition of 1 mM magnesium chloride, manganese chloride and especially calcium chloride restores the activity (Fig. 3B). In order to determine whether the increased activity unequivocally stemmed from these divalent ions, further investigations using compounds that were not chloride salts were performed. Differences were clearly observed between addition of magnesium chloride, magnesium sulphate, calcium chloride, calcium sulphate and calcium acetate. The highest activities were obtained with $MgCl_2$ and $CaCl_2$, while the other values were similar to those found for the metal depleted enzyme (Fig. 3C). Varying concentrations of $CaCl_2$ (Fig. 3D) indicated that the enzyme progressively recovers its activity by addition of up to 2 mM $CaCl_2$. The sigmoidal shape of the observed activity curve is indicative of *Shw*SusB being an allosterically activated enzyme.

Enzyme stability

Thermal stability was investigated at 45, 50 and 55 °C using free enzyme and immobilized cells at pH 6.5 over 1 hour of incubation at these temperatures. Residual activity was measured at optimal reaction conditions. Our results indicate that the free enzyme is fully stable at temperatures below 45 °C. At 50 and 55 °C the enzyme stability decreased with half-lives of 60 and 30 min, respectively (Fig. 4A). The enzyme in immobilized cells was stable up to 50 °C for one hour, but was completely inactivated at 55 °C (Fig. 4B). pH stability studies at 45 and 50 °C at pH 5.5, 6, 6.5 and 7 (Fig. 5) indicated a fully stable enzyme at pH 6.5 and 45 °C, regardless of being free (Fig. 5A) or immobilized (Fig. 5B) in agreement with the thermostability study. At this temperature at the other pH values tested, the stability of the free enzyme rapidly decreased with half-lives of 5, 20 and 30 minutes at pH 5.5, 7 and 6, respectively (Fig. 5A). In contrast, when the enzyme was immobilized in cells, the stability was highly enhanced at these pH values with half-lives of 20, 60 and beyond 60 minutes at pH 5.5, 7 and 6, respectively (Fig. 5B). At 50 °C, the free enzyme retained 50% of its activity after 60 minutes incubation at the optimal pH whereas at pH 5.5, 6 and 7, the stability decreased abruptly within 10 minutes (Fig. 5C). When immobilizing cells in alginate beads, the enzyme retained almost all activity at pH 6.5 after 1 hour of incubation. At other pH values, the half-lives were 10, 12 and 25 minutes at pH 5.5, 7 and 6, respectively (Fig. 5D).

Substrate specificity

The activity of recombinant *Shw*SusB on maltooligosaccharides (degree of polymerization, DP, 2 to 7) was examined. Hydrolysis of terminal, non-reducing (1→4)-linked α-D-glucose residues was performed under optimal conditions (pH 6.5 at 50 °C) and in presence or absence of 1 mM $CaCl_2$. Whereas purified *Shw*SusB is slightly active on *p*-NPG, it has no activity on maltooligosaccharides without addition of 1 mM $CaCl_2$ (data not shown). The enzyme hydrolysed all tested maltooligosaccharides as seen in Fig. 6.

DISCUSSION

*Shw*SusB belongs to the intriguing glycoside hydrolase family 97 (GH97), which includes both inverting and retaining α-glycosidases. Despite being a unique GH family, so far only two enzymes have been characterized: an α-glucosidase *Bt*GH97a and an α-galactosidase *Bt*GH97b (23, 28, 29), from *B. thetaiotaomicron*, a symbiotic component of the microflora in the human gut. Multiple sequence alignment analysis highlights the presence of a HHET- and a K(T/S)GY motif which both have been reported to be consensus sequences and characteristic for inverting enzymes within the GH97 family. Size exclusion chromatography and crosslink/SDS-PAGE studies suggest that *Shw*SusB is monomeric in solution as opposed to the *B. thetaiotaomicron* α-glucosidase *Bt*GH97a, which has been shown to be a dimer in solution (28).

Though the role of *Shewanella* sp. ANA-3 α-glucosidase by analogy to its *B. thetaiotaomicron* counterpart, SusB, is to hydrolyse di- and trisaccharides produced by a neopullulanase, SusA (28, 30) (this latter is also present in *Shewanella* sp. ANA-3), all tested maltooligosaccharides (DP2 – DP7) were hydrolysed. The experiments determining metal ion dependency revealed that calcium ions play a key role in the activity. This is in agreement with earlier studies which suggest that the calcium ion in GH97 enzymes provides acid assistance to the catalytic acid during hydrolysis (23, 28). More unexpected the presence of chloride appears to be essential for the enzymatic activity of this cold-adapted enzyme. Both observations are consistent with the natural habitat of *Shewanella* sp. ANA-3 being brackish water with a calcium content of 400 mg/L and a chloride content of 17-18 g/L (31). Moreover, this kind of findings has been reported in earlier studies of enzymes from cold-adapted microorganisms as eg. the *Pseudoalteromonas haloplanktis* α-amylase for which chloride was shown to be an allosteric activator (32-35). The lack of activity on

maltooligosaccharides without addition of 1 mM $CaCl_2$ confirmed the essential role of both calcium- and chloride ions.

The optimal temperature for the recombinant α-glucosidase is significantly above the optimum growth temperature of *Shewanella* sp. ANA-3 or any other *Shewanella* strain (13). Despite having an optimal temperature around 50 °C, *Shw*SusB retains 50% of its activity at 5 °C. To the best of our knowledge, only *Aspergillus niger* α-glucosidase has been reported to display comparable activity at these low temperatures (14). Upon enzyme immobilization in whole cells, the lower relative activities observed at high cell concentrations are probably due to steric hindrance effects and diffusional limitations for the substrate within the matrix (36). Enzyme immobilization using the cell extract was also examined but no activity was detected probably due to leaking of the proteins out of the mesh of the beads as observed in earlier studies (37). Overall, immobilization of the studied enzyme in whole cells resulted in an important improvement of the relative activity at all pH values, suggesting that *Shw*SusB is less sensitive to alterations induced by pH when immobilized than when it is in the free enzyme state, as observed for earlier studied enzymes (36). This feature is more pronounced at alkaline pH values with a shift of the optimum pH from 6.5 to 7. The pH shift could be due to interactions between the enzyme and the polymeric matrix (ionic- and polar interactions, hydrogen bonding…) or due to an altered hydrogen ion concentration in the enzyme microenvironment (38, 39).

The observed maximal activity between 50 and 60 °C for the cell immobilized enzyme could perhaps be attributed to the requirement for a higher activation energy or conformational change(s) due to the charge and physical restraints of the matrix as reported in earlier studies (40). Interestingly, a reduction in the activity at low temperatures as shown in other studies (38, 40) was not observed for *Shw*SusB α-glucosidase. The loss in enzyme flexibility at lower temperatures may be offset by a high catalytic efficiency of the recombinant enzyme at temperatures below 30 °C. A direct link exists between the activity and stability and psychrophilic enzymes are known to be structurally flexible leading to thermolability (5, 33). Also at high temperatures, loss in α-glucosidase activity by the non-immobilized enzyme was clearly due to a progressive denaturation of enzyme caused by the unfolding of the enzyme (41). The improved stability of immobilized enzyme in whole cells is probably assisted by the matrix, which absorbs a part of the heat and protects the enzyme against denaturation. Thereby, enzyme entrapment in alginate beads diminishes unfolding by reducing the flexibility of the macromolecule, thus resulting in a gain of stability at different

temperatures. Our studies showed that whole cell immobilization in alginate beads results in an enhancement of the thermostability at temperatures up to 50 °C. The rapid decrease in activity observed at 55 °C is possibly due to important bead damage at this temperature.

In conclusion, our biochemical studies reveal interesting features, as activity at low temperatures and calcium- and chloride dependence of this enzyme, being in perfect agreement with the environmental parameters in which the *Shewanella* sp. ANA-3 has been isolated.

ACKOWLEDGEMENTS

This work was supported by the CNRS (Centre National de la Recherche Scientifique). We acknowledge Dr. Chad SALTIKOV from California Institute of Technology who generously provided the *Shewanella* sp. ANA-3 strain. Rimeh Ilhammami received financial support from the University of Gabes/Higher Institute of Applied Biotechnology of Mednine in Tunisia.

REFERENCES

1. **Gerday C, Aittaleb M, Arpigny JL, Baise E, Chessa JP, Garsoux G, Petrescu I, Feller G.** 1997. Psychrophilic enzymes: a thermodynamic challenge. Biochim Biophys Acta. **1342:**119-31.
2. **Fredrickson JK, Romine MF, Beliaev AS, Auchtung JM, Driscoll ME, Gardner TS, Nealson KH, Osterman AL, Pinchuk G, Reed JL, Rodionov DA, Rodrigues JL, Saffarini DA, Serres MH, Spormann AM, Zhulin IB, Tiedje JM.** 2008. Towards environmental systems biology of *Shewanella*. Nat Rev Microbiol. **6:**592-603.
3. **Hau HH, Gralnick JA.** 2007. Ecology and biotechnology of the genus *Shewanella*. Annu Rev Microbiol. **61:**237-58.
4. **Struvay C, Feller G.** 2012. Optimization to low temperature activity in psychrophilic enzymes. Int J Mol Sci. **13:**11643-65.
5. **Georlette D, Blaise V, Collins T, D'Amico S, Gratia E, Hoyoux A, Marx JC, Sonan G, Feller G, Gerday C.** 2004. Some like it cold: biocatalysis at low temperatures. FEMS Microbiol Rev. **28:**25-42.
6. **Paredes DI, Watters K, Pitman DJ, Bystroff C, Dordick JS.** 2011. Comparative void-volume analysis of psychrophilic and mesophilic enzymes: Structural bioinformatics of psychrophilic enzymes reveals sources of core flexibility. BMC Struct Biol. **11:**42.
7. **Rodionov DA, Yang C, Li X, Rodionova IA, Wang Y, Obraztsova AY, Zagnitko OP, Overbeek R, Romine MF, Reed S, Fredrickson JK, Nealson KH, Osterman AL.** 2010. Genomic encyclopedia of sugar utilization pathways in the *Shewanella* genus. BMC Genomics. **11:**494.
8. **Malasarn D, Keeffe JR, Newman DK.** 2008. Characterization of the arsenate respiratory reductase from *Shewanella* sp. strain ANA-3. J Bacteriol. **190:**135-42.
9. **Karpinets TV, Romine MF, Schmoyer DD, Kora GH, Syed MH, Leuze MR, Serres MH, Park BH, Samatova NF, Uberbacher EC.** 2010. Shewanella knowledgebase: integration of the experimental data and computational predictions suggests a biological role for transcription of intergenic regions. Database (Oxford). **2010:**baq012.
10. **Huang JH, Voegelin A, Pombo SA, Lazzaro A, Zeyer J, Kretzschmar R.** 2011. Influence of arsenate adsorption to ferrihydrite, goethite, and boehmite on the kinetics

of arsenate reduction by *Shewanella putrefaciens* strain CN-32. Environ Sci Technol. **45:**7701-9.

11. **Murphy JN, Durbin KJ, Saltikov CW.** 2009. Functional roles of arcA, etrA, cyclic AMP (cAMP)-cAMP receptor protein, and cya in the arsenate respiration pathway in *Shewanella* sp. strain ANA-3. J Bacteriol. **191:**1035-43.
12. **Janda JM, Abbott SL.** 2012. The genus Shewanella: from the briny depths below to human pathogen. Crit Rev Microbiol.
13. **Karpinets TV, Obraztsova AY, Wang Y, Schmoyer DD, Kora GH, Park BH, Serres MH, Romine MF, Land ML, Kothe TB, Fredrickson JK, Nealson KH, Uberbacher EC.** 2010. Conserved synteny at the protein family level reveals genes underlying *Shewanella* species cold tolerance and predicts their novel phenotypes. Funct Integr Genomics. **10:**97-110.
14. **Brizova K, Kralova B, Demnerova K, Vins I.** 1992. Isolation and characterization of alpha-glucosidase from *Aspergillus niger*. J Chromatogr. **593:**125-31.
15. **Chen W, Chen H, Xia Y, Yang J, Zhao J, Tian F, Zhang HP, Zhang H.** 2009. Immobilization of recombinant thermostable beta-galactosidase from *Bacillus stearothermophilus* for lactose hydrolysis in milk. J Dairy Sci. **92:**491-8.
16. **Rhimi M, Boisson A, Dejob M, Boudebouze S, Maguin E, Haser R, Aghajari N.** 2010. Efficient bioconversion of lactose in milk and whey: immobilization and biochemical characterization of a beta-galactosidase from the dairy *Streptococcus thermophilus* LMD9 strain. Res. Microbiol. **161:**515-25.
17. **Sambrook J, Fritsch EF, Maniatis T.** 1989. Molecular Cloning: a laboratory manual, vol. 1,2,3, New York.
18. **Kabsch W, Sander C.** 1984. On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. Proc Natl Acad Sci U S A. **81:**1075-8.
19. **Thompson JD, Higgins DG, Gibson TJ.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673-80.
20. **Gouet P, Robert X, Courcelle E.** 2003. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. **31:**3320-3.

21. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. **227:**680-5.

22. **Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. **37:**D233-8.

23. **Gloster TM, Turkenburg JP, Potts JR, Henrissat B, Davies GJ.** 2008. Divergence of catalytic mechanism within a glycosidase family provides insight into evolution of carbohydrate metabolism by human gut flora. Chem Biol. **15:**1058-67.

24. **Sharma S, Gupta MN.** 2001. Alginate as a macroaffinity ligand and an additive for enhanced activity and thermostability of lipases. Biotechnol Appl Biochem. **33:**161-5.

25. **Matsumoto M, Ohashi K.** 2003. Effect of immobilization on thermostability of lipase from *Candida rugosa*. Biochemical Engineering Journal. **14:**75-77.

26. **Cao L.** 2011. Immobilized Enzymes, p. 461-476. *In* M Moo-Young (ed.), Comprehensive Biotechnology (Second Edition), vol. 2, Burlington.

27. **Taqieddin E, Amiji M.** 2004. Enzyme immobilization in novel alginate-chitosan core-shell microcapsules. Biomaterials. **25:**1937-45.

28. **Kitamura M, Okuyama M, Tanzawa F, Mori H, Kitago Y, Watanabe N, Kimura A, Tanaka I, Yao M.** 2008. Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*. J Biol Chem. **283:**36328-37.

29. **Okuyama M, Kitamura M, Hondoh H, Kang MS, Mori H, Kimura A, Tanaka I, Yao M.** 2009. Catalytic mechanism of retaining alpha-galactosidase belonging to glycoside hydrolase family 97. J Mol Biol. **392:**1232-41.

30. **D'Elia JN, Salyers AA.** 1996. Contribution of a neopullulanase, a pullulanase, and an alpha-glucosidase to growth of *Bacteroides thetaiotaomicron* on starch. J Bacteriol. **178:**7173-9.

31. **Page IH.** 1928. Further observations on the chemical composition of Woods Hole sea water - The chlorine content and salt analysis. The Biological Bulletin. **55:**449-452.

32. **Feller G, Payan F, Theys F, Qian M, Haser R, Gerday C.** 1994. Stability and structural analysis of alpha-amylase from the antarctic psychrophile *Alteromonas haloplanctis* A23. Eur J Biochem. **222:**441-7.

33. **Feller G, Narinx E, Arpigny JL, Aittaleb M, Baise E, Genicot S, Gerday C.** 1996. Enzymes from psychrophilic organisms. FEMS Microbiology Reviews. **18:**189-202.

34. **Aghajari N, Feller G, Gerday C, Haser R.** 1998. Structures of the psychrophilic *Alteromonas haloplanctis* alpha-amylase give insights into cold adaptation at a molecular level. Structure. **6:**1503-16.

35. **Aghajari N, Feller G, Gerday C, Haser R.** 2002. Structural basis of alpha-amylase activation by chloride. Protein Sci. **11:**1435-41.

36. **Le-Tien C, Millette M, Lacroix M, Mateescu MA.** 2004. Modified alginate matrices for the immobilization of bioactive agents. Biotechnol Appl Biochem. **39:**189-98.

37. **Tanaka H, Kurosawa H, Kokufuta E, Veliky IA.** 1984. Preparation of immobilized glucomylase using ca-alginate gel coated with partially quaterized poly(ethleneimine). Biotechnol Bioeng. **26:**1393-4.

38. **Munjal N, Sawhney SK.** 2002. Stability and properties of mushroom tyrosinase entrapped in alginate, polyacrylamide and gelatin gels. Enzyme and Microbial Technology. **30:**613-619.

39. **Kumar S, Dwevedi A, Kayastha AM.** 2009. Immobilization of soybean (Glycine max) urease on alginate and chitosan beads showing improved stability: Analytical applications. Journal of Molecular Catalysis B: Enzymatic. **58:**138-145.

40. **Puri M, Marwaha SS, Kothari RM.** 1996. Studies on the applicability of alginate-entrapped naringinase for the debittering of kinnow juice. Enzyme and Microbial Technology. **18:**281-285.

41. **Iyer PV, Ananthanarayan L.** 2008. Enzyme stability and stabilization - Aqueous and non-aqueous environment. Process Biochemistry. **43:**1019–1032.

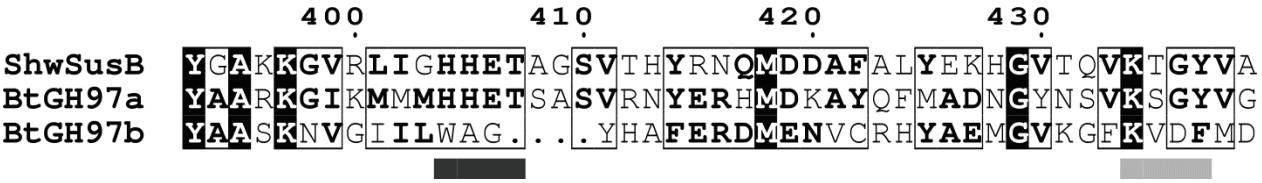


Figure 1. Partial sequence alignment of α-glucosidase from *Shewanella* sp. ANA-3 (*Shw*SusB), with *Bacteroides thetaiotaomicron* α-glucosidase (*Bt*GH97a) and *Bacteroides thetaiotaomicron* α-galactosidase (*Bt*GH97b). The amino acid numbering on the top of the alignment corresponds to the *Shewanella* enzyme and the HHET- and K(T/S)GY motifs found in GH97 members proceeding with inversion of stereochemistry is underlined in dark grey and light grey, respectively.

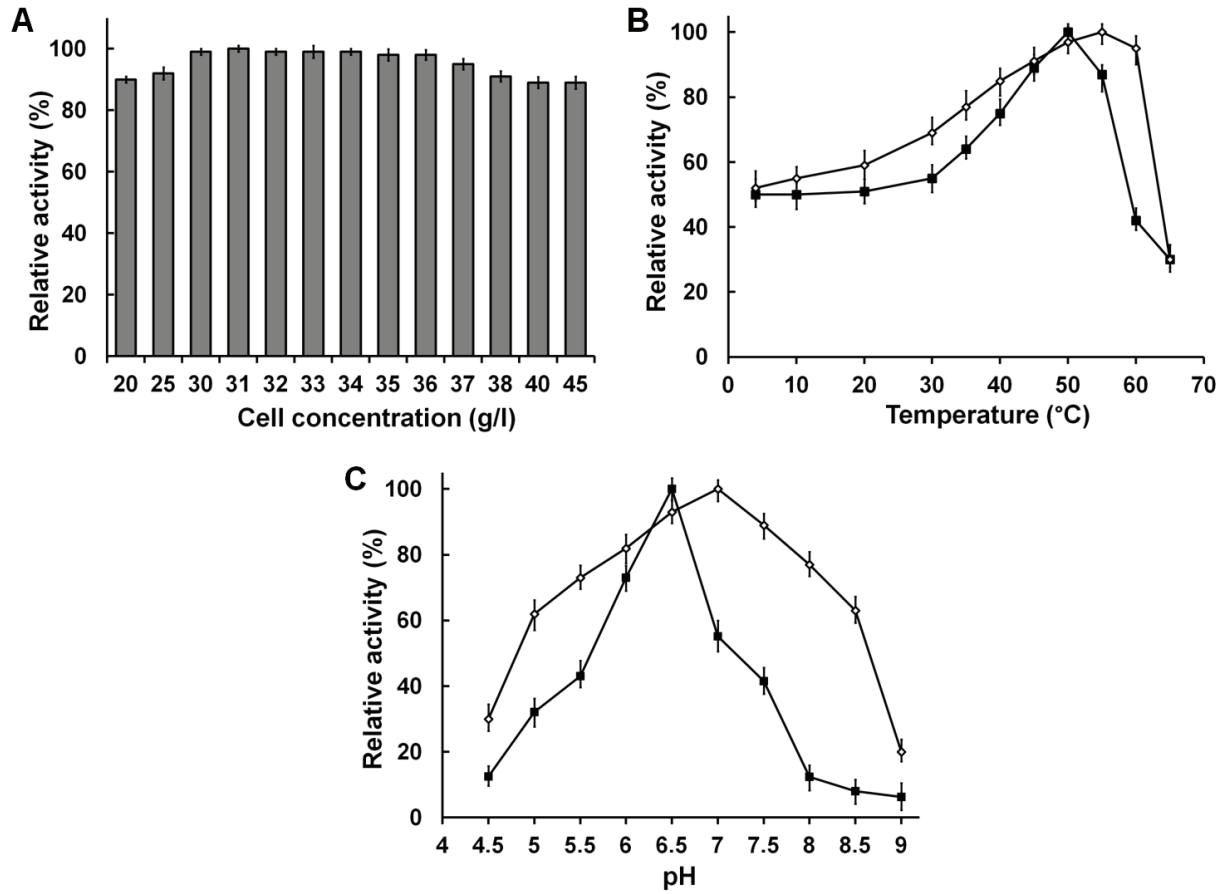


Figure 2. (A) Cell concentrations in alginate beads. (B) Temperature activity profiles of the free enzyme (bold squares) and whole cell immobilization (open rhombus). (C) pH activity profiles of the free enzyme (bold squares) and whole cell immobilization (open rhombus). Activities at optimal cell concentration, temperature and pH were defined as 100%. Errors bars represent the standard deviation from three independent experiments. For the free enzyme, 100% activity corresponds to 3125 U/mg.

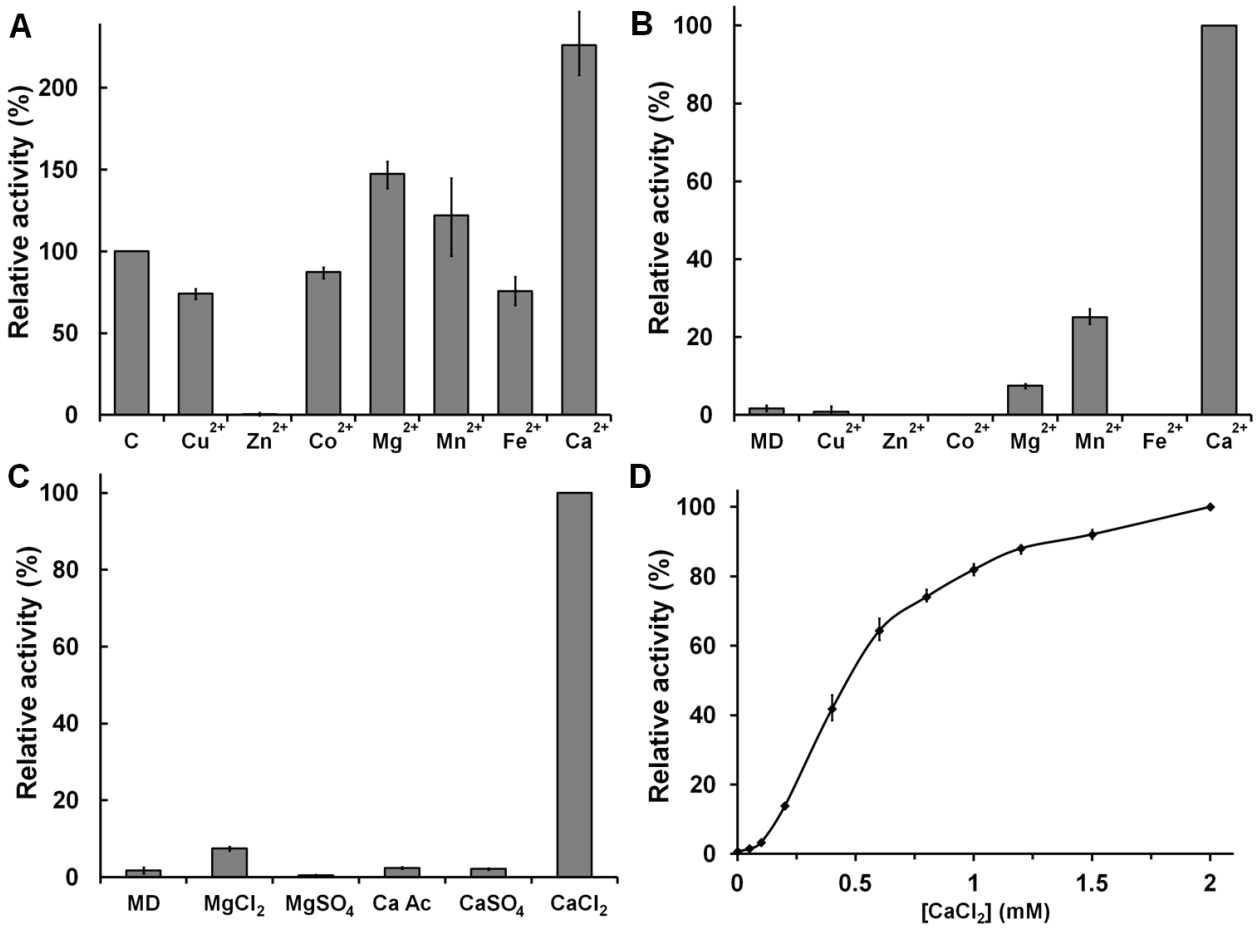


Figure 3. (A) Effect of divalent metal ions on the enzyme in the crude extract. "C" indicates enzyme activity measured in the crude extract without addition of metal ions and was defined as 100%. (B) Effect of divalent metal ions on the metal depleted (MD) purified enzyme. Activity in presence of calcium ions was defined as 100%. (C) Effect of different anions on the metal depleted (MD) purified enzyme. "Ca Ac" indicates calcium acetate. (D) Effect of various concentrations of $CaCl_2$. Activity at 2 mM was defined as 100%. Error bars represent the standard deviation from three independent experiments.

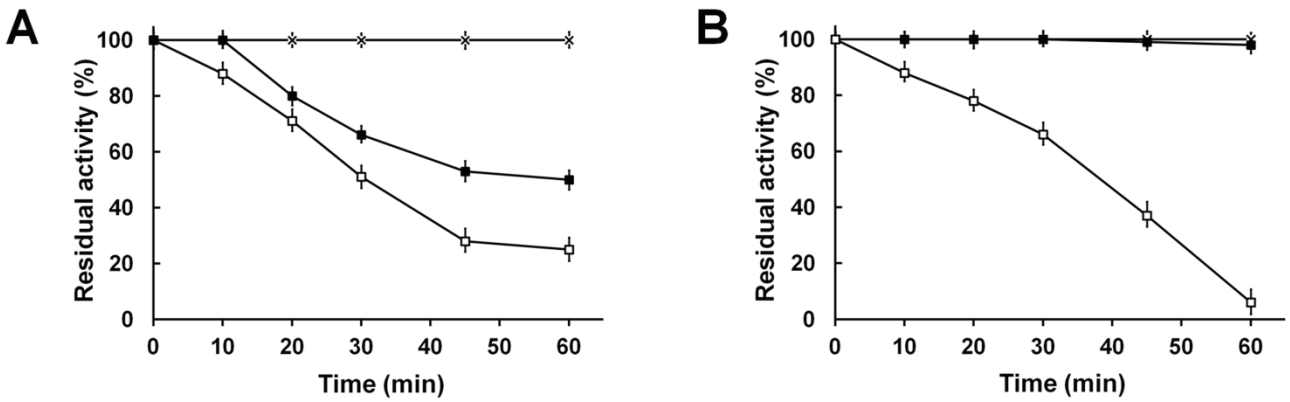


Figure 4. Thermostability profiles of the free enzyme (A) and whole cell immobilization (B). Crosses: 45 °C, bold squares: 50 °C, open squares: 55 °C. The initial activity was defined as 100%. Error bars represent the standard deviation from three independent experiments. For the free enzyme, 100% activity corresponds to 3125 U/mg.

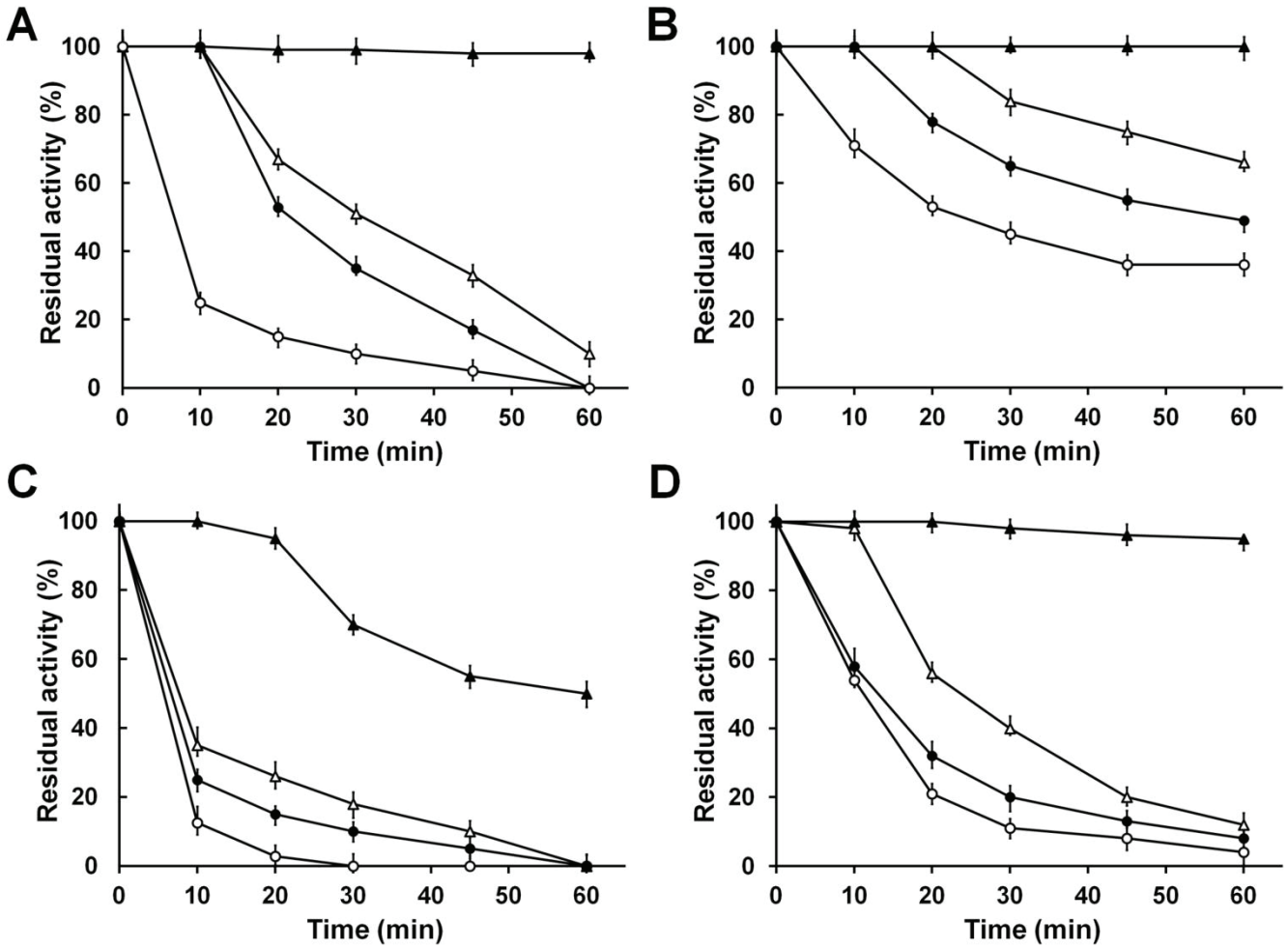


Figure 5. pH stability profiles of the free enzyme (A) and enzyme immobilized in whole cells (B) at 45 °C and of the free (C) and immobilized (D) enzyme at 50 °C. Open circles: pH 5.5; Open triangles: pH 6; bold triangles: pH 6.5; bold circles: pH 7. The initial activity was defined as 100%. Error bars represent the standard deviation from three independent experiments. For the free enzyme, 100% activity corresponds to 3125 U/mg.

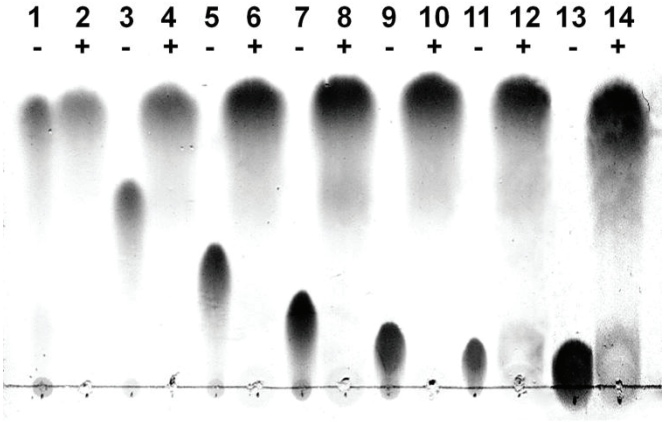


Figure 6. Thin-layer chromatography analysis of the reaction products of *Shw*SusB. The reactions were carried out in 12 µl of 100 mM MES and 1 mM Ca^{2+} buffer (pH 6.5) with 4 mM substrate and with (+) or without (-) 8.8 µg of purified enzyme at 50 °C. Lanes 1 and 2: glucose, lanes 3 and 4: maltose, lanes 5 and 6: maltotriose, lanes 7 and 8: maltotetraose, lanes 9 and 10: maltopentaose, lanes 11 and 12: maltohexaose, lanes 13 and 14: maltoheptaose.

INSTITUT PASTEUR

Research in Microbiology 161 (2010) 515–525

www.elsevier.com/locate/resmic

Efficient bioconversion of lactose in milk and whey: immobilization and biochemical characterization of a β-galactosidase from the dairy *Streptococcus thermophilus* LMD9 strain

Moez Rhimi [a], Anais Boisson [a], Magali Dejob [a], Samira Boudebouze [b], Emmanuelle Maguin [b], Richard Haser [a], Nushin Aghajari [a,*]

[a] *Laboratoire de BioCristallographie, Institut de Biologie et Chimie des Protéines, UMR 5086-CNRS/Université de Lyon, IFR128 "BioSciences Gerland – Lyon Sud", 7 Passage du Vercors, F-69367 Lyon cedex 07, France*

[b] *INRA, UMR 1319 Micalis, F-78350 Jouy-en-Josas, France*

Received 15 February 2010; accepted 29 April 2010
Available online 28 May 2010

Abstract

The gene encoding β-galactosidase from dairy *Streptococcus thermophilus* strain LMD9 was cloned, sequenced and expressed in *Escherichia coli*. The recombinant enzyme was purified and showed high specific activity of 464 U/mg. This protein displays a homotetrameric arrangement composed of four 118 kDa monomers. Monitoring of the activity showed that this enzyme was optimally active at a wide range of temperatures (25–40 °C) and at pH from 6.5 to 7.5. Immobilization of the recombinant *E. coli* in alginate beads clearly enhanced the enzyme activity at various temperatures, including 4 and 50 °C, and at pH values from 4.0 to 8.5. Stability studies indicated that this biocatalyst has high stability within a broad range of temperatures and pH. This stability was improved not only by addition of 1 mM of Mn^{2+} and 1.2 mM Mg^{2+}, but essentially through immobilization. The remarkable bioconversion rates of lactose in milk and whey at different temperatures revealed the attractive catalytic efficiency of this enzyme, thus promoting its use for lactose hydrolysis in milk and other dairy products.
© 2010 Elsevier Masson SAS. All rights reserved.

Keywords: β-galactosidase; Probiotic; Immobilization; Milk; Whey; Lactose hydrolysis

1. Introduction

Strong industrial interests are ascribed to β-galactosidases due to their hydrolase and transferase activities (Mahoney, 1998). As a consequence of their transglycosylation activity, these enzymes are widely used for production of oligosaccharides, some of which are very promising prebiotic agents (Coulier et al., 2009; Rycroft et al., 2001). Furthermore, via their hydrolytic activity, β-galactosidases are widely employed in the agrofood industry, essentially for reducing the lactose concentration in milk products with the aim of overcoming lactose intolerance (Scrimshaw and Murray, 1988). Lactose cleavage in milk represents the first probiotic effect conferred upon *Streptococcus thermophilus* and *Lactobacillus bulgaricus*, used routinely for yoghurt production (Guarner et al., 2005). Several reports described the molecular and biochemical characterization of β-galactosidases from lactic acid bacteria (Barrett et al., 2004; Ramana Rao and Dutta, 1981; Schmidt et al., 1989). Thermo-active and cold-adapted β-galactosidases from various bacteria such as *Bacillus licheniformis*, *Lactobacillus acidophilus* R22, *Geobacillus stearothermophilus*, *Sulfolobus solfataricus*, *Enterobacter agglomerans* B1, *Thermotoga maritima*, *Alicyclobacillus acidocaldarius* subsp. *Rittmannii*, *S. thermophilus* and *Arthrobacter psychrolactophilus* strain F2 were described as suitable biocatalysts for many industrial applications (Kim et al., 2004; Lu et al., 2007; Nakagawa et al., 2006; Nguyen et al., 2007a; Pisani et al., 1990; Phan Tran et al., 1998;

* Corresponding author. Tel.: +33 (0) 4 72 72 26 33; fax: +33 (0) 4 72 72 26 16.
E-mail address: n.aghajari@ibcp.fr (N. Aghajari).

0923-2508/$ - see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.resmic.2010.04.011

Ramana Rao and Dutta, 1981). Immobilization of β-galactosidases using different methods, for industrial purposes has also been reported (Hinberg et al., 1974; Heng and Glatz, 1994). Indeed, immobilization improves enzyme thermoactivity and thermostability and reduces glucose inhibition (Mateo et al., 2004; Chen et al., 2009; Elnashar and Yassin, 2008). Use of immobilized β-galactosidases for lactose hydrolysis in milk/whey as an efficient bioconverter on an industrial scale was reported (Haider and Husain, 2007, 2009a,b,c). For hydrolysis of lactose in whey and many fermented milk products, it would be of interest to determine a β-galactosidase efficient within a broad temperature range and at acidic pH values (Robinson, 1994). This would allow lactose bioconversion not only during the milk pasteurization step and storage of the products, but also during fermentation of milk. Whey is an inexpensive raw material with a pH between 5.5 and 6.5 containing a high amount of lactose, an attractive source of low calorie sweetener production (Rhimi et al., 2007). Hence, for efficient hydrolysis of the extracted lactose, the use of a β-galactosidase active at acidic pH would be particularly suitable.

Here we report the cloning, expression, purification and biochemical characterization of a β-galactosidase from the probiotic and food-grade *S. thermophilus* LMD9 strain. We also investigate the effect of immobilization of the enzyme on biochemical properties. The efficiency of free and immobilized enzymes on lactose hydrolysis in milk and whey is reported.

2. Materials and methods

2.1. Bacterial strains, plasmids and media

Growth of the *S. thermophilus* LMD9 (*S. thermophilus*) strain under aerobic conditions at 42 °C in M17 medium (Difco) with 1% (w/v) glucose was previously described (Makarova et al., 2006; Terzaghi and Sandine, 1975). *Escherichia coli* NovaBlue (DE3) was used as host strain. Culture of the *E. coli* strain was done in Luria Bertani (LB) medium (Sambrook et al., 1989) supplemented, when required, with ampicillin (100 μg/ml), X-gal (100 μg/ml) and IPTG (isopropyl-β-D-thiogalactopyranoside) at 160 μg/ml. The pGEMT-Easy Vector plasmid (Promega) was used following the manufacturer's instructions.

2.2. DNA manipulation and PCR

S. thermophilus LMD9 total DNA was prepared as described by Pospiech and Neumann (1995). Preparation of plasmid DNA was carried out using the Gene JET™ Plasmid Miniprep kit (Fermentas). DNA digestion with restriction endonucleases and separation of fragments by agarose gel electrophoresis were performed as described (Sambrook et al., 1989). Polymerase chain reactions were carried out in a Gene Amp® PCR System 9700 (Applied Biosystems). The amplification reaction mixtures (100 μl) contained Phusion™ High-Fidelity DNA polymerase amplification buffer, 10 pmol of each primer, 100 ng of DNA template, and 2 units of Phusion HF enzyme (Finnzymes). Cycling parameters were 94 °C for 5 min, followed by 35 cycles at 94 °C for 30 s, at 55 °C for 60 s, and 72 °C for 180 s and finally 15 min at 72 °C.

2.3. Cloning, overexpression and sequence analysis of the S. thermophilus *LMD9* lacZ *gene*

Based on the DNA sequence of *S. thermophilus* LMD9 β-galactosidase reported in the NCBI database (accession number YP_820735), two oligonucleotides, BGLMD9-D (5′GTTGGAGAAAAGGTAGAAGCTTTGTCTGAAGTTAT-TACCTTC3′) and BGLMD9-R (5′ACTAGATTCTTAATAAC TACCTAATTTAGTGGTTCAATCATG3′), were conceived. Chromosomal DNA isolated from *S. thermophilus* LMD9 was used as template and Phusion HF DNA polymerase for amplification. The resulting PCR product, with the expected size of nearly 3 kb, was purified using the QIAquick® Gel Extraction kit (Qiagen), following the manufacturer's instructions. The obtained fragment was treated with Taq DNA polymerase in the presence of 0.2 mM of dATP as recommended by the suppliers (Promega). Then, the resulting fragment was cloned in the pGEMT-Easy Vector (Promega) and transformed into *E. coli* NovaBlue (DE3) competent cells with the genotype: *endA1 hsdR17*(*r*K122*m*K121) *supE44 thi-1 recA1 gyrA96 relA1lac* [F9 *proA1B1 lacI*qZD*M15*::Tn*10* (Tcr)] (Novagen). The recombinant strains were selected by enzymatic restriction. The strain (MRS35) harboring the *S. thermophilus lacZ* gene under the control of the PT7 promoter was used for activity assays and grew at 37 °C to an OD_{600} of 1.0 followed by induction with IPTG (160 μg/ml). Subsequently, protein expression was allowed to continue overnight in culture at 37 °C. Notably, the amplified fragment contains the *lacZ* gene encoding full-length *S. thermophilus* LMD9 β-galactosidase and its RBS (ribosome binding site) region. The *lacZ* gene sequence was determined using an automated 3100 Genetic Analyzer (Applied Biosystems). Protein multiple sequence alignments were done using the CLUSTAL W program (Thompson et al., 1994).

2.4. Crude cell lysate preparation and enzyme purification

MRS35 cells were harvested by centrifugation (7500 g, 10 min at 4 °C) and the pellets were suspended in 100 mM HEPES buffer pH 6.5. Then, cell suspensions were incubated for 2 h on ice in the presence of 5 mg/ml lysozyme supplemented with one tablet of Complete™ EDTA-free protease inhibitor (Roche Applied Science) per 50 ml. Cell disruption was achieved by sonication at 4 °C for 6 min (pulsation of 3 s, Amplify-90) using a Vibra-Cell™ 72405 sonicator and cell debris were removed by centrifugation (30,000×*g*, for 30 min at 4 °C). For purification, crude cell extract from the MRS35 *E. coli* strain was heat-treated (55 °C, 60 min) followed by centrifugation at 30,000×*g*, for 30 min at 4 °C. Proteins were precipitated by ammonium sulfate saturation (40–50%), suspended in 100 mM HEPES buffer pH 6.5, concentrated and

desalted in centrifugal microconcentrators (Amicon, Inc) with a 50 kDa membrane cut-off. In a second purification step, size exclusion chromatography was performed (flow rate of 0.5 ml/min) using a Superdex 200 column (10/300GL, GE Healthcare) and 100 mM HEPES buffer at pH 6.5. Purification to homogeneity was completed by fast-performance liquid chromatography (FPLC-Amersham-Bioscience) using a Mono-Q anion-exchange column (Mono-Q™ 5/50GL, GE Healthcare) equilibrated with the same 100 mM HEPES buffer. The proteins were eluted at a flow rate of 2 ml/min using a linear NaCl gradient ranging from 0 to 0.5 M in the HEPES buffer. The pooled fractions exhibiting β-galactosidase activity were desalted and concentrated.

2.5. Protein quantification, electrophoresis, molecular mass determination and MALDI-TOF MS analysis

The protein concentration was determined using Bradford's method Bradford (1976) with bovine serum albumin as standard. Protein samples were separated by SDS-PAGE (12%) according to the method of Laemmli (1970), and visualized by Coomassie brilliant blue R-250 (Biorad) staining. The molecular mass of the purified *S. thermophilus* LMD9 β-galactosidase was confirmed by size exclusion HPLC using a Bio-Sil SEC 250-5 column (30 cm × 0.78 cm). Proteins were separated by isocratic elution at a flow rate of 0.5 ml/min with a 100 mM HEPES buffer (pH 6.5) containing 150 mM NaCl and detected by an UV–VIS spectrophotometer (SHIMADZU) at 280 nm.

An Applied Biosystems Voyager-DE™ PRO MALDI-TOF mass spectrometer was used in the positive linear mode. Spectra were acquired in the 5–100 kDa mass range at a 25 kV accelerating voltage and with a 750 ns delay time. The matrix solution consisted of 1 g/l sinapinic acid (LaserBio Labs) dissolved in a water/acetonitrile mixture (1 part each) containing 1 mL/L TFA. The dried droplet technique was employed for depositing the sample on the target.

2.6. Preparation of cells and immobilization

Cell immobilization was carried out by mixing 10 ml of recombinant *E. coli* strain overexpressing the *S. thermophilus* β-galactosidase with 10 ml of 6.0% (w/v) sodium alginate (Fluka) to obtain an alginate–cell suspension with a cell density of 3×10^5 cell ml^{-1} of alginate. Then, using a syringe the obtained mixture was dropped slowly into a cooled and stirred 200 mM calcium chloride ($CaCl_2$) solution (200 ml). The integrity of the beads was improved by renewing the $CaCl_2$ bath for an overnight incubation at 4 °C. Hereafter, beads were washed twice in 100 mM HEPES buffer (pH 6.5) at 4 °C for 1 h. The average bead size was approximately 0.7 mm.

2.7. Enzyme assays

β-Galactosidase activity was detected by measuring the production of D-glucose or *o*-nitrophenol under the following standard conditions. The reaction mixture contained 50 µl of enzyme preparation at 1 mg/ml, 100 mM of lactose or 3 mM of ONPG and 100 mM HEPES buffer (pH 6.5) to bring the final volume to 1 ml.

Activity in immobilized cell preparations was determined in 100 mg of alginate beads suspended in 100 mM HEPES buffer at pH 6.5. Hydrolysis was performed in 2 ml reaction mixture under the above described assay conditions for 10 min.

For *o*-nitrophenol, the reaction was stopped with 0.6 M sodium carbonate and the absorbance was measured at 420 nm. For lactose, the reaction was stopped by heating the reaction mixture to 99 °C for 5 min and the amount of produced D-glucose was determined colorimetrically using the GOD/POD assay (Kunst et al., 1988). Cleavage of lactose was confirmed by HPLC using an HPAEC-PAD column (Dionex Co.). The products were separated by isocratic elution with water at a flow rate of 0.3 ml/min and detected with a refractive index detector (SHIMADZU, RID-10A). Solutions of D-galactose, D-glucose and lactose at 5 g/l each were used as standards.

One unit of β-galactosidase activity was defined as the amount of enzyme catalyzing the formation of 1 µmol D-glucose or *o*-nitrophenol per min under the above specified conditions.

The conversion rate represents the ratio between the formed D-glucose (equivalent to D-galactose) concentration and the initial lactose concentration.

The productivity represents the concentration of the formed D-glucose (equivalent to D-galactose) per h.

2.8. Effect of cell concentration, temperature and pH

To optimize the operating conditions, the effect of cell concentrations, temperature and pH were investigated. The cell concentration inside the alginate capsules varied between 20 and 45 g/l and the activity was measured. The effect of temperature on activity was measured by incubating free and immobilized enzymes, respectively, at temperatures between 4 and 65 °C. The *S. thermophilus* LMD9 β-galactosidase pH profile was obtained by measuring the activity at various pH values ranging from 4.0 to 8.5 (4.0–6.0 in acetate buffer and 6.5–8.5 in HEPES buffer). Measurement of thermal and pH stabilities of the β-galactosidase was carried out after incubation of the free and immobilized enzymes at temperatures comprised between 4 and 60 °C and pH values between 4.5 and 7.5. Samples were withdrawn at periodic intervals for determination of residual activity under standard conditions.

2.9. Effect of divalent ions on enzyme activity

β-galactosidase was purified as described above, but without addition of metal ions to the culture media or buffer, and dialyzed against 100 mM HEPES buffer (pH 6.5) containing 10 mM EDTA for 48 h at 4 °C. Then the enzyme was incubated in a solution containing divalent ions (1 mM) for 30 min followed by monitoring of β-galactosidase activity.

Determination of concentrations of metal ions required for optimal enzyme activity was achieved by incubating the enzyme samples at various divalent metal ion concentrations.

2.10. Determination of kinetic parameters

Kinetic parameters were determined from Lineweaver–Burk plots. Assays were carried out in 100 mM HEPES buffer (pH 6.5), 1.2 mM $MgCl_2$, 1 mM $MnCl_2$ and 0.05–80 mM ONPG or 0.5–200 mM lactose. The enzyme reaction was performed under standard conditions and stopped by heating to 99 °C. The amount of released D-glucose or *o*-nitrophenol was determined as described above.

2.11. Milk and whey hydrolysis under batch conditions

In the case of whey, milk was skimmed by centrifuging at 8000 g for 20 min and treated with HCl until the pH reached 4.8 (Haider and Husain, 2009a). Casein was removed by centrifugation and the obtained whey was stored at 4 °C for subsequent use as earlier described (Roy and Gupta, 2003). In the batch process, milk or whey solutions (100 ml) were fed into the bioreactor in the presence of immobilized beads (300 mg). Each mixture was stirred continuously in a water circulator at 4 °C and 45 °C for milk and whey, respectively. Samples (300 µl) were then withdrawn at periodic intervals for 6 h and lactose hydrolysis rates were determined as previously described in the enzyme assay section.

2.12. Operating in a packed-bed bioreactor

Beads were packed in a thermostat column, with a diameter/height of 2.7/24 cm and used for lactose hydrolysis. Whey or milk was fed into the bioreactor at a dilution rate of 1.2 h^{-1} using a peristaltic pump (Gilson). D-Glucose and D-galactose products were monitored from the effluent. The reaction temperature was maintained at 4 and 45 °C using a water circulator. In order to optimize lactose hydrolysis operating conditions, the effect of the dilution rate on the conversion level and productivity was explored.

3. Results and discussion

3.1. Molecular cloning, expression, sequence analysis and purification

A PCR-amplified fragment of nearly 3 kb corresponding to the *S. thermophilus* LMD9 *lacZ* gene was purified and cloned in the pGEMT-Easy vector and transformed in the *E. coli* BL21 strain. Many recombinant functional clones harboring the *lacZ* gene were obtained on LB agar medium containing X-gal and analyzed using restriction enzymes. A clone (pMR35) carrying the *lacZ* gene under control of the T7 promoter was used for liquid culture and β-galactosidase activity was monitored in the crude extract. The Novablue/pMR35 strain (MRS35) had clearly detectable activity evaluated to 9.6 U/mg, proving the cloning and expression of the *S. thermophilus* LMD9 *lacZ* gene. The latter gene encodes a 1026 amino acid residue protein with a calculated molecular mass of 117.04 kDa. It should be noted that several reports described the molecular cloning of genes encoding for β-galactosidases from *S. thermophilus* strains along with their expression in different host strains including *E. coli* and *Bacillus subtilus* (Herman and McKay, 1986; Molotov et al., 1990; Ramana Rao and Dutta, 1981).

The primary structure of *S. thermophilus* LMD9 β-galactosidase showed significant identities with β-galactosidases from *S. thermophilus* LMG18311 (99%), *Lactobacillus salivarius* (94%), *Bacillus halodurans* C-125 (48%), *L. bulgaricus* ATCC 11842 (48%), *L. acidophilus* (45%), *Bifidobacterium adolenscentis* ATCC15703 (44%) and *T. maritima* MSB8 (40%).

The protein was purified to homogeneity and size exclusion chromatography showed a unique elution peak with a size corresponding to a molecular weight of 470 kDa (Fig. 1B). SDS-PAGE revealed a single band at a molecular weight of ~118 kDa (Fig. 1A), which was analyzed by MALDI-TOF MS and shown to correspond to *S. thermophilus* LMD9 β-galactosidase (data not shown). These results support the notion that, as expected, *S. thermophilus* LMD9 β-galactosidase is a tetramer composed of four identical subunits. This homotetrameric arrangement is similar to previously characterized β-galactosidases from *T. maritima*, *L. bulgaricus* ATCC 11842, *E. coli* and *Pseudomonas* sp. 22b (Coker et al.,

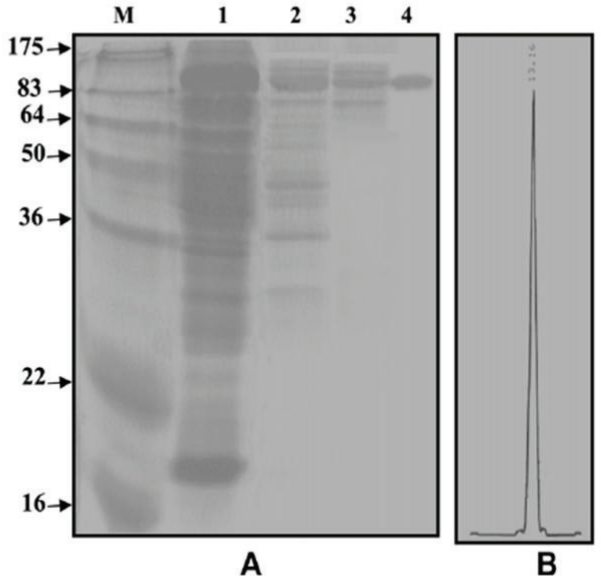


Fig. 1. Electrophoretic and size exclusion HPLC analysis of the purified *S. thermophilus* β-galactosidase. (A) Coomassie brilliant blue-stained SDS-PAGE gel showing the various steps in the purification procedure under reducing conditions. Lane M, protein markers (molecular masses in kDa); lane 1, total protein cell extract; lane 2, sample from lane 1 after sulfate precipitation (40–50%); lane 3, sample from lane 2 after gel filtration; lane 4, purified recombinant β-galactosidase after anion-exchange chromatography. (B) Size exclusion HPLC analysis of the purified recombinant β-galactosidase using protein markers of 669 (RT = 8.755 min), 440 (RT = 13.165 min), 232 (RT = 14.02 min), 140 (RT = 17.163 min) and 66 kDa (RT = 21.438 min), shows a single peak corresponding to a 470 kDa protein (RT = 13.16 min).

2003; Cieslinski et al., 2005; Kim et al., 2004; Rhimi et al., 2009) and differs from that of *L. acidophilus* R22 which is a heterodimer consisting of two subunits of 35 and 72 kDa each (Nguyen et al., 2007a). The purified protein displayed specific activities of 464 and 126 U/mg on ONPG and lactose, respectively. These specific activities were highest when compared to those of *A. psychrolactophilus* F2, *Arthrobacter* sp. HB7, *Bacillus stearothermophilus*, *A. acidocaldarius* and *L. acidophilus* R22 enzymes for which the specific activities on ONPG were evaluated to 20.4, 108.7, 125, 112.7 and 230 U/mg, respectively (Chen et al., 2008; Gul-Guven et al., 2007; Nakagawa et al., 2006; Nguyen et al., 2007a; Trimbur et al., 1994).

3.2. Biochemical characterization

The effect of the concentration of cells inside the immobilized beads on lactose hydrolysis was explored. The results shown in Fig. 2A demonstrate an increasing activity proportional with the cell concentration up to 34 g/l. At higher cell concentrations the relative activity values decreased, possibly due to steric hindrance effects. Studies of the optimal temperature of the purified *S. thermophilus* LMD9 β-galactosidase indicated a wide range comprised between 25 and 40 °C (Fig. 2B). Interestingly, this enzyme was highly active at low temperatures and displayed 70, 72 and 75% of its maximal relative activity at 4, 10 and 15 °C, respectively. At temperatures above 55 °C, the enzyme was rapidly inactivated (Fig. 2B). As compared hereto, a β-galactosidase from another *S. thermophilus* strain being fully active at 57 °C has been reported (Ramana Rao and Dutta, 1981). Immobilization did not affect the optimal temperature, but increased the relative activity at both low and high temperatures (Fig. 2B). For example, at 4 °C, the free enzyme and the enzyme immobilized in cells display 70 and 80% of their relative activities, respectively. The wide optimal temperature range distinguishes this enzyme from the majority of previously reported β-galactosidases, including those from psychrophiles (Table 1). Indeed, β-galactosidases from *Arthrobacter* sp C2-2, *Pseudoalteromonas haloplanktis* and *Pseudoalteromonas* sp. 22b retain only 12.5, 25 and 8% of their maximal activity at 5 °C, respectively (Cieslinski et al., 2005), whereas *A. psychrolactophilus* F2 β-galactosidase retains approximately 90% of its relative activity at 4 °C (Nakagawa et al., 2006). This property makes β-galactosidase from the food-grade *S. thermophilus* LMD9 highly attractive as an efficient biocatalyst for lactose hydrolysis in milk. Analysis of the activity profile at different pH values indicated that soluble and immobilized enzymes were optimally active at pH values comprised between 6.5 and 7.5 (Fig. 2C). These results are similar to

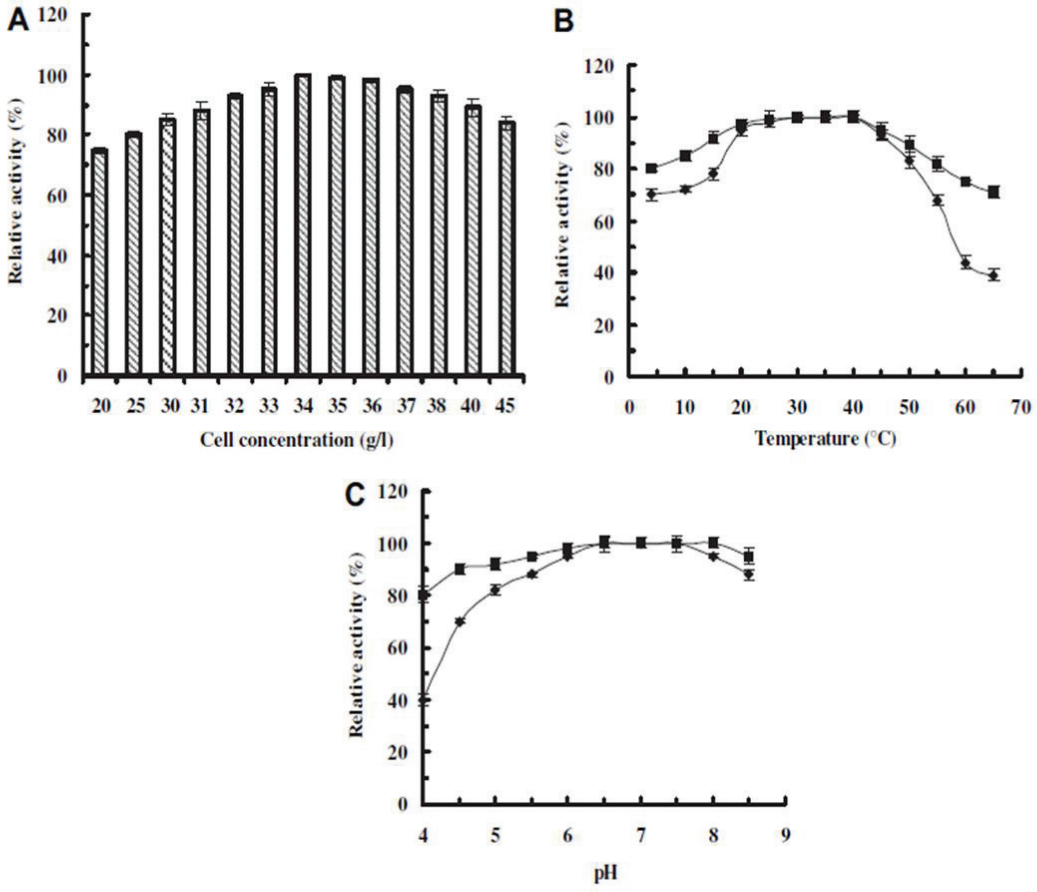


Fig. 2. Cell concentrations (A), temperature (B) and pH (C) activity profiles of the purified soluble (◆) and immobilized (■) β-galactosidase. Activities at optimal cell concentrations, pH and temperature were defined as 100%. Error bars represent the standard deviation from three separate experiments.

Table 1
Biochemical properties of β-galactosidases.

Strain	T_{opt} (°C)	pH	K_m (mM)	V_{max} (U/mg)	k_{cat} (s^{-1})	Metal ion requirement	References
Streptococcus thermophilus LMD9	25–40	6.5–7.5	1.1	464	234	Mn^{2+}, Mg^{2+}	This work
Lactobacillus bulgaricus	35–50	5.0–5.5	0.98	72.5	45.7	Mn^{2+}, Co^{2+}	(Rhimi et al., 2009)
Pseudoalteromonas haloplanktis	50	–	2.4	–	33	–	(Hoyoux et al., 2001)
Lactobacillus acidophilus	55	6.5–8	4.4	28.8	–	Mg^{2+}	(Nguyen et al., 2007a)
Arthrobacter psychrolactophilus F2	10	8.0	42.1	–	–	–	(Nakagawa et al., 2006)
Alicyclobacillus acidocaldarius	65	6.0	–	–	–	–	(Gul-Guven et al., 2007)
Bifidobacterium infantis	60	7.5	2.6	262	–	Na^{+}, K^{+}	(Hung and Lee, 2002)
Lactobacillus reuteri	50	7					(Nguyen et al., 2007b)
Thermus sp. A4	70	6.5	19	–	12.5	Zn^{2+}, Mn^{2+}, Co^{2+}	(Ohtsu et al., 1998)
Bacillus licheniformis	45	5.7					(Phan Tran et al., 1998)
Klyveromyces lactis	40	7.0	20	570	–	Mn^{2+}	(Kim et al., 2003)
Thermus aquaticus YT1	80	5.5	6.6	–	5.7	Ca^{2+}	(Berger et al., 1997)
Flavobacterium sp	42	7.5	0.65	62	–	–	(Sorensen et al., 2006)
Thermus sp. T2	80	5.5	13	–	–	–	(Vian et al., 1998)
E. coli	55	7.2	0.2	52.2	–	Mg^{2+}	(Coker et al., 2003)
Bacillus stearothermophilus	70	7	–	–	–	–	(Chen et al., 2008)
Thermotoga maritima	80	6.5	1.6	63.3	–	Mn^{2+}	(Kim et al., 2004)
Pseudoalteromonas sp 22b	40	6.0–8.0	3.3	–	157	Na^{+}, K^{+}, Mg^{2+}, Mn^{2+}	(Cieslinski et al., 2005)

those reported for a *S. thermophilus* β-galactosidase, which was active at pH 6.8–7.0 and differed from those obtained with *Kluyveromyces fragilis* β-galactosidase, where immobilization shifted the optimal temperature from 35 to 50 °C and the optimal pH from 7.0 to 6.0 (Ramana Rao and Dutta, 1981; Szczodrak, 2000). Notably, immobilization highly improved the relative activity values, particularly at acidic pH; indeed the relative activity values increased from 40 to 80% at pH 4.0 and from 70 to 90% at 4.5. Immobilization therefore clearly enhances the enzyme activity under extreme conditions of temperatures and pH. Furthermore, immobilized cells harboring *S. thermophilus* LMD9 β-galactosidase showed relative activities higher than 80% at pH values ranging from 4.0 to 8.5. This behavior matches that of the *L. bulgaricus* and *Aspergillus oryzae* β-galactosidases that are active at acidic pH including pH 4.0, and differs from other β-galactosidases of *Lactobacillus reuteri, Thermus* sp. A4, *A. psychrolactophilus* strain F2, *Pseudoaltermonas* sp. 22b, *Arthrobacter* sp. C 2-2, *Bifidobacterium longum* CCRC 15708 and *A. acidocaldarius,* which are inactive at this pH (Cieslinski et al., 2005; Hsu et al., 2005; Karasova-Lipovova et al., 2003; Nakagawa et al., 2006; Nguyen et al., 2007a; Ohtsu et al., 1998). The efficiency of the *S. thermophilus* LMD9 β-galactosidase in a wide pH range makes this enzyme attractive for industrial applications as a lactose bioconverter. This takes on more importance when one considers that an appropriate enzyme for lactose hydrolysis in milk and whey should be active at pH 6.7–6.8 and 5.5–6.0, respectively (Szczodrak, 2000; Trimbur et al., 1994).

Metal ion dependency was investigated by treating the purified protein with EDTA, resulting in a 48% activity loss. Addition of metal ions and especially manganese and magnesium restores the activity (Fig. 3A). As shown in Fig. 3B, for maximal activity, *S. thermophilus* β-galactosidase required only 1 mM Mn^{2+} and 1.2 mM Mg^{2+}. Divalent Ca^{2+}, Cu^{2+}, Co^{2+}, Fe^{2+} and Zn^{2+} ions are poor activators and Ni^{2+} is an inhibitor of this enzyme (Fig. 3A). Notably, a previously described β-galactosidase from *S. thermophilus* was inhibited by Cu^{2+} (Ramana Rao and Dutta, 1981). Thus, like the majority of β-galactosidases previously described, that of *S.*

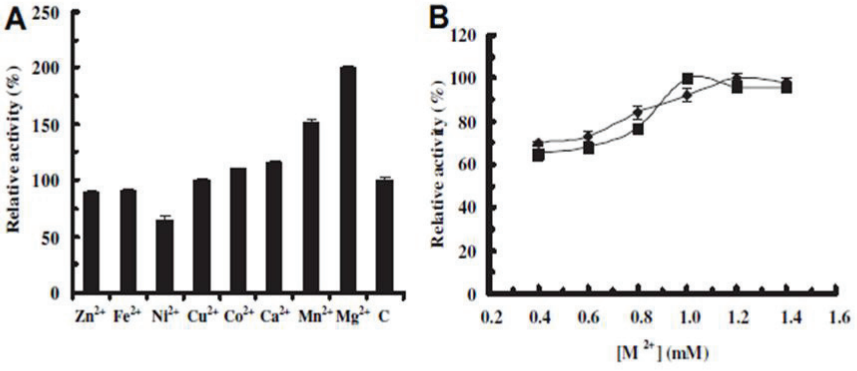


Fig. 3. *S. thermophilus* β-galactosidase activity (A): Effect of divalent metal ions. (B): Effect of various concentrations of Mg^{2+} (◆) and Mn^{2+} (■). "C" indicates enzyme activity measured for the EDTA-treated and purified enzyme. Activities at optimal metal ion concentrations were defined as 100%. Error bars represent the standard deviation from three separate experiments.

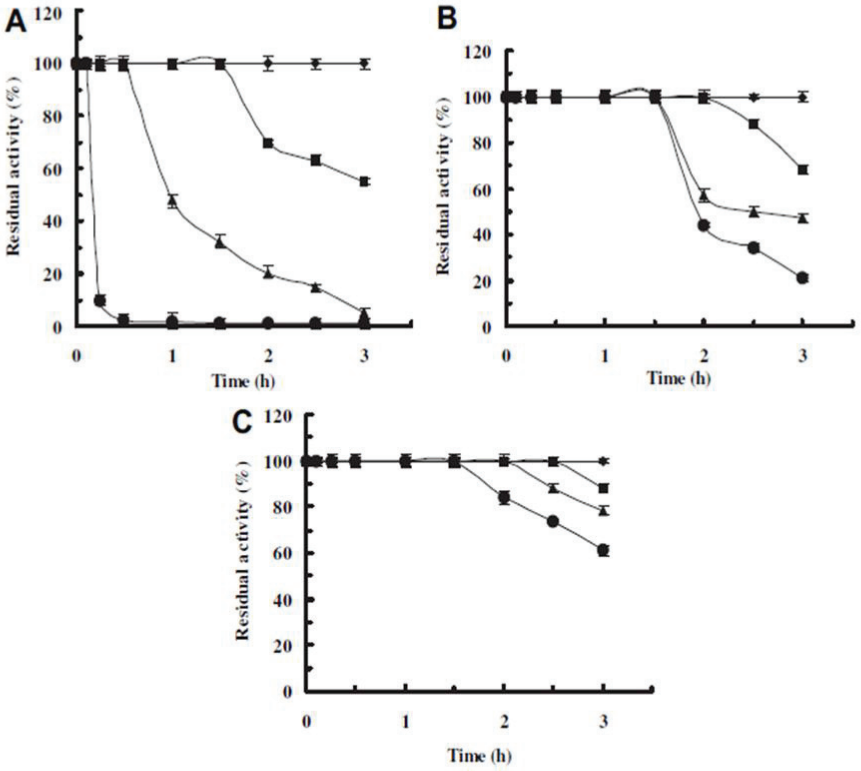


Fig. 4. Thermostability profiles of the metal depleted *S. thermophilus* LMD9 β-galactosidase at different temperatures: (◆) 45 °C, (■) 50 °C, (▲) 55 °C and 60 °C (●) for (A) non-immobilized enzyme (B) non-immobilized enzyme in the presence of Mg^{2+} (1.2 mM) and Mn^{2+} (1 mM) and (C) immobilized enzyme. The initial activity was defined as 100%. Error bars represent the standard deviation from three separate experiments.

thermophilus LMD9 is a metalloprotein requiring metal ions for its activity (Table 1).

3.3. Thermal and pH stabilities

Stability is a crucial parameter when using an enzyme on the industrial scale. Thermal stability was explored using EDTA-treated protein and we found that *S. thermophilus* LMD9 β-galactosidase was stable at temperatures including 4–45 °C for 24 h (data not shown). At higher temperatures, protein thermostability decreased displaying half-lives of 3, 1 and 0.16 h at 50, 55 and 60 °C, respectively (Fig. 4A). Adding 1 mM of Mn^{2+} and 1.2 mM Mg^{2+} did not affect thermostability at temperatures below 45 °C, whereas in the presence of these divalent ions and at temperatures exceeding 45 °C, enzyme thermostability was clearly enhanced (Fig. 4B). After 2.5 h incubation at 50 °C, the enzyme retained more than 85% of its initial activity and the half-lives were increased to approximately 2.5 h and 2 h at 55 and 60 °C, respectively (Fig. 4B). Also, the immobilized enzyme was more stable than free β-galactosidase in the presence of Mn^{2+} and Mg^{2+}. Indeed, as shown in Fig. 4C, in immobilized cells, the enzyme preserved more than 80% of its initial activity at 50 and 55 °C after 3 h of incubation. These results demonstrate that *S. thermophilus* LMD9 β-galactosidase is completely thermostable at temperatures from 4 °C to 45 °C without addition of divalent ions. At temperatures above 45 °C, addition of 1 mM of Mn^{2+} and 1.2 mM Mg^{2+} enhanced enzyme stability, thus highlighting their stabilizing effect at high temperatures. In addition hereto, cell immobilization in alginate beads clearly increased enzyme thermostability. Taken together, these behaviors make this enzyme attractive for industrial applications.

Study of the pH stability of the soluble enzyme indicated that half-lives decreased around pH 6.5, notably when going from pH 6.5 to 4.5 rather than from 6.5 to 7.5 (Table 2). This clearly indicated that *S. thermophilus* LMD9 β-galactosidase is more stable at pH 7.0 and 7.5 than at acid pH. Amongst the β-galactosidases reported thus far, *L. bulgaricus* ATCC 11842 and *S. thermophilus* LMD9 enzymes (the latter described herein) were highly active within a large pH range (4.0–7.0). However, the *S. thermophilus* protein is more stable than that of *L. bulgaricus* (Rhimi et al., 2009). Furthermore, as

Table 2
pH stability of soluble and immobilized (between parentheses) *S. thermophilus* LMD9 β-galactosidase as a function of pH at different temperatures.

pH	Temperature (°C)			
	4	10	30	45
	Half-lives ($t_{1/2}$, h)			
4.5	24(29)	25(30)	24(29)	23(29)
5.0	39(41)	41(42)	40(41)	40(42)
5.5	43(44)	42(44)	42(43)	43(43)
6.0	66(74)	67(73)	67(72)	66(74)
6.5	72(86)	73(88)	73(89)	72(88)
7.0	67(80)	68(79)	69(80)	69(80)
7.5	68(74)	67(73)	67(72)	68(73)

Table 3
Hydrolysis of lactose (%) in milk and whey using soluble and immobilized *S. thermophilus* LMD9 β-galactosidase, respectively, at 45 °C. Values in parentheses correspond to hydrolysis at 4 °C.

Time (h)	Milk		Whey	
	Soluble	Immobilized	Soluble	Immobilized
1	30(29)	31(32)	37(33)	38(32)
2	46(38)	48(43)	58(44)	58(49)
3	62(52)	63(58)	69(58)	70(64)
4	74(61)	75(69)	81(64)	82(72)
5	83(70)	83(77)	87(71)	88(76)
6	84(72)	85(77)	88(73)	89(78)

compared to the free enzyme, when immobilized and at pH 6.5 and 7.0, the half-lives were 87 and 78 h, respectively, versus 74 and 67 h for the free enzyme. This behavior is very interesting when one considers that lactose hydrolysis in milk requires stability and activity at pH 6.7–6.8, thus perfectly matching with the enzyme studied herein.

In order to gain further insight into this issue, we studied pH stability at different temperatures. As shown in Table 2, the half-lives did not significantly change from 4 to 45 °C, underlining remarkable protein pH stability in this large range of temperatures. Moreover, higher half-lives were obtained when using immobilized cells at various temperatures and pH values (Table 2), demonstrating that immobilization increases enzyme pH stability, thereby promoting industrial application of this biocatalyst. β-galactosidase from *A. oryzae* is highly stable at pH 4.6, but preserves only 45% of its initial activity after 5 h incubation at pH 7.4 (Haider and Husain, 2008). Hence, this enzyme seems to be more efficient for lactose hydrolysis in whey than in milk. The substantial *S. thermophilus* LMD9 β-galactosidase stability over a wide pH range (4.5–7.0) further promotes the use of this enzyme for lactose hydrolysis not only in whey but also in milk.

3.4. Kinetic studies and lactose hydrolysis

Study of enzyme kinetic parameters revealed that *S. thermophilus* LMD9 β-galactosidase has remarkable activity toward lactose when compared to the previously reported β-galactosidases; this is related to its high enzyme turnover of 234 s^{-1} (Table 1). The efficiency of this enzyme for lactose bioconversion was explored using this biocatalyst for hydrolysis of lactose in whey and milk under batch conditions (Table 3). At 4 °C bioconversion rates of 72 and 77% were achieved in milk with the soluble and immobilized enzymes, respectively, whereas in whey these values were 73 and 78%, respectively. In general, the highest lactose hydrolysis rates were obtained for the immobilized enzyme in whey, as in milk. When using free or immobilized enzyme at 45 °C, similar bioconversion rates (~89%) were obtained for milk and whey (Table 3), whereas at 4 °C, bioconversion rates were lower at 45 °C (Table 3), related to the lower enzyme activity at this temperature. Indeed, at 4 °C, soluble and immobilized enzymes preserved 70 and 80% of their relative activities, respectively; however at 45 °C, both free and immobilized enzymes were fully active (Fig. 2A). Similar bioconversion rates (~90%) were achieved with *K. fragilis* β-galactosidase immobilized in porous glass, *A. oryzae* β-galactosidase immobilized in a PCV-silica sheet and *Klyveromyces lactis* immobilized in corn grits (Siso and Doval, 1994; Szczodrak, 2000; Wierzbicki et al., 1974). In order to further investigate the enzyme's efficiency for lactose bioconversion, lactose hydrolysis was carried out in whey (pH 4.5, 45 °C) and milk (pH 6.5, 4 °C) in a packed-bed bioreactor. With the aim of optimizing the operating parameters, conversion yield and productivity were investigated at various dilution rates. As illustrated in Fig. 5A, in whey, a dilution rate of 0.2 h^{-1} resulted in the highest conversion yield, evaluated at 91%, and the lowest productivity, 0.12 g/l h. A dilution rate of 8 h^{-1} resulted in the lowest conversion yield, estimated at 30%, combined with maximal productivity of 8.8 g/l h. In milk, similar variations were observed and a dilution rate of 0.2 h^{-1} provided the highest conversion yield of 78% and the lowest productivity of 0.19 g/l h. (Fig. 5B). The highest dilution rate of 8 h^{-1} corresponds to the lowest hydrolysis rate of 31% and the highest productivity evaluated at 7.9 g/l h. Therefore, decreasing the dilution rate enables increasing the hydrolysis rate due to the longer sojourn time course in the column, and thereby longer contact time between substrate and enzyme. In contrast, the productivity parameter increased with increasing dilution rates. This phenomenon may be related to the decrease in the cycle time as the dilution rate increases. Indeed, the increase in the flow rate was generally limited by both velocities of mass transfer and the hydrolysis reaction. The appropriate dilution rate for optimal productivity and the

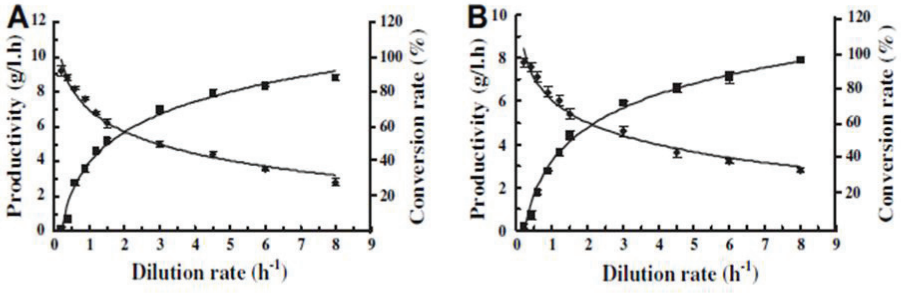


Fig. 5. Effect of the dilution rate on the productivity (■) and the bioconversion rate (◆) in a packed-bed bioreactor. Assays were carried out at 45 and 4 °C in whey (A) and milk (B), respectively. Error bars represent the standard deviation from three separate experiments.

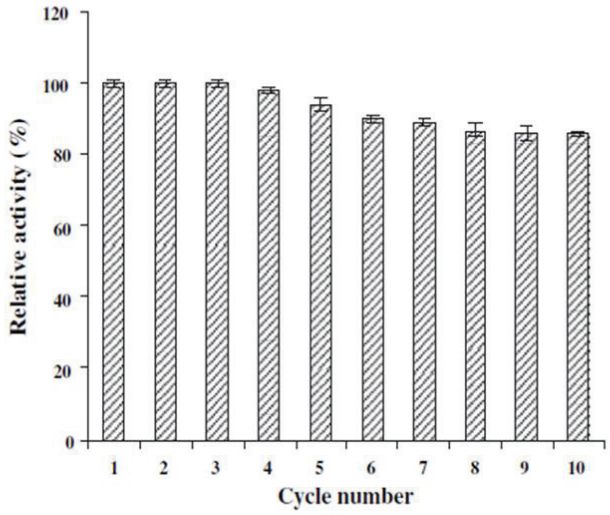


Fig. 6. Re-use stability of *S. thermophilus* LMD9 β-galactosidase immobilized in alginate beads. Initial activities were defined as 100%. Error bars represent the standard deviation from three separate experiments.

optimal conversion rate corresponds to the intersection of the two curves in Fig. 5A and B. Thus, the optimal dilution rate, productivity and conversion rate were evaluated at 2 h^{-1}, 5.8 g/l h and 58%, respectively, in whey, and 2.15 h^{-1}, 5 g/l h and 50%, respectively, in milk.

Due to its great efficiency at both 4 °C and at 45 °C, *S. thermophilus* LMD9 β-galactosidase appears to be an attractive enzyme enabling lactose hydrolysis under conditions which are compatible to storage of milk or during milk fermentation. Furthermore, lactose hydrolysis products can be used for production of novel and low calorie syrups with high sweetness (Rhimi et al., 2007).

3.5. Re-use stability

The re-use stability of an immobilized biocatalyst is an essential parameter in industry. Immobilized cells were recycled and used many times, followed by measurement of β-galactosidase activity. As shown in Fig. 6, after ten cycles of use, more than 85% of the initial β-galactosidase activity was retained. This clearly demonstrates that immobilization of *E. coli* cells overexpressing *S. thermophilus* LMD9 β-galactosidase is useful for continuous conversion of lactose from milk or whey.

We report the molecular and biochemical characterization of probiotic *S. thermophilus* LMD9 β-galactosidase. Its high activity at low temperatures (4–40 °C), along with its wide pH range (4–8.5), distinguishes it from previously characterized β-galactosidases. Enzyme activity and thermostability were remarkably enhanced by adding metal ions and, in particular, by immobilization. The use of free and immobilized enzymes for lactose bioconversion, in milk and whey, in batch and packed-bed bioreactor systems, underlines the great efficiency of this protein. The important bioconversion rates of lactose in milk and whey at different conditions promotes its use for lactose hydrolysis in milk and other dairy products. Indeed, this enzyme appears as a powerful biocatalyst for lactose extracted from whey, which constitutes a cheap source of D-galactose and D-glucose. This takes more importance when one considers that these monosaccharides are used for the production of hypocaloric sweeteners. Moreover, this protein efficiently cleaves lactose in milk, and may thus contribute to overcome the lactose intolerance problem generated by milk products.

Acknowledgements

The authors would like to express their gratitude to members of the Laboratory of Microbiological Genetics (INRA/Jouy-en-Josas) for their generous help and support. We thank Dr. Michel Becchi for mass spectrometry analysis. Moez Rhimi received a financial support from CNRS/France.

References

Barrett, E., Stanton, C., Zelder, O., Fitzgerald, G., Ross, R.P., 2004. Heterologous expression of lactose- and galactose-utilizing pathways from lactic acid bacteria in *Corynebacterium glutamicum* for production of lysine in whey. Appl. Environ. Microbiol. 70, 2861–2866.

Berger, J.L., Lee, Byong H., Lacroix, C., 1997. Purification, properties and characterization of a high-molecular-mass β-galactosidase isoenzyme from *Thermus aquaticus* YT-I. Biotechnol. Appl. Biochem. 25, 29–41.

Bradford, M.M., 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Chen, W., Chen, H., Xia, Y., Yang, J., Zhao, J., Tian, F., Zhang, H.P., Zhang, H., 2009. Immobilization of recombinant thermostable β-galactosidase from *Bacillus stearothermophilus* for lactose hydrolysis in milk. J. Dairy Sci. 92, 491–498.

Chen, W., Chen, H., Xia, Y., Zhao, J., Tian, F., Zhang, H., 2008. Production, purification, and characterization of a potential thermostable galactosidase for milk lactose hydrolysis from *Bacillus stearothermophilus*. J. Dairy Sci. 91, 1751–1758.

Cieslinski, H., Kur, J., Bialkowska, A., Baran, I., Makowski, K., Turkiewicz, M., 2005. Cloning, expression, and purification of a recombinant cold-adapted β-galactosidase from Antarctic bacterium *Pseudoalteromonas* sp. 22b. Protein Expr. Purif. 39, 27–34.

Coker, J.A., Sheridan, P.P., Loveland-Curtze, J., Gutshall, K.R., Auman, A.J., Brenchley, J.E., 2003. Biochemical characterization of a beta-galactosidase with a low temperature optimum obtained from an *Arthrobacter* isolate. J. Bacteriol. 185, 5473–5482.

Coulier, L., Timmermans, J., Bas, R., Van Den Dool, R., Haaksman, I., Klarenbeek, B., Slaghek, T., Van Dongen, W., 2009. In-depth characterization of prebiotic galacto-oligosaccharides by a combination of analytical techniques. J. Agric. Food Chem. 57, 8488–8495.

Elnashar, M.M., Yassin, M.A., 2008. Lactose hydrolysis by β-galactosidase covalently immobilized to thermally stable biopolymers. Appl. Biochem. Biotechnol. 159, 426–437.

Guarner, F., Perdigon, G., Corthier, G., Salminen, S., Koletzko, B., Morelli, L., 2005. Should yoghurt cultures be considered probiotic? Br. J. Nutr. 93, 783–786.

Gul-Guven, R., Guven, K., Poli, A., Nicolaus, B., 2007. Purification and some properties of a β-galactosidase from the thermoacidophilic *Alicyclobacillus acidocaldarius* subsp. *rittmannii* isolated from Antarctica. Enzyme Microb. Technol. 40, 1570–1577.

Haider, T., Husain, Q., 2007. Calcium alginate entrapped preparations of *Aspergillus oryzae* β-galactosidase: its stability and applications in the hydrolysis of lactose. Int. J. Biol. Macromol. 41, 72–80.

Haider, T., Husain, Q., 2008. Concanavalin A layered calcium alginate-starch beads immobilized β-galactosidase as a therapeutic agent for lactose intolerant patients. Int. J. Pharm. 359, 1–6.

Haider, T., Husain, Q., 2009a. Hydrolysis of milk/whey lactose by β-galactosidase: a comparative study of stirred batch process and packed bed reactor prepared with calcium alginate entrapped enzyme. Chem. Eng. Proc.: Process Intensification 48, 576–580.

Haider, T., Husain, Q., 2009b. Immobilization of β-galactosidase from *Aspergillus oryzae* via immunoaffinity support. Biochem. Eng. J. 43, 307–314.

Haider, T., Husain, Q., 2009c. Immobilization of β-galactosidase by bioaffinity adsorption on Concanavalin A layered calcium alginate-starch hybrid beads for the hydrolysis of lactose from whey/milk. Int. Dairy J. 19, 172–177.

Heng, M.H., Glatz, C.E., 1994. Ion exchange immobilization of changed β-galactosidase fusions for lactose hydrolysis. Biotechnol. Bioeng. 44, 745–752.

Herman, R.E., McKay, L.L., 1986. Cloning and expression of the β-D-galactosidase gene from *Streptococcus thermophilus* in *Escherichia coli*. Appl. Environ. Microbiol. 52, 45–50.

Hinberg, I., Korus, R., O'Driscoll, K.F., 1974. Gel entrapped enzymes: kinetic studies of immobilized beta-galactosidase. Biotechnol. Bioeng. 16, 943–963.

Hoyoux, A., Jennes, I., Dubois, P., Genicot, S., Dubail, F., Francois, J.M., Baise, E., Feller, G., Gerday, C., 2001. Cold-adapted β-galactosidase from the Antarctic psychrophile *Pseudoalteromonas haloplanktis*. Appl. Environ. Microbiol. 67, 1529–1535.

Hsu, C.A., Yu, R.C., Chou, C.C., 2005. Purification and characterization of a sodium-stimulated β-galactosidase from *Bifidobacterium longum* CCRC 15708. World J. Microbiol. Biotechnol. 22, 355–361.

Hung, M.N., Lee, B.H., 2002. Purification and characterization of a recombinant β-galactosidase with transgalactosylation activity from *Bifidobacterium infantis* HL96. Appl. Microbiol. Biotechnol. 58, 439–445.

Karasova-Lipovova, P., Strnad, H., Spiwok, V., Mala, S., Kralova, B., Russell, N.J., 2003. The cloning, purification and characterization of a cold-active β-galactosidase from the psychrotolerant Antarctic bacterium *Arthrobacter* sp. C2-2. Enzyme Microb. Technol. 33, 836–844.

Kim, C.S., Ji, E.S., Oh, D.K., 2004. Characterization of a thermostable recombinant β-galactosidase from *Thermotoga maritima*. J. Appl. Microbiol. 97, 1006–1014.

Kim, C.S., Ji, E.S., Oh, D.K., 2003. Expression and characterization of *Kluyveromyces lactis* β-galactosidase in *Escherichia coli*. Biotechnol. Lett. 25, 1769–1774.

Kunst, A., Draeger, B., Ziegernhorn, J., 1988. Colorimetric methods with glucose oxidase and peroxidase. In: Bergmeyer, H.U., Bergmeyer, J., Grassl, M. (Eds.), Methods of Enzymatic Analysis, third ed.. VCH Publishers, Weinheim, Germany, pp. 178–185.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

Lu, L., Xiao, M., Xu, X., Li, Z., Li, Y., 2007. A novel beta-galactosidase capable of glycosyl transfer from *Enterobacter agglomerans* B1. Biochem. Biophys. Res. Commun. 356, 78–84.

Mahoney, R.R., 1998. Galactosyl-oligosaccharide formation during lactose hydrolysis: a review. Food Chem. 63, 147–154.

Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., Shakhova, V., Grigoriev, I., Lou, Y., Rohksar, D., Lucas, S., Huang, K., Goodstein, D.M., Hawkins, T., Plengvidhya, V., Welker, D., Hughes, J., Goh, Y., Benson, A., Baldwin, K., Lee, J.H., Díaz-Muñiz, I., Dosti, B., Smeianov, V., Wechter, W., Barabote, R., Lorca, G., Altermann, E., Barrangou, R., Ganesan, B., Xie, Y., Rawsthorne, H., Tamir, D., Parker, C., Breidt, F., Broadbent, J., Hutkins, R., O'Sullivan, D., Steele, J., Unlu, G., Saier, M., Klaenhammer, T., Richardson, P., Kozyavkin, S., Weimer, B., Mills, D., 2006. Comparative genomics of the lactic acid bacteria. Proc. Natl. Acad. Sci. U.S.A. 103, 15611–15616.

Mateo, C., Monti, R., Pessela, C.C.B., Fuentes, M., Torres, R., Guisán, J. M., Fernández-Lafuente, R., 2004. Immobilization of lactase from *Kluyveromyces lactis* greatly reduces the inhibition promoted by glucose: full hydrolysis of lactose in milk. Biotechnol. Prog. 20, 1259–1262.

Molotov, S.V., Duzhii, D.E., Danilevich, V.N., Sukhodolets, V.V., 1990. Cloning of the *Streptococcus thermophilus* β-galactosidase gene and its expression in *Escherichia coli* and *Bacillus subtilis* cells. Mol. Gen. Mikrobiol. Virusol. 4, 10–14.

Nakagawa, T., Fujimoto, Y., Ikehata, R., Miyaji, T., Tomizuka, N., 2006. Purification and molecular characterization of cold-active β-galactosidase from *Arthrobacter psychrolactophilus* strain F2. Appl. Microbiol. Biotechnol. 72, 720–725.

Nguyen, T.H., Splechtna, B., Krasteva, S., Kneifel, W., Kulbe, K.D., Divne, C., Haltrich, D., 2007a. Characterization and molecular cloning of a heterodimeric β-galactosidase from the probiotic strain *Lactobacillus acidophilus* R22. FEMS Microbiol. Lett. 269, 136–144.

Nguyen, T.H., Splechtna, B., Yamabhai, M., Haltrich, D., Peterbauer, C., 2007b. Cloning and expression of the β-galactosidase genes from *Lactobacillus reuteri* in *Escherichia coli*. J. Biotechnol. 129, 581–591.

Ohtsu, N., Motoshima, H., Goto, K., Tsukasaki, F., Matsuzawa, H., 1998. Thermostable β-galactosidase from an extreme thermophile, *Thermus* sp. A4: enzyme purification and characterization, and gene cloning and sequencing. Biosci. Biotechnol. Biochem. 62, 1539–1545.

Phan Tran, L.S., Szabo, L., Fulop, L., Orosz, L., Sik, T., Holczinger, A., 1998. Isolation of a β-galactosidase-encoding gene from *Bacillus licheniformis*: purification and characterization of the recombinant enzyme expressed in *Escherichia coli*. Curr. Microbiol. 37, 39–43.

Pisani, F.M., Rella, R., Raia, C.A., Rozzo, C., Nucci, R., Gambacorta, A., De Rosa, M., Rossi, M., 1990. Thermostable β-galactosidase from the archaebacterium *Sulfolobus solfataricus*. Purification and properties. FEBS Eur. J. Biochem 187, 321–328.

Pospiech, A., Neumann, B., 1995. A versatile quick-prep of genomic DNA from gram-positive bacteria. Trends Genet. 11, 217–218.

Ramana Rao, M.V., Dutta, S.M., 1981. Purification and properties of β-galactosidase from *Streptococcus thermophilus*. J. Food Sci. 46, 1419–1423.

Rhimi, M., Aghajari, N., Jaouadi, B., Juy, M., Boudebbouze, S., Maguin, E., Haser, R., Bejar, S., 2009. Exploring the acidotolerance of β-galactosidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*: an attractive enzyme for lactose bioconversion. Res. Microbiol. 160, 775–784.

Rhimi, M., Messaoud, E.B., Borgi, M.A., Khadra, K.B., Bejar, S., 2007. Co-expression of L-arabinose isomerase and D-glucose isomerase in *E. coli* and development of an efficient process producing simultaneously D-tagatose and D-fructose. Enzyme Microb. Technol. 40, 1531–1537.

Robinson, R.K., 1994. Modern Dairy Technology, second ed., vol. 1. Chapman & Hall, London.

Roy, I., Gupta, M.N., 2003. Lactose hydrolysis by Lactozym(TM) immobilized on cellulose beads in batch and fluidized bed modes. Process Biochem. 39, 325–332.

Rycroft, C.E., Jones, M.R., Gibson, G.R., Rastall, R.A., 2001. A comparative *in vitro* evaluation of the fermentation properties of prebiotic oligosaccharides. J. Appl. Microbiol. 91, 878–887.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: a Laboratory Manual, second ed.. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Schmidt, B.F., Adams, R.M., Requadt, C., Power, S., Mainzer, S.E., 1989. Expression and nucleotide sequence of the *Lactobacillus bulgaricus* beta-galactosidase gene cloned in *Escherichia coli*. J. Bacteriol. 171, 625–635.

Scrimshaw, N.S., Murray, E.B., 1988. The acceptability of milk and milk products in populations with a high prevalence of lactose intolerance. Am. J. Clin. Nutr. 48, 1079–1159.

Siso, M.I.G., Doval, S.S., 1994. *Kluyveromyces lactis* immobilization on corn grits for milk whey lactose hydrolysis. Enzyme Microb. Technol. 16, 303–310.

Sorensen, H.P., Porsgaard, T.K., Kahn, R.A., Stougaard, P., Mortensen, K.K., Johnsen, M.G., 2006. Secreted beta-galactosidase from a *Flavobacterium*

sp. isolated from a low-temperature environment. Appl. Microbiol. Biotechnol. 70, 548–557.

Szczodrak, J., 2000. Hydrolysis of lactose in whey permeate by immobilized β-galactosidase from *Kluyveromyces fragilis*. J. Mol. Catal. B: Enzymatic 10, 631–637.

Terzaghi, B.E., Sandine, W.E., 1975. Improved medium for lactic Streptococci and their Bacteriophages. Appl. Microbiol. 29, 807–813.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Trimbur, D.E., Gutshall, K.R., Prema, P., Brenchley, J.E., 1994. Characterization of a psychrotrophic *Arthrobacter* gene and its cold-active β-galactosidase. Appl. Environ. Microbiol. 60, 4544–4552.

Vian, A., Carrascosa, A.V., Garcia, J.L., Cortes, E., 1998. Structure of the beta-galactosidase gene from *Thermus* sp. strain T2: expression in *Escherichia coli* and purification in a single step of an active fusion protein. Appl. Environ. Microbiol. 64, 2187–2191.

Wierzbicki, L.E., Edwards, V.H., Kosikowski, F.V., 1974. Hydrolysis of lactose in acid whey using β-galactosidase immobilized on porous glass particles: preparation and characterization of a reusable catalyst for the production of low-lactose dairy products. Biotechnol. Bioeng. 16, 397–411.