



**HAL**  
open science

# Identification et caractérisation du système alginolytique de la bactérie marine *Zobellia galactanivorans*

François Thomas

## ► To cite this version:

François Thomas. Identification et caractérisation du système alginolytique de la bactérie marine *Zobellia galactanivorans*. Microbiologie et Parasitologie. Paris 6, 2011. Français. NNT: . tel-01110859

**HAL Id: tel-01110859**

<https://hal.sorbonne-universite.fr/tel-01110859v1>

Submitted on 29 Jan 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.

**THESE DE DOCTORAT DE  
L'UNIVERSITE PIERRE ET MARIE CURIE**

Spécialité  
**Microbiologie et Biochimie**

Ecole Doctorale Interdisciplinaire pour le Vivant

Présentée par

**M. François THOMAS**

Pour obtenir le grade de

**DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE**

# **Identification et caractérisation du système alginolytique de la bactérie marine *Zobellia galactanivorans***

soutenue le 23 septembre 2011

devant le jury composé de :

M. Gurvan MICHEL  
Mme Nicole COTTE-PATTAT  
M. Mirosław CYGLER  
M. Jean-François BERNARDET  
Mme Corine SANDSTRÖM  
M. Germain TRUGNAN

CNRS (Roscoff)  
CNRS (Villeurbanne)  
NRC-BRI (Montreal)  
INRA (Jouy-en-Josas)  
SLU (Uppsala)  
UPMC (Paris)

Directeur de thèse  
Rapportrice  
Rapporteur  
Examinateur  
Examinatrice  
Examinateur



après utilisation être with  
régions gène transcription voie  
acides structure permet trois  
action BamHI mL algues were  
galactanivora  
was AlyA5 an peut gènes  
from génome impliqués  
conditions bactéries marine solution  
for système eau fonction résidus  
expression activité module  
motifs analyse nombre substrat Tableau données aminés  
comme protéines  
lyase foie produits mL The  
enzymes protéine  
dégrader résultats two milieu  
différentes débit peuvent que entre autres famille  
réaction Cette différents que oligosaccharides autre exemple chez is  
by lyases présence is enzyme paroi  
cibles matière métabolisme permis étude mM AlyA1  
Chapitre NA min également polypeptide perception  
polysaccharide this brunes glucose phase  
are cette utilisant genes source  
zobellia pH séquences carbone  
plus signal blocs forme acide ADN montre sous  
pourraient pourraient culture voies  
dont Bacteroidetes





---

*People want to understand the Earth, but they don't, so they build a model. Then they have two things they don't understand.*

*Pr. Phil England (University of Oxford)*

*Dans la vie, il faut rebondir.*

*Tigrou*



---

# Remerciements

---

Eh bien voilà, voici venu le temps d'écrire les remerciements... En cinq ans de présence à la Station, forcément, ça fait du monde à remercier !

Je souhaiterais tout d'abord remercier M. Bernard Kloareg pour m'avoir accueilli à la Station Biologique de Roscoff. Cet entretien en juillet 2006 a indiscutablement été déterminant dans la suite de mon parcours. J'adresse également ma plus vive gratitude à Mme Catherine Boyen pour m'avoir permis de réaliser ma (mes ?) thèse(s) dans son Unité. Merci de m'avoir fait confiance !

Je remercie Mme Nicole Cotte-Pattat, Mme Corine Sandström, M. Mirek Cygler, M. Jean-François Bernardet et M. Germain Trugnan d'avoir accepté de juger ce travail. Merci également à Mme Coralie Bompard, M. Gérald Peyroche et M. Eric Duchaud pour leurs précieux conseils lors de mon comité de thèse.

Un énorme MERCI à Mme Mirjam Czjzek et M. Gurvan Michel, mes co-directeurs de thèse. Gurvan, Mirjam, comment vous dire à quel point je vous suis reconnaissant de m'avoir fait confiance, de m'avoir permis d'arriver où j'en suis aujourd'hui, de m'avoir écouté, de m'avoir supporté. Gurvan, merci pour ton soutien et tes soixante idées à la minute. Travailler avec toi est hyper stimulant ! Merci de m'avoir fait m'épanouir dans la Zobelliomique intégrative. Ah oui j'allais oublier ! Merci d'avoir passé ton HDR avant ma thèse ! ;-)  
Mirjam, merci pour ta disponibilité, ton calme, ton positivisme, ta philosophie.

J'adresse par ailleurs un merci chaleureux et sincère à Mme Catherine Leblanc et M. Philippe Potin. Catherine, Philippe, merci pour votre compréhension. Je vous dois beaucoup.

Merci à tous les membres passés ou présents de l'équipe Glycobiologie marine. Ça a été une joie d'être un Cheese-Cake boy. Et je ne doute pas qu'Aurore va lancer une nouvelle mode de Gurvanette. Aurore : l'important, c'est de s'amuser ! Merci à Tristan (mon modèle de microbiologiste), Lionel et Cécile pour les discussions enrichissantes. Murielle et Alex, le labo ne tournerait pas sans vous ! Merci pour le temps que vous avez pris pour me former quand j'en avais besoin. Merci aussi à Justyna, Jan-Hendrik, Etienne et Richard.

Je souhaite remercier tous les membres de l'UMR 7139 avec qui j'ai pu interagir et travailler (mais pas que) pendant ces années à la Station. Un merci tout particulier à Pi, Sylvie, Delphine, Sabine J., Gaëlle et Ludo pour votre disponibilité et votre sourire. Merci aussi à (dans l'ordre des étages) : Maryvonne, Mark, Olivier, Susana (on va enfin avoir le temps de le travailler ce trio !), William, Thierry (attention à pas trop te faire contaminer par Gurvan), Agnès, Sophie, Nathalie, Laurence MC, Zofia, Jessica, Andres (ah ton sens de l'humour va me manquer !), Ben, Maude, Sabine C., Bernard B., Laurence D., Jonas, Stéphane, JB, Zhanru, Aurélie, Fred, Alok. Merci pour vos conseils et les discussions enrichissantes. Merci aussi pour votre générosité et tous les bons moments, les petits murets (mais si, les travaux finiront un jour !), les concours du meilleur gâteau (pour moi, je crois que ça reste la tarte banane-chocolat), bref, toutes ces choses qui font qu'on se sent bien dans l'unité ! Je n'oublie pas non plus les (nombreux) thésards ou post-docs qui sont passés par le

labo et partis depuis mais sans qui l'ambiance n'aurait pas été la même : Stéphanie, Diane, Maud, Marion, Simon, Sarah, Aude, Jan-Hendrik, Etienne, PO, Alexandra. Un merci tout particulier à Audrey, qui n'imagine sans doute pas à quel point je le suis reconnaissant. Merci, entre autres, pour les bons moments au labo (40 cm ?? non, attends on va dessiner pour voir), sur le terrain (miam, un helcion), autour d'un barbecue (ah.....l'orgasme givré !). Merci aussi d'être entrée en salle 119 ce jour-là.

Merci à tout le personnel de la Station Biologique de Roscoff, un lieu où il fait bon travailler et vivre. Merci aux Services Techniques qui font un boulot monstre pour que tout ça tourne rond. Merci à Marie-Noëlle pour ses petits mots tous les matins. Pour tous les bons moments, merci  $\times 10^{10}$  à Vincent, Laure et Daniella. Merci à Christophe, sans qui ma flûte traversière aurait probablement moisi au fond d'un tiroir ! Merci aussi à Xavier, Valeria, Catharina, Fanny, Ines, Caroline, Sarah, Antonio, Dianne, Christian, Joana, Mathieu, Anne-Laure, Jean-Charles, Claire, Gauthier, Jihane, Peter, Sakina, Gildas et sûrement encore beaucoup de monde que j'oublie.

Merci à toute l'équipe pédagogique et technique du Département BGB de l'ENS Cachan, avec qui ça a été un plaisir d'enseigner ces trois dernières années. Particulièrement, merci à Valérie, Catherine et Isa. Et bien sûr, merci MONSIEUR Gérald, un vrai modèle pour moi.

Cécile, merci d'avoir eu l'idée d'habiter dans la rue la plus chaude de Saint Pol ! Merci de supporter mon hyper-activité, de trouver mes gâteaux bons, de me faire découvrir tous les jours des groupes (comment dire, éclectiques ?), de savoir utiliser une perceuse, allumer un barbecue, tailler une haie, servir l'apéro, dire « le chameau et le renard sont grands mais ils ne sont pas gros » en arabe (si, si, il paraît que c'est utile !!) et faire la soupe au poireau la plus originale du monde. Vivre dans la maison anti-clichés aura été un vrai plaisir.

Merci également à tous mes amis qui, de près ou de loin, m'ont soutenu. De près, merci à l'école de danse Coïnci'Danse de Morlaix, Françoise, Delphine, Oriane, Bernard, Julie, Maude, et toutes les filles de la danse qui ne liront probablement jamais ces lignes. Sans vous, ces cinq ans m'auraient paru bien longs ! De plus loin géographiquement mais pas dans le cœur, merci à tout le groupe 41 et assimilés, vous me manquez !! Comprenez-vous un jour que je ne travaille pas sur les moules ?! Merci aussi à Alexandrine, Jean et Guillaume.

Et bien sûr, un incommensurable merci à toute ma famille, en particulier à Maman, Papa, Olivier, mes grands-parents et Isa. Merci de croire en moi, de me faire aller de l'avant, d'être toujours derrière moi même si c'est un peu difficile à suivre...

Et enfin, Jérôme, merci. En dire plus ici n'aurait pas de sens, ou plutôt trop de sens. Alors le plus simplement, mais surtout le plus sincèrement du monde, MERCI...

**MERCI A TOUS !**

# Sommaire

<b>Remerciements</b> .....	<b>3</b>
<b>Sommaire</b> .....	<b>5</b>
<b>Chapitre 1 : Introduction générale</b> .....	<b>15</b>
<b>I- Ecologie microbienne marine</b> .....	<b>17</b>
I-1 Importance des bactéries hétérotrophes dans le cycle du carbone .....	17
I-2 Les régions côtières : des écosystèmes particuliers .....	25
<b>II- La paroi des algues brunes</b> .....	<b>28</b>
II-1 Place des algues brunes dans l'évolution des Eucaryotes.....	28
II-2 Les parois végétales : considérations générales .....	30
II-3 Architecture de la paroi des algues brunes.....	32
II-4 L'alginate : un composé majoritaire.....	34
II-5 Rôle de la paroi des algues brunes dans l'interaction avec les bactéries.....	43
<b>III- <i>Zobellia galactanivorans</i> : un modèle d'interaction algues/bactéries</b> .....	<b>46</b>
III-1 <i>Z. galactanivorans</i> et le genre <i>Zobellia</i> .....	46
III-2 <i>Zobellia</i> et les algues: "Je t'aime...moi non plus !" .....	48
III-3 <i>Z. galactanivorans</i> : un dégradeur d'algues marines .....	49
<b>IV- La dégradation de l'alginate</b> .....	<b>53</b>
IV-1 Les organismes dégradeurs .....	53
IV-2 La voie catabolique de l'alginate .....	56
IV-3 Les alginate lyases .....	58
<b>V- Présentation du travail</b> .....	<b>69</b>
<b>Chapitre 2 : Bacteroidetes intestinaux et environnementaux : "The food connection" .....</b>	<b>71</b>
<b>Chapitre 3 : Evaluation des gènes de référence pour la RT-PCR quantitative</b>	
<b>chez <i>Z. galactanivorans</i>.....</b>	<b>91</b>
<b>Chapitre 4 : Etude du système algolytique de <i>Z. galactanivorans</i> .....</b>	<b>103</b>
<b>Chapitre 5 : Analyse transcriptomique à l'échelle du génome .....</b>	<b>173</b>
<b>I- Introduction</b> .....	<b>174</b>
<b>II- Matériels et Méthodes</b> .....	<b>179</b>
II-1 Construction de la puce oligonucléotidique.....	179
II-2 Culture de <i>Z. galactanivorans</i> et préparation des ARN.....	179
II-3 Synthèse des ADN complémentaires double brin.....	180
II-4 Marquage et hybridation des échantillons d'ADNc.....	181
II-5 Traitement et analyses statistiques des données transcriptomiques .....	181
<b>III- Résultats et Discussion</b> .....	<b>183</b>
III-1 Une puce à ADN pangénomique pour <i>Z. galactanivorans</i> .....	183

III-2	Contrôles qualité et validation des données transcriptomiques .....	185
III-3	Analyse globale de l'expression dans les différentes conditions .....	190
III-4	L'alginate induit une régulation massive du transcriptome .....	197
III-5	Les gènes surexprimés en présence d'alginate .....	202
III-6	Les gènes sous-exprimés en présence d'alginate.....	216
<b>IV-</b>	<b>Vers un nouveau modèle de régulation .....</b>	<b>220</b>
<b>Chapitre 6 Clonage et surexpression à moyen débit de protéines impliquées dans la dégradation de l'alginate.....</b>		<b>223</b>
<b>I-</b>	<b>Introduction.....</b>	<b>224</b>
I-1	Contexte .....	224
I-2	Philosophie de la méthode de surexpression à moyen débit.....	225
<b>II-</b>	<b>Matériels et Méthodes .....</b>	<b>227</b>
II-1	Analyses bioinformatiques des séquences .....	227
II-2	Dessin des amorces oligonucléotidiques .....	227
II-3	Préparation du matériel biologique .....	231
II-4	Clonage.....	234
II-5	Tests d'expression .....	235
<b>III-</b>	<b>Résultats et Discussion .....</b>	<b>236</b>
III-1	Choix des cibles à cloner et analyse de séquences .....	236
III-2	Clonage à moyen débit.....	240
III-3	Discussion.....	245
<b>IV-</b>	<b>Bilan et perspectives.....</b>	<b>247</b>
<b>Chapitre 7 Caractérisation phylogénétique, biochimique et cristallographique d'alginate lyases de la famille PL7.....</b>		<b>250</b>
<b>I-</b>	<b>Présentation .....</b>	<b>251</b>
<b>II-</b>	<b>Matériels et Méthodes .....</b>	<b>252</b>
II-1	Substrats.....	252
II-2	Analyses de séquences et phylogénie.....	252
II-3	Clonage des gènes d'alginate lyases.....	252
II-4	Procédures d'expression et de purification des protéines.....	253
II-5	Tests enzymatiques .....	254
II-6	Chromatographie d'exclusion de taille des produits de dégradation par AlyA5.....	255
II-7	Purification d'oligo-alginates insaturés .....	255
II-8	Analyse des oligosaccharides par électrophorèse .....	255
II-9	Réduction des oligosaccharides .....	256
II-10	Suivis des réactions enzymatiques par spectrométrie RMN .....	256
II-11	Analyse des produits finaux de réaction par RMN et spectrométrie de masse.....	257
II-12	Cristallogénèse .....	258
II-13	Cristallographie et résolution des structures .....	259
<b>III-</b>	<b>Résultats et Discussion .....</b>	<b>261</b>
III-1	Analyses de séquences et phylogénie.....	261
III-2	Surexpression et purification des alginate lyases recombinantes .....	265
III-3	Caractérisation biochimique d'AlyA1 .....	266

---

III-4	Caractérisation biochimique d'AlyA5 .....	275
III-5	Etude structurale des deux alginate lyases AlyA1 et AlyA5 .....	283
<b>IV-</b>	<b>Conclusions .....</b>	<b>297</b>
	<b>Conclusions et_Perspectives générales .....</b>	<b>298</b>
<b>I-</b>	<b>Principaux résultats .....</b>	<b>299</b>
<b>II-</b>	<b>Perspectives .....</b>	<b>302</b>
II-1	Du pain sur la planche... ..	302
II-2	Le potentiel biotechnologique de <i>Z. galactanivorans</i> .....	303
II-3	<i>Z. galactanivorans</i> , un nouvel organisme dégradeur modèle .....	305
	<b>Annexe 1 : Gènes surexprimés en présence d'alginate .....</b>	<b>308</b>
	<b>Annexe 2 : Gènes sous-exprimés en présence d'alginate .....</b>	<b>314</b>
	<b>Bibliographie .....</b>	<b>322</b>



# Table des illustrations

Figure 1-1: Abondance et distribution des différents groupes microbiens marins .....	18
Figure 1-2: Le cycle du carbone dans la mer, selon ZoBell (1946) .....	19
Figure 1-3 : Le rôle des bactéries dans le cycle du carbone océanique .....	21
Figure 1-4 : Le continuum de taille de la matière organique dans l'eau de mer .....	22
Figure 1-5 : Les neiges marines, des <i>hotspots</i> de la colonisation bactérienne. ....	24
Figure 1-6 : L'étendue et la diversité des régions côtières.....	26
Figure 1-7 : Arbre phylogénétique des Eucaryotes (Baldauf 2008) .....	29
Figure 1-8 : Modèle hypothétique de la paroi d'une algue brune. ....	32
Figure 1-9 : Changements dans la composition en polysaccharides de parois isolées de <i>Fucus vesiculosus</i> à différents temps après la fertilisation de l'embryon .....	33
Figure 1-10 : Structure et conformation des constituants monosaccharidiques de l'alginate .....	34
Figure 1-11 : Organisation d'un polysaccharide d'alginate .....	35
Figure 1-12 : La voie de biosynthèse de l'alginate chez les bactéries ( <i>Pseudomonas aeruginosa</i> ) et les algues brunes ( <i>Ectocarpus siliculosus</i> ) .....	36
Figure 1-13 : Les propriétés gélifiantes des alginates .....	38
Figure 1-14 : Exemples d'utilisations de l'alginate dans l'industrie.....	42
Figure 1-15 : Communautés microbiennes associées à l'algue brune <i>Laminaria digitata</i> .....	43
Figure 1-16 : Schéma des réponses de défense des algues brunes face à une attaque bactérienne... ..	44
Figure 1-17 : La bactérie marine <i>Z. galactanivorans</i> , isolée de l'algue rouge <i>D. sanguinea</i> .....	46
Figure 1-18 : Modèle spéculatif de la mobilité par glissement chez les Flavobactéries. ....	47
Figure 1-19 : Les sites d'isolement des membres du genre <i>Zobellia</i> .....	48
Figure 1-20 : Interaction entre <i>Z. galactanivorans</i> et l'algue brune <i>Ectocarpus siliculosus</i> .....	50
Figure 1-21 : Biofilms de <i>Z. galactanivorans</i> en présence d'exsudats de l'algue brune <i>L. digitata</i> .....	51
Figure 1-22 : Encapsulation et germination chez <i>A. vinelandii</i> .....	53
Figure 1-23 : Voie catabolique de l'alginate chez les bactéries.....	56
Figure 1-24 : Système d'import et de dégradation de l'alginate chez <i>Sphingomonas</i> sp.A1 .....	58
Figure 1-25 : Nomenclature des sous-sites dans une polysaccharidase. ....	59
Figure 1-26 : Mécanisme de $\beta$ -élimination d'une alginate lyase .....	60
Figure 1-27 : Les 4 types de repliement connus dans les familles d'alginate lyases.....	63
Figure 1-28 : Exemple de topologies adoptées par les polysaccharidases .....	66
Figure 1-29 : Schéma présentant l'objectif de ce travail de thèse .....	69
Figure 5-1 : Schéma de principe d'une expérience transcriptomique utilisant des puces à ADN .....	174
Figure 5-2 : Nombre annuel de publications traitant de puces à ADN depuis les années 1990.....	175
Figure 5-3 : Principe de la synthèse de puces à ADN par photochimie.....	176
Figure 5-4 : Structure des polysaccharides utilisés dans cette étude .....	177
Figure 5-5 : Analyse des profils des ADNc double-brin avant le marquage .....	185
Figure 5-6 : Graphique en boîte à moustaches des données brutes .....	186
Figure 5-7 : Coefficient de corrélation de Pearson des données brutes avec l'échantillon médian... ..	187
Figure 5-8 : Graphique en boîtes à moustaches des données normalisées.....	188

Figure 5-9 : Représentation graphique de l'analyse PCA sur les données normalisées, selon les deux axes principaux.....	189
Figure 5-10 : Pourcentage de cibles significativement exprimées dans les différentes conditions de culture (p-value < 0.01) .....	190
Figure 5-11 : Analyse de groupement du transcriptome des sept conditions de culture.....	191
Figure 5-12 : Schéma simplifié de l'assimilation du sulfate inorganique chez les bactéries.....	194
Figure 5-13 : Diagramme en Volcano-Plot de la comparaison des valeurs d'expression en présence d'alginate ou de glucose.....	197
Figure 5-14 : Comparaison des taux de variation obtenus en RT-qPCR et par les puces à ADN.....	199
Figure 5-15: Détection d'ARNm polycistroniques par l'approche <i>microarray</i> .....	200
Figure 5-16 : Représentation en diagramme de Venn des gènes dont l'expression est régulée en présence d'alginate ou de laminarine .....	201
Figure 5-17 : Schéma des gènes sur-exprimés en présence d'alginate et de laminarine .....	204
Figure 5-18 : Schéma des systèmes bactériens de transduction du signal impliquant des transferts de phosphate.....	205
Figure 5-19 : Structure chimique de quelques composés phénoliques de végétaux .....	208
Figure 5-20 : Principe de la régulation par un système sigma/anti-sigma .....	209
Figure 5-21 : Place des dérivés du gluconate dans le catabolisme de l'alginate.....	211
Figure 5-22 : Clusters de gènes de fonction inconnue surexprimés en présence d'alginate .....	213
Figure 5-23 : Alignement de séquences de protéines de fonction inconnue surexprimées en présence d'alginate .....	215
Figure 5-24 : Région génomique autour d' <i>alyA7</i> , sous-exprimée en présence d'alginate .....	218
Figure 5-25 : Modèle de la réponse à l'alginate chez <i>Z. galactanivorans</i> .....	221
Figure 6-1 : Principe de l'expression dans le vecteur pFO4.....	226
Figure 6-2 : Résultat de l'analyse de la séquence protéique d'AlyA1 sur le serveur CDD (NCBI) .....	238
Figure 6-3 : Principe de la méthode HCA (d'après Gaboriaud et al., 1987) .....	239
Figure 6-4 : Analyse modulaire de l'alginate lyase AlyA1.....	240
Figure 6-5 : Analyse en SDS-PAGE des fractions purifiées des extraits cellulaires d' <i>E. coli</i> . .....	244
Figure 6-6 : Comparaison des taux de réussite dans les trois stratégies de clonage et surexpression à moyen débit .....	245
Figure 6-7 : Analyse en SDS-PAGE des extraits cellulaires d' <i>E. coli</i> exprimant la protéine AlyA7 en milieu Graffinity.....	247
Figure 7-1: Diagramme de phase du principe de la croissance cristalline .....	258
Figure 7-2 : Techniques de cristallogenèse par goutte assise (A) ou suspendue (B) .....	259
Figure 7-3 : Alignement de séquences multiples des alginate lyases PL7 de <i>Z. galactanivorans</i> avec A1-II' (Q75WP3) de <i>Sphingomonas</i> sp. A1 et alyPG (Q9RB42) de <i>Corynebacterium</i> sp. ALY-1. ....	262
Figure 7-4: Arbre phylogénétique non raciné de 39 enzymes de la famille PL7.....	263
Figure 7-5: Purification des protéines recombinantes AlyA1 et AlyA5. ....	265
Figure 7-6: Effet de la température et de la concentration en NaCl sur l'activité d'AlyA1.....	266
Figure 7-7: Effet du pH sur l'activité d'AlyA1.....	267
Figure 7-8 : Relations entre les constantes cinétiques d'AlyA1 et le contenu en guluronate dans l'alginate. ....	268
Figure 7-9 : Suivi <sup>1</sup> H-RMN de la dégradation d'alginate par AlyA1 .....	269
Figure 7-10 : Suivi <sup>1</sup> H-RMN de la dégradation de blocs de motifs connus par AlyA1 .....	270
Figure 7-11: Analyse en C-PAGE des produits de la dégradation d'alginate par AlyA1 .....	271

Figure 7-12 : Chromatographie d'exclusion de taille des produits terminaux de la dégradation par AlyA1.....	272
Figure 7-13 : Spécificité de reconnaissance des sous-sites d'AlyA1.....	274
Figure 7-14 : Action d'AlyA5 dimère sur les oligo-alginates saturés.....	276
Figure 7-15: Effet du pH sur l'activité d'AlyA5.....	277
Figure 7-16: Chromatographie d'exclusion de taille des oligosaccharides insaturés résultant de la dégradation partielle par AlyA1. ....	278
Figure 7-17: Dégradation d'oligosaccharides insaturés par AlyA5.....	278
Figure 7-18 : Suivi en <sup>1</sup> H-RMN de l'action d'AlyA5 sur les blocs polyG.....	280
Figure 7-19 : Suivi en <sup>1</sup> H-RMN de l'action d'AlyA5 sur l'oligosaccharide ΔGG.....	281
Figure 7-20 : Schéma d'interprétation de l'activité d'AlyA5 sur des oligogulonates saturés et insaturés.....	282
Figure 7-21 : Cristaux des deux alginates lyases recombinantes.....	283
Figure 7-22 : Clichés de diffraction obtenus pour les cristaux d'AlyA1 (A) et AlyA5 (B).....	284
Figure 7-23 : Comparaison des structures d'AlyA1 (A) et AlyA5(B) avec celle d'A1-II' de <i>Sphingomonas</i> sp.A1 (C) obtenue en complexe avec un oligosaccharide.....	288
Figure 7-24 : Evolution du facteur B des résidus le long de la séquence d'AlyA1.....	289
Figure 7-25 : Comparaison des surfaces des enzymes A1-II' et AlyA1.....	289
Figure 7-26 : Surface de l'enzyme AlyA5 de <i>Z. galactanivorans</i> . ....	290
Figure 7-27 : Etude de la flexibilité de la boucle couvrant le site actif d'AlyA5.....	291
Figure 7-28 : Visualisation du dimère d'AlyA5 le plus probable.....	292
Figure 7-29 : Mode de reconnaissance du substrat chez A1-II' de <i>Sphingomonas</i> sp.A1.....	293
Figure 7-30 : Modélisation du substrat dans le site catalytique d'AlyA1.....	294
Figure 7-31 : Modélisation du substrat dans le site catalytique d'AlyA5.....	295
Figure 7-32 : Mode spéculatif d'extraction des produits de réaction dans AlyA5.....	296
Figure 8-1 : Modèle hypothétique final du système alginolytique de <i>Z. galactanivorans</i> .....	301

# Liste des tableaux

Tableau 1-1 : Principaux polysaccharides de la paroi des plantes et des macroalgues .....	31
Tableau 1-2 : Proportion de motifs M et G dans des alginates de différentes sources.....	39
Tableau 1-3 : Masses d'algues récoltées dans le monde en 2001 .....	40
Tableau 1-4 : Les polysaccharidases caractérisées de <i>Z. galactanivorans</i> .....	50
Tableau 1-5 : Classification et caractéristiques des alginate lyases .....	62
Tableau 5-1 : Résultat de la recherche des sondes dans les régions codantes.....	183
Tableau 5-2 : Listes des gènes de fonction connue non représentés sur la puce Nimblegen .....	184
Tableau 5-3 : Résultat de la recherche de sondes dans les régions intergéniques.....	184
Tableau 5-4 : Listes des gènes de <i>Z. galactanivorans</i> fortement exprimés dans toutes les conditions .....	193
Tableau 5-5 : Répartition des cibles différentielles détectées lors de la comparaison de l'expression en présence d'alginate (référence = glucose) .....	197
Tableau 5-6 : Détection de termes GO sur-représentés parmi les gènes surexprimés en présence d'alginate .....	202
Tableau 5-7 : Familles de régulateurs transcriptionnels dont l'expression est induite en présence d'alginate .....	207
Tableau 5-8 : Détection de termes GO sur-représentés parmi les gènes sous-exprimés en présence d'alginate ou de laminarine.....	217
Tableau 6-1: Tableau récapitulatif des cibles clonées et des amorces oligonucléotidiques utilisées	230
Tableau 6-2 : Amorces universelles des vecteurs de clonage .....	235
Tableau 6-3 : Prédiction de peptides signaux et d'hélices transmembranaires dans les séquences..	237
Tableau 6-4 : Résultats du Blastp pour les deux modules d'AlyA1 .....	238
Tableau 6-5 : Résultats des différentes étapes du clonage et de l'expression à moyen débit .....	242
Tableau 7-1 : Propriétés des trois alginate lyases PL7 de <i>Z. galactanivorans</i> .....	251
Tableau 7-2 : Pourcentages d'identité de séquences entre les PL7 alignées.....	261
Tableau 7-3 : Effet d'agents chélatants sur l'activité d'AlyA1 .....	267
Tableau 7-4 : Constantes cinétiques d'AlyA1 .....	268
Tableau 7-5 : Proportions des différents oligosaccharides produits par AlyA1. ....	273
Tableau 7-6 : Comparaison des activités spécifiques d'AlyA1 et AlyA5 sur les polysaccharides d'alginate .....	275
Tableau 7-7 : Données de la collecte et statistiques de l'affinement pour AlyA1 et AlyA5.....	285
Tableau 7-8 : Analyse des interfaces protéine-protéine au sein du cristal d'AlyA5.....	292

---

# Liste des abréviations

---

$\Delta$  : acide uronique 4-désoxy-L-érythro-hex-4-enopyranosyluronique  
ADN : acide désoxyribonucléique  
ADNc : ADN complémentaire  
AMAC : 2-aminoacridone  
ANOVA : analyse de la variance  
APS : 5'-phosphosulfate  
ARN : acide ribonucléique  
ATP : adénosine triphosphate  
BLAST : *Basic Local Alignment Search Tool*  
BSA : albumine de sérum bovin  
CAZyme : enzyme active sur les sucres  
CBM : module de liaison aux sucres  
CE : Carbohydate Esterase  
CFU : unité formant colonie  
CM : membrane cytoplasmique  
COHV : composés organiques halogénés volatils  
C-PAGE : électrophorèse de carbohydrates sur gel de polyacrylamide  
DP : degré de polymérisation  
DEH : acide uronique 4-désoxy-L-erythro-5-hexoseulose  
DLS : diffusion dynamique de la lumière  
DMS : diméthylsulphide  
DMSO : diméthylsulfoxyde  
DNase : désoxyribonucléase  
dNTP : désoxyribonucléotide triphosphate  
DOC : carbone organique dissous  
DOM : matière organique dissoute  
EC : *Enzyme Commission*  
EDTA : acide éthylène diamine tétracétique  
EGTA : acide éthylène glycol tétracétique  
FACE : électrophorèse de carbohydrates couplés à un fluorophore  
FC : variation d'expression  
FDR : *False Discovery Rate*  
G : acide  $\alpha$ -1,4-L-guluronique  
GDP : guanosine diphosphate  
GDP-ManA : acide mannuronique-guanoside diphosphate  
GH : glycoside hydrolase  
GO : ontologie de gènes  
GST : glutathion-S-transférase  
HCA : méthode des amas hydrophobes  
HEPES : acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique  
HMM : modèle de Markov caché  
IPTG : isopropyl  $\beta$ -D-1-thiogalactopyranoside  
KDG : acide cétodésoxygluconique  
KDPG : acide cétodésoxyphosphogluconique  
LPS : lipopolysaccharide

M : acide  $\beta$ -1,4-D-mannuronique  
MOPS : acide 2-morpholinoéthanesulfonique  
NADPH : nicotinamide adenine dinucleotide phosphate  
OM : membrane externe  
ORF : cadre ouvert de lecture  
PAGE : électrophorèse sur gel de polyacrylamide  
PAPS : 3'-phosphoadénylyl sulfate  
PCA : analyse en composantes principales  
PCR : réaction de polymérisation en chaîne  
PDB : *Protein Data Bank*  
PG : peptidoglycane  
PGA : propylène glycol alginate  
PL : polysaccharide lyase  
POC : carbone organique particulaire  
POM : matière organique particulaire  
PUL : locus d'utilisation de polysaccharides  
RI : index de réfraction  
RMN : résonance magnétique nucléaire  
RNase : ribonucléase  
rpm : rotations par minute  
RT : transcription inverse  
RT-qPCR : PCR quantitative couplée à la transcription inverse  
SDS : sodium dodécylsulfate  
SF : sous-famille  
SPase : signal peptidase  
Sus : *starch utilization system*  
TEP : exopolymères transparents  
TBDR : récepteur TonB dépendant  
Tris : trishydroxyméthylaminométhane  
U : unité  
UV : lumière ultraviolette  
vHPO : haloperoxydase à vanadium



# Chapitre 1

## Introduction générale

---





## I- Ecologie microbienne marine

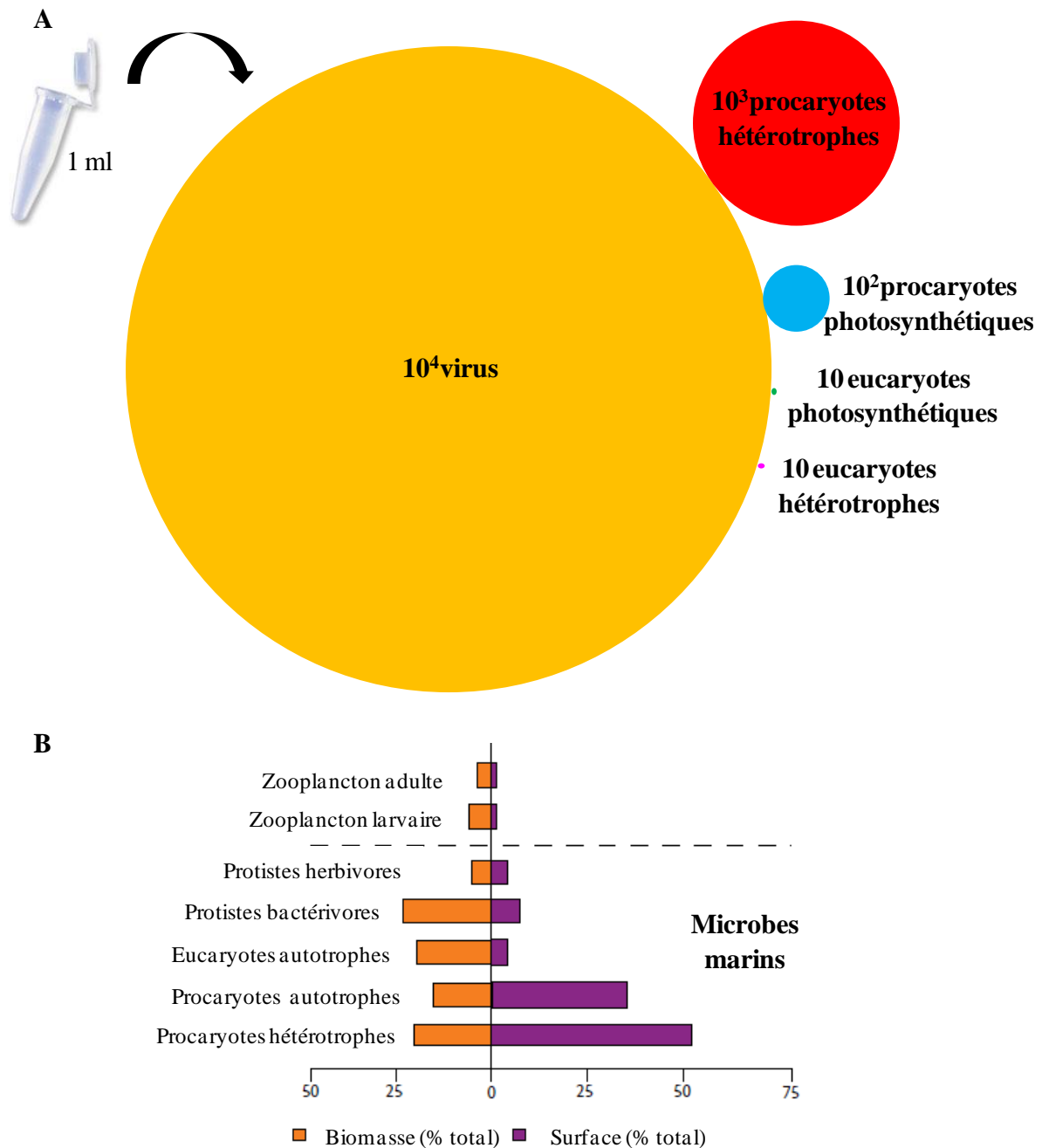
### I-1 Importance des bactéries hétérotrophes dans le cycle du carbone

Microbe marin : (*expression nominale, masculin*) organisme plus petit qu'environ 100  $\mu\text{m}$ , ne pouvant pas être vu à l'œil nu, et vivant dans un environnement marin.

Cette définition simple cache en fait une multitude d'êtres vivants connus et inconnus, appartenant à différents groupes évolutifs et jouant des rôles écologiques très variés, mais tous fondamentaux pour la structure et le fonctionnement des écosystèmes marins. Il s'agit de procaryotes (archées et bactéries) et d'eucaryotes unicellulaires (protistes). Les océans et mers du globe recouvrent environ 66% de la surface de la Terre et représentent des environnements très variés. Les microbes marins ont colonisé tous les habitats, des profondeurs du plancher océanique jusqu'aux premiers millimètres à la surface de l'eau. Ils prospèrent dans des conditions très diverses, du très froid (dans les canaux d'eau sur la banquise en Arctique) au très chaud (près des sources hydrothermales), sous des pressions importantes (sur le plancher océanique), dans des eaux très salées (Mer Morte) ou au contraire plus douces (Mer Baltique), dans des zones riches en nutriments, dites eutrophes (régions côtières) ou pauvres en nutriments, dites oligotrophes (océan ouvert), etc. Cette extrême versatilité fait des microbes marins le groupe d'organismes le plus abondant sur la planète, mais également explique leur implication dans pratiquement tous les processus biogéochimiques de l'océan (Kirchman 2008).

Le nombre total de procaryotes dans l'océan est estimé à  $10^{29}$  cellules, dont environ un tiers dans les premiers 200 mètres de la surface, et deux tiers en-dessous (Whitman *et al.*, 1998). Les densités cellulaires sont très élevées, variant de  $10^4$  à  $10^7$  cellules par millilitre. De manière simpliste, chaque millilitre d'eau de mer de surface contient 10000 virus, 1000 bactéries ou archées hétérotrophes, respectivement 100 et 10 cellules de bactéries photosynthétiques *Prochlorococcus* et *Synechococcus*, 10 eucaryotes unicellulaires photosynthétiques et 10 protistes (Azam and Malfatti 2007). Basée sur des estimations, la biomasse totale des microbes marins excède celle de tous les animaux et zooplancton réunis. Cette biomasse représente 95% de la respiration totale. Si on considère de plus que c'est par leur surface que les organismes interagissent avec le milieu, et donc que ce paramètre est plus pertinent que la masse pour évaluer la contribution relative au fonctionnement d'un

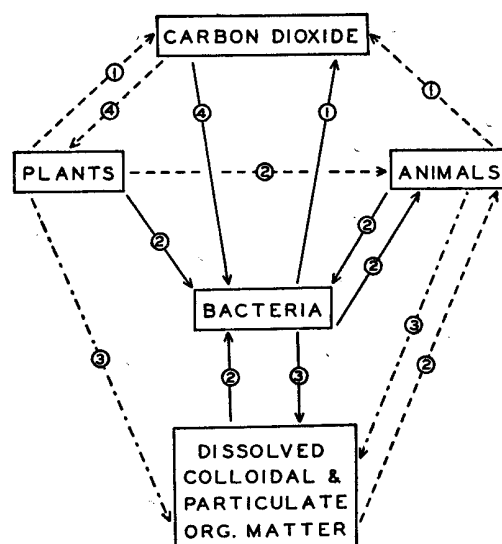
écosystème, on constate même que les bactéries surpassent tous les autres groupes de microbes (Figure 1-1).



**Figure 1-1: Abondance et distribution des différents groupes microbiens marins**

**A.** Abondance des différents groupes dans 1 ml d'eau de mer de surface. **B.** Distribution de la biomasse et de la surface des différents groupes de la fraction planctonique dans la zone euphotique de l'océan. Les calculs sont basés sur des estimations moyennes dans différentes régions océaniques. La surface est calculée en admettant une géométrie sphérique des cellules. Adapté de Pomeroy *et al.* (2007)

Un des grands enjeux de l'écologie marine est de comprendre les flux de matière organique entre les différents compartiments de l'écosystème. En effet, la moitié de la production primaire globale a lieu dans les océans (Field *et al.*, 1998). Ce phénomène, dont la photosynthèse est un exemple bien connu, permet l'incorporation sous forme de molécules organiques du carbone contenu dans le CO<sub>2</sub>. Il est réalisé par les organismes phototrophes (bactéries, plantes ou algues) et certains microorganismes chimiosynthétiques. Une petite partie de cette matière organique (~2%) est séquestrée dans les sédiments océaniques et a contribué à faire baisser le taux de dioxyde de carbone dans l'atmosphère, tout en fournissant la fameuse énergie fossile (Hedges *et al.*, 2001). Cependant, la grande majorité de cette matière organique est recyclée au travers de chaînes alimentaires. Des travaux pionniers par Waksman (1888 – 1973) et ZoBell (1904 – 1989) ont permis d'appréhender l'importance des bactéries dans la décomposition de la matière organique (McGraw 2006). Ces deux auteurs ont bâti des modèles de fonctionnement des écosystèmes pélagiques, où les groupes fonctionnels d'organismes sont représentés par une boîte, avec des entrées et des sorties de matière (Figure 1-2). Même si les méthodes de l'époque ne permettaient pas encore de tester expérimentalement les modèles, ces approches ont posé les bases de la compréhension du rôle prépondérant des bactéries hétérotrophes dans les flux de carbone.

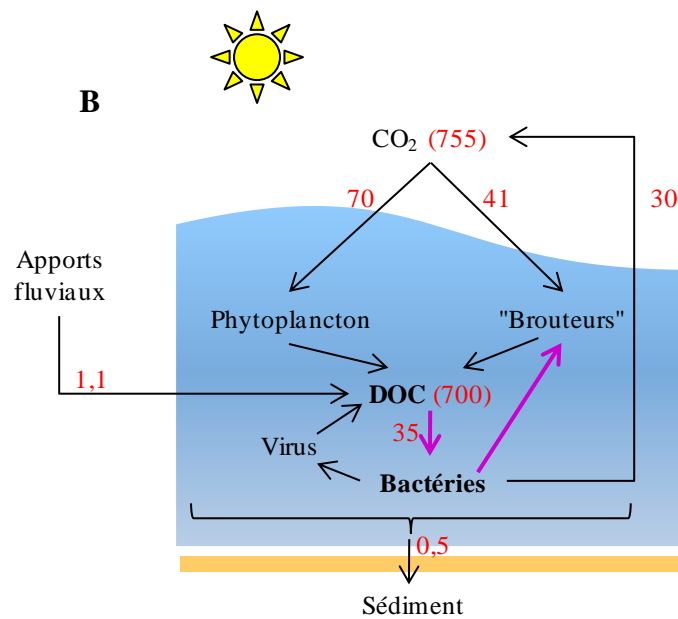
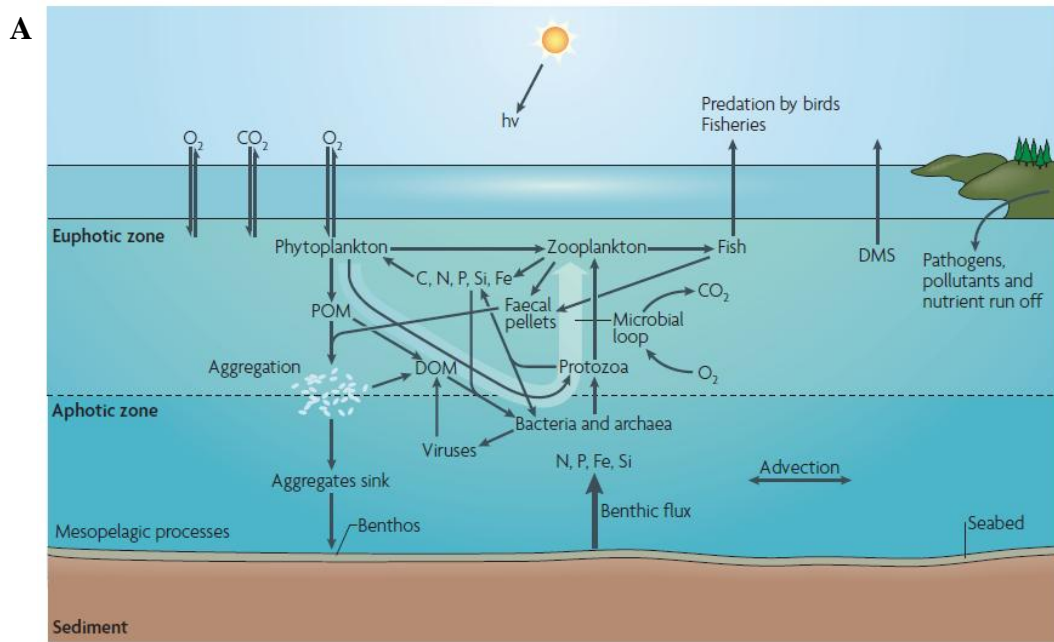


**Figure 1-2: Le cycle du carbone dans la mer, selon ZoBell (1946)**

Reproduction de la Figure 10 du chapitre 10 "Transformation de la matière organique" dans le manuel *Marine Microbiology* (ZoBell 1946). Dans ce schéma, les lignes continues représentent des processus où les bactéries sont les seuls organismes impliqués, les lignes pointillées doubles ceux où les bactéries sont susceptibles de participer et les lignes pointillées simples ceux auxquels les bactéries ne participent pas. 4 types de processus sont pris en compte: (1) respiration, (2) nutrition, (3) décomposition, (4) fixation du CO<sub>2</sub>.

D'après Zobell (1946), la fonction majeure des bactéries dans le cycle du carbone est la décomposition de la matière organique en CO<sub>2</sub>, eau, ammonium et minéraux. Une seconde fonction est la conversion de la matière organique présente dans l'eau de mer en "substance cellulaire" qui peut ensuite être assimilée par exemple par des organismes filtreurs. Troisièmement, certaines bactéries de la flore intestinale (microbiome) d'animaux marins peuvent participer à la digestion de la nourriture qu'ils ingèrent. Enfin, les bactéries autotrophes photo- ou chimiosynthétiques sont des producteurs primaires.

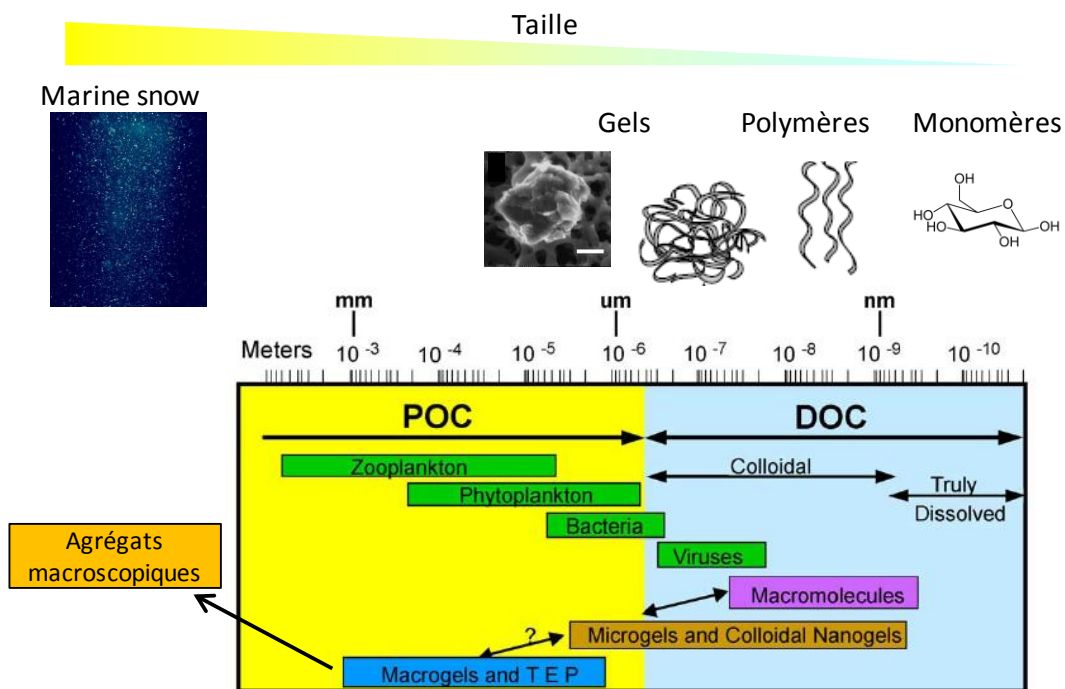
Le diagramme de ZoBell (Figure 1-2) souligne l'existence de la matière organique sous deux formes: dissoute (*Dissolved Organic Matter*, DOM) ou particulaire (*Particulate Organic Matter*, POM). Techniquement, la DOM est définie comme la fraction du matériel organique qui peut traverser les pores d'un filtre d'environ 0,7 µm, et la POM celle qui est retenue. Par conséquent, la DOM peut contenir en plus de la matière véritablement dissoute des colloïdes, des virus, de petites bactéries ou archées, ainsi que certaines particules de faible diamètre (Kirchman 2002). La DOM contient un pool de carbone réduit (*Dissolved Organic Carbon*, DOC), mais aussi d'autres éléments importants tels que l'azote, le phosphore ou le soufre. La concentration de ces différents éléments est de l'ordre du micro- au nanomolaire. La quantité totale de carbone dissous dans l'océan est estimée à environ  $7.10^{17}$  g (Hansell *et al.*, 2009), soit 200 fois plus que la biomasse marine totale, et autant que de CO<sub>2</sub> dans l'atmosphère. Les microbes ont un rôle crucial pour le recyclage de la DOM. En effet, ce gigantesque stock de matière n'est pratiquement accessible qu'aux bactéries et archées hétérotrophes. Les microbes ayant assimilé le carbone dissous sont à leur tour soit lysés par des virus soit consommés par des protistes, eux-mêmes à la base du régime alimentaire d'animaux marins. L'ensemble de ces processus microbiens de transfert de matière et d'énergie de la DOM vers les niveaux trophiques supérieurs est connu sous le nom de boucle microbienne (Azam *et al.*, 1983; Azam 1998) (Figure 1-3 A). On sait aujourd'hui que cette boucle est indispensable au recyclage du carbone présent dans la DOM, qui sinon serait complètement perdu pour le réseau trophique. Elle représente un flux de  $35.10^{15}$  g de carbone par an (Figure 1-3 B), et on estime que 40% de la matière organique produite par les producteurs primaires ("*Phytoplankton*" dans la Figure 1-3) sont recyclés via cette boucle.



**Figure 1-3 : Le rôle des bactéries dans le cycle du carbone océanique**

**A.** Structure schématique d'un écosystème océanique ouvert, qui met en évidence le rôle de la boucle microbienne (flèche blanche épaisse) dans le recyclage de la matière organique. POM : *Particulate Organic Matter*. DOM : *Dissolved Organic Matter*. DMS : Diméthylsulphide. Schéma copié de Azam et Malfatti (2007). **B.** Flux de carbone entre différents compartiments de l'écosystème. Les valeurs entre parenthèses représentent des stocks de carbone (en  $10^{15}$  g). Les autres valeurs sont des flux en  $10^{15}$  g par an. La boucle microbienne est mise en évidence par les flèches épaisses violettes. DOC : *Dissolved Organic Carbon*. Adapté de Kirchman (2008).

En dehors de la dégradation par les bactéries, la DOM peut subir des phénomènes d'agrégation, pour aboutir à des particules jusqu'à plusieurs millimètres de diamètre. Il existe en fait un continuum de taille entre la matière réellement dissoute et les agrégats (Nagata 2008) (Figure 1-4). La transition entre ses deux états est réalisée par la formation de gels colloïdes, dont la masse représente environ 10% de la DOM, soit  $70 \cdot 10^{15}$  g de carbone. La présence de polysaccharides gélifiants dans la DOM favorise ainsi l'apparition de microgels dans la colonne d'eau, qui sont des réseaux tridimensionnels de biopolymères (Verdugo *et al.*, 2004). Ces gels peuvent se former spontanément dans la DOM en quelques minutes à plusieurs heures (Chin *et al.*, 1998). L'eau piégée à l'intérieur de ces gels les empêche de collapser et constitue également un micro-environnement en équilibre thermodynamique avec l'extérieur. Tous les polysaccharides gélifiants, synthétisés par exemple par des bactéries ou des algues, sont susceptibles de former ces microgels avec parfois la nécessité de cations comme  $K^+$  ou  $Ca^{2+}$ . En plus de représenter des réservoirs de carbone, ces gels colloïdes piègent de nombreux composés inorganiques dissous dans l'eau de mer.



**Figure 1-4 : Le continuum de taille de la matière organique dans l'eau de mer**

Les tailles des différentes formes de la matière organique, des monomères totalement dissous aux macrogels, sont données sur une échelle logarithmique. Ces macrogels peuvent être à l'origine de la formation de neige marine (*marine snow*). Par comparaison, les gammes de tailles des microbes marins sont indiquées dans les cadres verts. POC : *Particulate Organic Carbon*. DOC : *Dissolved Organic Carbon*. TEP : *Transparent Exopolymers*.

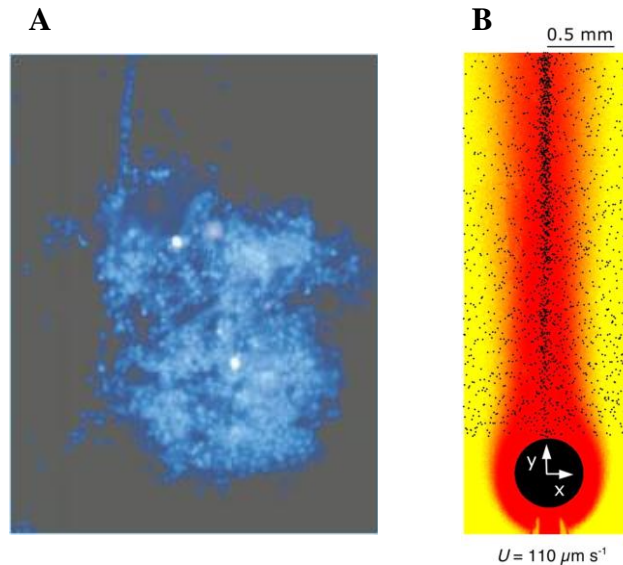
Les macrogels peuvent s'agréger pour former de petites particules, qui englobent à la fois des organismes vivants et des détritiques (Figure 1-4). Les plus petites particules sont constituées de micro-agrégats, de fèces, de phytoplancton et de débris organiques. Elles ont la propriété de pouvoir s'associer pour former de plus grosses particules. Ce processus d'agrégation est favorisé notamment par la présence d'exsudats, d'exopolymères ou de débris résultant d'une lyse cellulaire, qui peuvent agir comme une "colle" et former une matrice dans laquelle les particules sont engluées. Ces particules jouent un rôle prépondérant dans les cycles de matière organique car elles ont tendance à sédimenter dans la colonne d'eau. Ce phénomène est un acteur de la "pompe biologique", qui contribue à la séquestration d'une partie du carbone organique dans les profondeurs de l'océan. Le flux de carbone associé est estimé à  $10^{16}$  g par an (Nagata 2008). Ces particules sont aussi connues sous le nom de neige marine (*marine snow*) car elles tombent au fond de l'océan telles des flocons (Alldredge and Silver 1988).

Toutes ces particules, des microgels aux neiges marines, sont des *hotspots* de matière organique et sont rapidement colonisées par des communautés bactériennes (Kiorboe *et al.*, 2003). Les densités de population sur les neiges marines peuvent atteindre  $10^8$ - $10^9$  cellules par ml (Figure 1-5 A). Comme pour la DOM, l'activité de dégradation de la POM par les bactéries hétérotrophes joue un rôle crucial dans le cycle du carbone. L'expression d'enzymes de dégradation de la matière agrégée (polysaccharides, protéines, etc.) leur permet d'assimiler une partie de la ressource carbonée piégée. D'autre part, la dégradation de la particule diminue sa densité et donc abaisse sa vitesse de sédimentation, ce qui minore le taux de carbone séquestré dans les profondeurs de l'océan. Par ailleurs, tous les produits de la dégradation enzymatique ne sont pas consommés par les bactéries fixées sur la particule. Elle laisse donc derrière elle un panache de matière organique dissoute, qui peut alimenter les communautés microbiennes libres dans la colonne d'eau (Figure 1-5 B).

Les petits composés (< 600 Da) peuvent rentrer dans les cellules directement grâce à l'expression à leur surface de protéines de transport. Cela concerne par exemple les monosaccharides ou les acides aminés libres. Par contre, les composés de plus haut poids moléculaire (> 1000 Da) doivent d'abord être dégradés par des enzymes spécifiques avant d'être internalisés (Kujawinski 2011). Les enzymes impliquées peuvent être soit sécrétées, soit ancrées à la membrane externe. Le rôle des polysaccharidases dans l'acquisition du carbone est prépondérant, puisqu'il a été montré que les polysaccharides pouvaient représenter jusqu'à 50% des composés de haut poids moléculaire dans les eaux de surface, et environ 25% en



profondeur (Benner *et al.*, 1992). L'expression de ces différents acteurs cellulaires de la dégradation est finement régulée. En effet, les bactéries sont capables de détecter la présence de matière organique, et de s'y adapter. Il a par exemple été montré que des bactéries marines "sentent" le panache de matière dissoute qui suit les neiges marines en sédimentation, et sont capables de nager vers ce *hotspot* de nutriments par un mécanisme de chimiotactisme (Stocker *et al.*, 2008) (Figure 1-5 B).



**Figure 1-5 : Les neiges marines, des *hotspots* de la colonisation bactérienne.**

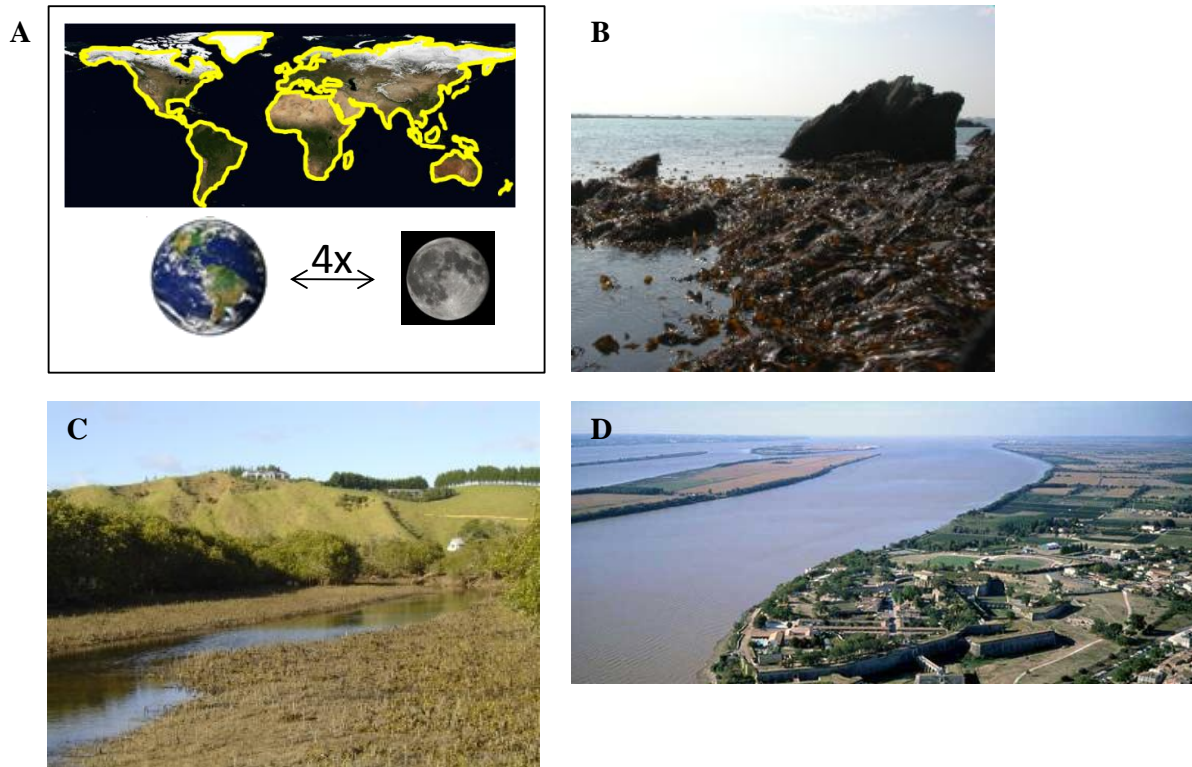
A. Une neige marine collectée dans la baie de Monterey et observée au microscope à épifluorescence. Les bactéries ont été colorées au préalable par le fluorochrome DAPI, d'où la couleur bleue. La taille de la particule est de  $37 \times 27 \mu\text{m}$ . Image copiée de Azam et Long (2001). B. Illustration de la réponse chimiotactique des bactéries à la présence d'un panache de nutriments derrière une neige marine. Les points noirs représentent les bactéries, qui se concentrent là où la disponibilité en nutriments est la plus élevée (en rouge) derrière la particule qui coule à  $110 \mu\text{m.s}^{-1}$ . Image copiée de Stocker *et al.* (2008).

Par ailleurs, il semble que des mécanismes de *quorum sensing* puissent contrôler l'expression de différents caractères chez les bactéries résidant sur les particules, tels que l'adhésion ou la production d'enzymes (Gram *et al.*, 2002). L'utilisation du *quorum sensing* pourrait permettre aux bactéries de ne commencer à dégrader la particule que lorsque leur densité cellulaire atteint un certain seuil. On sait également aujourd'hui que les différentes espèces de bactéries marines qui dégradent la matière organique ne possèdent pas toutes les mêmes capacités métaboliques, et notamment sécrètent différentes enzymes (Martinez *et al.*, 1996). Ainsi, la composition qualitative de la DOM ou POM pourrait favoriser une population par rapport à une autre dans les communautés microbiennes. Des phyla bactériens entiers semblent spécialisés dans la prise en charge d'un type de matière. Par exemple, le phylum des

*Bacteroidetes* est considéré comme spécialiste de la dégradation de la matière de haut poids moléculaire, notamment les polysaccharides (cf. Chapitre 2). La composition de la matière organique influence également l'expression des gènes chez les bactéries. Ceci révèle des mécanismes de perception finement contrôlés afin de répondre le plus efficacement possible à la présence de nutriments dans le milieu hautement compétitif qu'est l'environnement marin. Ainsi, il a été montré que l'ajout de DOM de haut poids moléculaire à des populations microbiennes naturelles induit la transcription de gènes spécifiques, impliqués dans des systèmes de perception, de chimiotaxie, de mobilité et du métabolisme des sucres, des acides gras, du phosphate ou de l'azote (McCarren *et al.*, 2010). Dans une autre étude, Poretsky *et al.* (2010) ont également mis en évidence que l'expression des transporteurs localisés dans les membranes des bactéries est contrôlée par la nature et la source de la DOM (ex. dérivée de phytoplancton ou de plantes vasculaires).

## **I-2 Les régions côtières : des écosystèmes particuliers**

Les régions côtières sont à l'interface entre le continent et l'océan, et peuvent donc être considérées comme des zones "tampons" ayant des apports de ces deux écosystèmes. Il n'existe pas de définition unique d'une zone côtière. Certains auteurs, revenant sur ce concept de zone "tampon", les définissent comme "la partie du continent qui est la plus affectée par sa proximité avec la mer, et la partie de l'océan qui est la plus affectée par sa proximité avec la terre". D'autres études donnent un cadre plus précis comme "les aires intertidales et subtidales du plateau continental (jusqu'à 200 m de profondeur) qui sont régulièrement inondées par l'eau de mer, et les terres immédiatement adjacentes" (Burke *et al.*, 2001). Quelle que soit la définition choisie, les régions côtières sont des zones particulières sur la Planète. Bien qu'elles ne représentent que 7% de la surface des océans et 20% de la surface des continents, elles sont peuplées de 2,2 milliards de personnes, soit 39% de la population mondiale (en 1995, à moins de 100 km d'une côte) et ce chiffre est en augmentation (Burke *et al.*, 2001). La longueur cumulée des traits de côtes est estimée à 1,6 million de km, soit 4 fois la distance de la Terre à la Lune...! Ceci regroupe évidemment des environnements très différents, tels que les côtes bretonnes, les mangroves, les barrières de coraux, les estuaires, etc. (Figure 1-6). Toutes ces régions, d'une remarquable productivité et accessibilité, ont attiré l'Homme depuis des millénaires. Elles sont aujourd'hui à la base d'innombrables activités professionnelles ou de loisirs, et soutiennent donc l'activité économique.



**Figure 1-6 : L'étendue et la diversité des régions côtières.**

A. Représentation schématique de la longueur des régions côtières, qui représente 4 fois la distance Terre-Lune. B. Les Duons, Baie de Roscoff. C. Mangrove dominée par l'espèce *Avicenna marina*, Matapouri, Nouvelle-Zélande. D. Estuaire de la Gironde.

La zone côtière joue un rôle considérable dans les cycles biogéochimiques pour plusieurs raisons (Gattuso *et al.*, 1998; Gazeau *et al.*, 2004). Tout d'abord, comme il a déjà été mentionné, c'est une zone tampon. Elle reçoit par conséquent des apports massifs de composés organiques d'origine terrestre, et échange des quantités considérables de matière et d'énergie avec l'océan ouvert. Elle constitue par ailleurs l'une des zones les plus productives de la planète, tant en termes géochimiques que biologiques.

Les écosystèmes côtiers renferment des communautés végétales abondantes, comprenant des macro-algues (dominantes sur les côtes rocheuses, les estuaires et les barrières de coraux) et des plantes marines (dominantes dans les systèmes intertidaux tels que les mangroves, les marais salants et les zones sableuses). L'étendue de ces communautés végétales est estimée à 2 millions de km<sup>2</sup> (Whittaker and Likens 1973), et elles peuvent contribuer jusqu'à 50% de la production primaire dans les zones non turbides où la lumière pénètre facilement (Gattuso *et al.*, 1998). Il a été montré que les zones côtières dominées par des végétaux sont des sites privilégiés de séquestration de matière organique. En effet, le temps de renouvellement des végétaux marins est relativement long par rapport au

phytoplancton (de l'ordre de l'année par rapport à quelques jours), et leur biomasse jusqu'à 400 fois plus importante. Les composants de cette biomasse sont donc moins rapidement ré-injectés dans les réseaux trophiques (Smith 1981). Par conséquent, les zones côtières sont considérées comme des puits de carbone, retenant environ  $200 \cdot 10^{12}$  g de carbone par an (Duarte *et al.*, 2005). Cependant, une grande partie de la matière organique produite par les végétaux marins est exportée vers le large soit sous forme détritique, soit après dégradation par des organismes hétérotrophes. Cet apport de matière des zones côtières vers l'océan ouvert pourrait représenter jusqu'à  $3 \cdot 10^{15}$  g de carbone par an. Le régime alimentaire d'un grand nombre d'animaux marins herbivores (poissons, invertébrés) est basé sur la consommation de ces végétaux, qui supportent donc le réseau trophique local. Les bactéries hétérotrophes jouent également un rôle prépondérant dans la dégradation de cette matière organique et sa ré-injection dans le réseau trophique. Par exemple, dans les forêts de laminaires typiques de la baie de Roscoff, il a été montré que la voie détritique est le principal mode de transfert de matière vers des niveaux trophiques supérieurs, plus important que la consommation directe par les herbivores (Schaal *et al.*, 2009). La prévalence de débris d'algues brunes dans la POM en suspension dans ces zones est principalement due à la colonisation de ces particules par des bactéries spécialisées capables notamment de dégrader les polysaccharides complexes. On sait par ailleurs que les tissus d'algues frais ne sont pas une source de nourriture appropriée pour la plupart des consommateurs, à cause de leur contenu pauvre en azote et de la présence de métabolites secondaires qui repoussent les herbivores. Les processus de dégradation bactériens augmentent la valeur nutritionnelle des tissus, qui peuvent ensuite être ingérés par des organismes filtreurs (Norderhaug *et al.*, 2003; Schaal *et al.*, 2010).

Les éléments présentés dans cette partie soulignent donc l'importance des bactéries hétérotrophes dans les cycles biogéochimiques océaniques. Par leur activité de dégradation de la DOM et de la POM, ces organismes contrôlent la quantité de carbone réinjectée dans les chaînes alimentaires ou au contraire séquestrée dans les sédiments. Ce phénomène, prépondérant pour l'équilibre du système océanique, est finement contrôlé. L'élucidation de ces mécanismes est une étape clé de la compréhension du rôle joué par les bactéries hétérotrophes dans le fonctionnement des écosystèmes marins. Pour ce faire, deux approches sont possibles. On peut étudier les communautés microbiennes naturelles et tenter de déchiffrer comment elles interagissent avec la matière organique, par des techniques de métagénomique ou métatranscriptomique par exemple (Moran 2008; Moran 2009). De manière complémentaire, on peut également s'intéresser à la dégradation de la matière par des

organismes modèles, dont la culture est maîtrisée en laboratoire. La dissection fine des mécanismes de perception, de régulation et de dégradation par de tels organismes permet de bâtir des modèles qui facilitent l'interprétation des données sur les communautés microbiennes naturelles.

Dans le cadre de ce travail de thèse, je me suis intéressé une bactérie marine hétérotrophe cultivée au laboratoire : *Zobellia galactanivorans* et à l'utilisation de la matière organique issue d'algues brunes.

## **II- La paroi des algues brunes**

### **II-1 Place des algues brunes dans l'évolution des Eucaryotes**

Les organismes communément dénommés "algues" représentent en fait un assemblage très divers d'organismes autotrophes. On compte parmi les algues des lignées eucaryotes unicellulaires (microalgues) et multicellulaires (macroalgues). Les caractéristiques communes des macroalgues sont la multicellularité, la photosynthèse, l'absence de tiges, de feuilles ou de racines et l'environnement exclusivement marin. Cependant, les macroalgues ne peuvent être considérées comme un groupe d'organismes monophylétique. On distingue ainsi trois lignées : les algues vertes Chlorophyta, les algues rouges Rhodophyta et les algues brunes Phaeophyceae. Les algues rouges et vertes appartiennent au phylum des Archaeplastida qui contient également les plantes terrestres (Figure 1-7). Les algues brunes appartiennent quant à elles au phylum des Straménopiles, qui du point de vue de l'évolution est aussi distant des Archaeplastida que des animaux (Baldauf 2008). En plus des algues brunes, on trouve les oomycètes et les diatomées dans ce phylum.

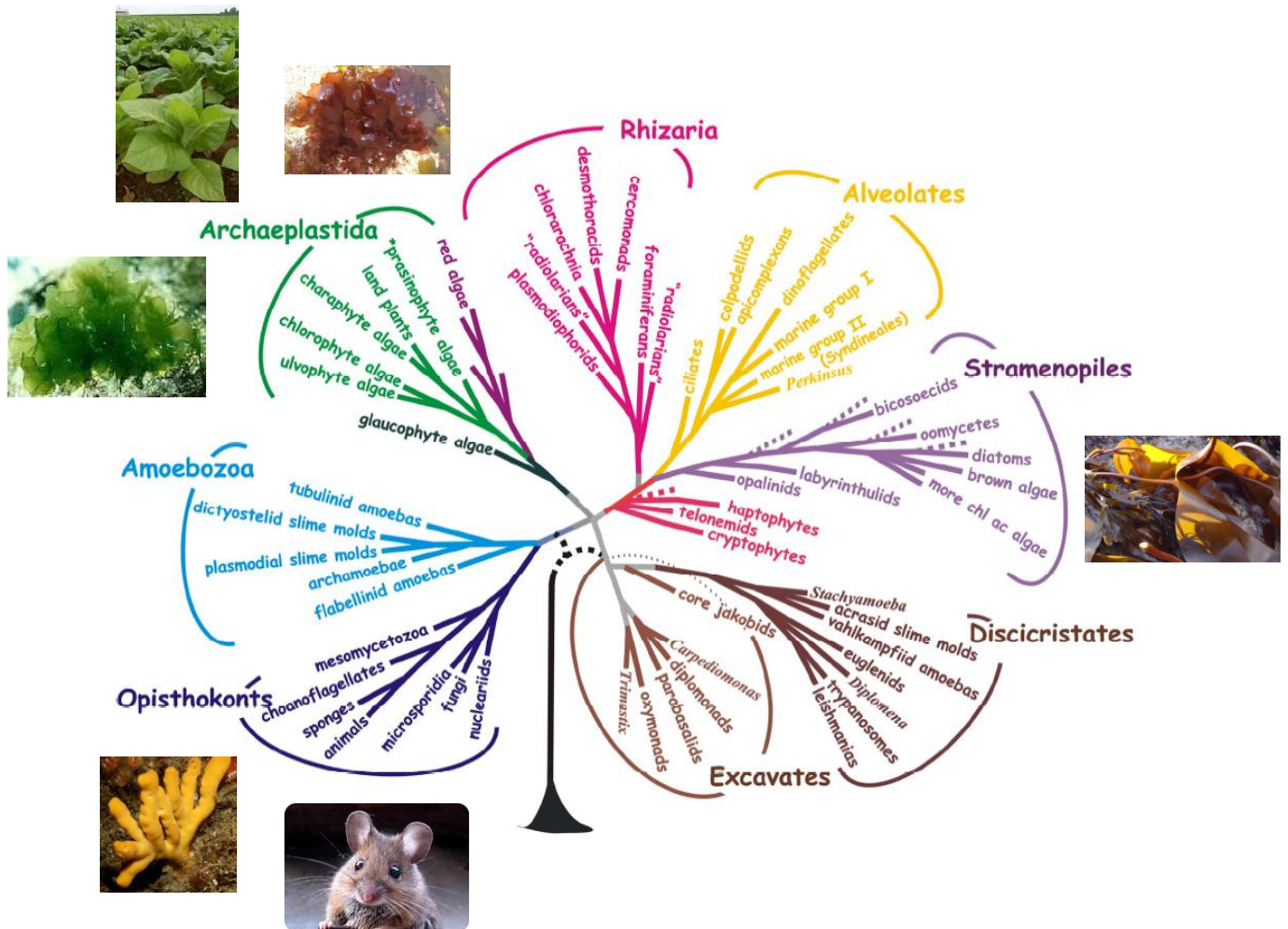


Figure 1-7 : Arbre phylogénétique des Eucaryotes (Baldauf 2008)

L'émergence du groupe eucaryote des Straménopiles est datée à environ un milliard d'années. Elle résulte d'un phénomène connu sous le nom d'endosymbiose secondaire, au cours duquel une algue rouge unicellulaire a été capturée par un protiste ancestral (Reyes-Prieto *et al.*, 2007). La radiation des algues brunes daterait quant à elle de 200 millions d'années. Au sein des Straménopiles, les algues brunes sont les seules à avoir développé une multicellularité complexe. Le développement d'une matrice extracellulaire adhérente a été proposé comme un des évènements évolutifs majeurs ayant permis le développement d'organismes multicellulaires (Popper *et al.*, 2010).

La position phylogénétique particulière des algues brunes en fait des modèles d'étude particulièrement intéressants. Les parois des algues brunes ont été largement étudiées, notamment pour les informations qu'elles peuvent apporter sur l'évolution de la multicellularité chez les êtres vivants, mais aussi pour leurs applications industrielles.

## II-2 Les parois végétales : considérations générales

Comme chez les plantes terrestres, les cellules des macroalgues sont entourées d'une paroi. Ce compartiment extracellulaire complexe peut être décrit comme un système à deux phases : une phase cristalline squelettique baignant dans une phase plus amorphe, la matrice. Que ce soit chez les végétaux terrestres ou marins, la paroi remplit des fonctions analogues. Elle joue un rôle dans la croissance et la forme des organismes, la résistance mécanique, la régulation osmotique ou ionique. Toujours en contact avec le milieu extérieur, elle sert d'interface entre le végétal et son environnement. Elle est ainsi le siège de phénomènes de perception / signalisation et est en première ligne lors des réactions de défense face à des stress biotiques (herbivores, pathogènes, etc.) ou abiotiques (rayonnement ultraviolet, dessiccation, etc.). Cependant, même si leurs fonctions sont similaires, la composition des parois des végétaux terrestres et marins présente des différences majeures (Kloareg and Quatrano 1988). Ainsi, les composants de la phase matricielle dominent largement la composition des parois des algues par rapport à la phase squelettique. Chez les plantes terrestres, les polysaccharides de la phase squelettique sont plus abondants. Ces caractéristiques peuvent être reliées à l'habitat des deux types de végétaux. La paroi des plantes terrestres, vivant à l'air libre, doit être plus rigide pour compenser les effets de la gravité. Ceci explique l'abondance de constituants de la phase squelettique, qui s'organisent en un solide réseau. Au contraire, les algues vivent en milieu aquatique et sont donc portées par la poussée d'Archimède. La dominance de la phase amorphe confère aux thalles d'algues une plus grande flexibilité, leur permettant de résister aux stress mécaniques que constituent les courants et les battements des vagues. Les cycles de marées les soumettent par ailleurs à des périodes de dessiccation fréquentes, qui peuvent durer jusqu'à plusieurs heures pour les individus vivant en haut de l'estran. Les algues doivent donc retenir un maximum d'eau dans leurs tissus, d'où une plus grande proportion relative de la phase amorphe. On sait par exemple que l'alginate, un des constituants majoritaires de la phase matricielle chez les algues brunes, peut absorber jusqu'à 140 fois son volume en eau. Une autre différence majeure entre la paroi des plantes terrestres et des algues réside dans la nature même des composés pariétaux. Ainsi, les polysaccharides matriciels des algues sont majoritairement anioniques, avec une contribution mineure des polysaccharides neutres (Tableau 1-1). Les algues possèdent également des polysaccharides sulfatés, qui sont totalement absents des plantes terrestres. Cependant, des galactanes sulfatés sont présents dans la paroi d'angiospermes



marines, qui sont de la même lignée que les plantes terrestres (Aquino *et al.*, 2005). La biosynthèse de polysaccharides sulfatés semble être une adaptation au milieu marin, dont la concentration en sulfate est de 28 mM alors qu'elle est de l'ordre de 10  $\mu$ M à 1 mM dans l'eau douce et dans l'eau interstitielle des sols. Au niveau évolutif, il a été proposé que la capacité de synthèse de polysaccharides sulfatés ait été perdue par les plantes lors de colonisation du milieu terrestre (Michel *et al.*, 2010).

		<b>Plantes terrestres</b>	<b>Chlorophyta Algues vertes</b>	<b>Rhodophyta Algues rouges</b>	<b>Phaeophyceae Algues brunes</b>
<b>Phase cristalline</b>		Cellulose	Cellulose	Cellulose Mannane Xylane	Cellulose
<b>Phase matricielle</b>	<b>Neutre</b>	Hémicelluloses	Hémicelluloses	Hémicelluloses	Hémicelluloses
	<b>Acide</b>	Pectines	Ulvanes		Alginates
	<b>Sulfaté</b>	-	Ulvanes	Agars Carraghénanes Porphyrane	Fucanes

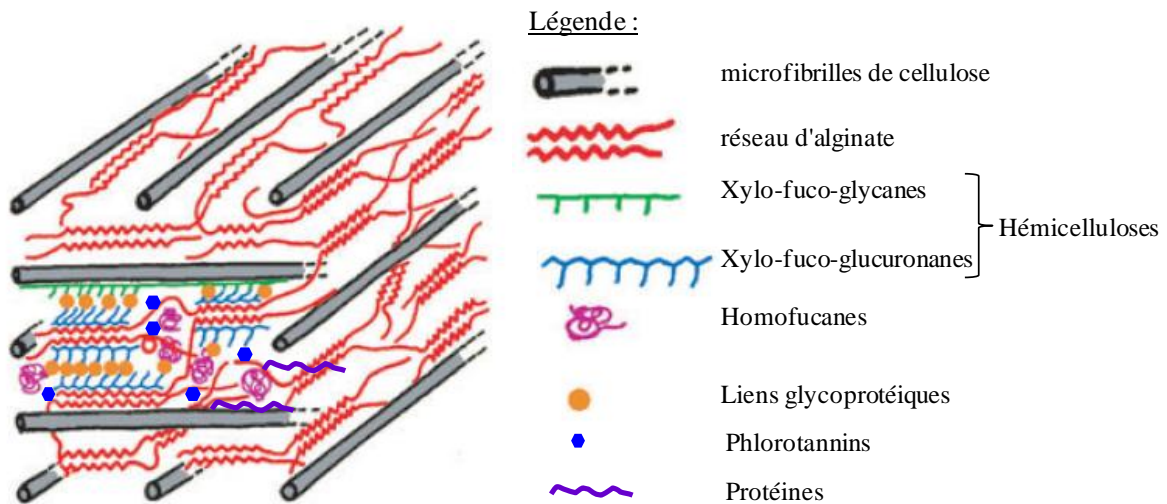
**Tableau 1-1 : Principaux polysaccharides de la paroi des plantes et des macroalgues**

D'après Popper *et al.* (2010)



### II-3 Architecture de la paroi des algues brunes

La paroi des algues brunes peut être modélisée comme un réseau tridimensionnel de microfibrilles de cellulose cristalline baignant dans une phase gélifiée moins structurée composée principalement d'un réseau d'alginate (Figure 1-8).



**Figure 1-8 : Modèle hypothétique de la paroi d'une algue brune.**

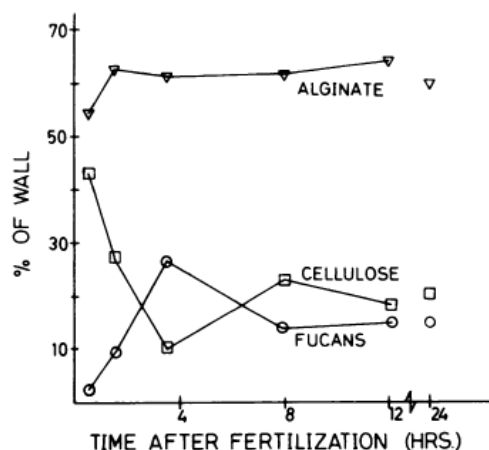
Image adaptée de Michel *et al.* (2010) d'après Kloareg et Quatrano (1988)

La phase squelettique des algues brunes est principalement constituée de cellulose, un polymère de  $\beta$ -1,4 glucane. Le taux de cellulose est compris entre 1 et 8% du poids sec du thalle des algues brunes, alors que cette proportion est classiquement de 30% chez les plantes terrestres (Cronshaw *et al.*, 1958). Des études menées sur des parois d'algues brunes isolées ont confirmé la faible proportion de cellulose. Le contenu en cellulose d'échantillons provenant de zygotes de *Fucus* et d'algues adultes (Laminariales et Fuciales) ne dépasse pas les 20% du poids sec (Quatrano and Stevens 1976; Kloareg 1984; Mabeau and Kloareg 1987).

La phase amorphe de la paroi des algues brunes est constituée majoritairement d'alginate, un polymère d'acides uroniques dont la structure sera présentée en détail dans la suite de ce chapitre. Selon l'espèce considérée, les alginates peuvent représenter de 10 à 45% du poids sec d'un thalle d'algue brune (Kloareg and Quatrano 1988). La paroi est également constituée de fucanes, des polysaccharides sulfatés composés de résidus de fucose. Il existe différents types de fucanes. Ceux regroupés sous le nom d'hémicelluloses sont très branchés, possèdent peu de groupements sulfate et contiennent une proportion significative d'autres sucres tels que

le xylose, le galactose et le mannose. Les homofucanes sont plus homogènes, beaucoup plus sulfatés et fortement branchés. Le nombre et la position des groupements sulfate varient en fonction des espèces des algues (Michel *et al.*, 2010). Le ratio massique moyen de l'alginate, des fucanes et de la cellulose est typiquement de 3 : 1 : 1 chez les algues brunes adultes (Mabeau and Kloareg 1987). Au sein de la paroi, les hémicelluloses assurent l'ancrage entre le réseau d'alginate et les microfibrilles de cellulose grâce à des liens de type glycoprotéique (Figure 1-8). De plus, le réseau pariétal renferme des phlorotannins, qui sont des composés phénoliques halogénés et/ou sulfatés, et environ 5% de protéines (Quatrano and Stevens 1976; Schoenwaelder and Wiencke 2000).

L'assemblage initial de la paroi a été étudié chez des zygotes d'algues et chez des protoplastes, qui sont des cellules dont la paroi a été digérée. Chez *Laminaria japonica*, la régénération de la paroi des protoplastes est détectable après deux jours, et complète au bout d'une semaine (Sawabe and Ezura 1996). Les embryons de *Fucus* ne possèdent pas de paroi avant la fertilisation. Il a été montré que la synthèse de la paroi débute dès 40 minutes après la fertilisation. Durant les deux premières heures, il y a un rapide dépôt d'alginate (Novotny and Forman 1975). Au cours de la génération de la paroi, l'alginate et la cellulose sont synthétisés en premier (au cours de la première heure), puis des fucanes sulfatés sont ajoutés au détriment de la cellulose (Figure 1-9). Le taux d'alginate reste quant à lui élevé et constant tout au long du processus (Quatrano and Stevens 1976).



**Figure 1-9 : Changements dans la composition en polysaccharides de parois isolées de *Fucus vesiculosus* à différents temps après la fertilisation de l'embryon**

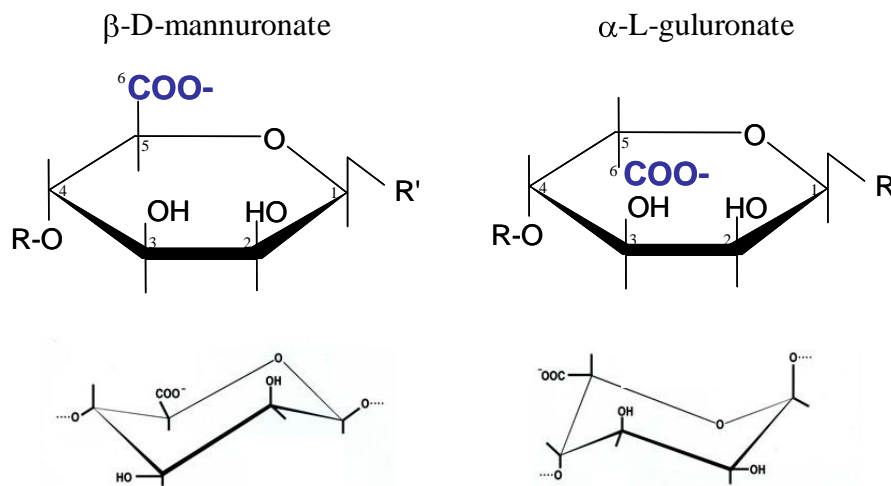
Image copiée de Quatrano et Stevens (1976)

Les données présentées ci-dessus montrent que la paroi est un assemblage complexe de polysaccharides, de protéines et de phlorotannins. Elle peut être considérée comme un véritable compartiment dynamique, dont la synthèse est finement contrôlée. Au cours de cette thèse, je me suis particulièrement intéressé au composé majoritaire de la paroi des algues brunes : l'alginate.

## II-4 L'alginate : un composé majoritaire

### II-4-a Structure chimique de l'alginate

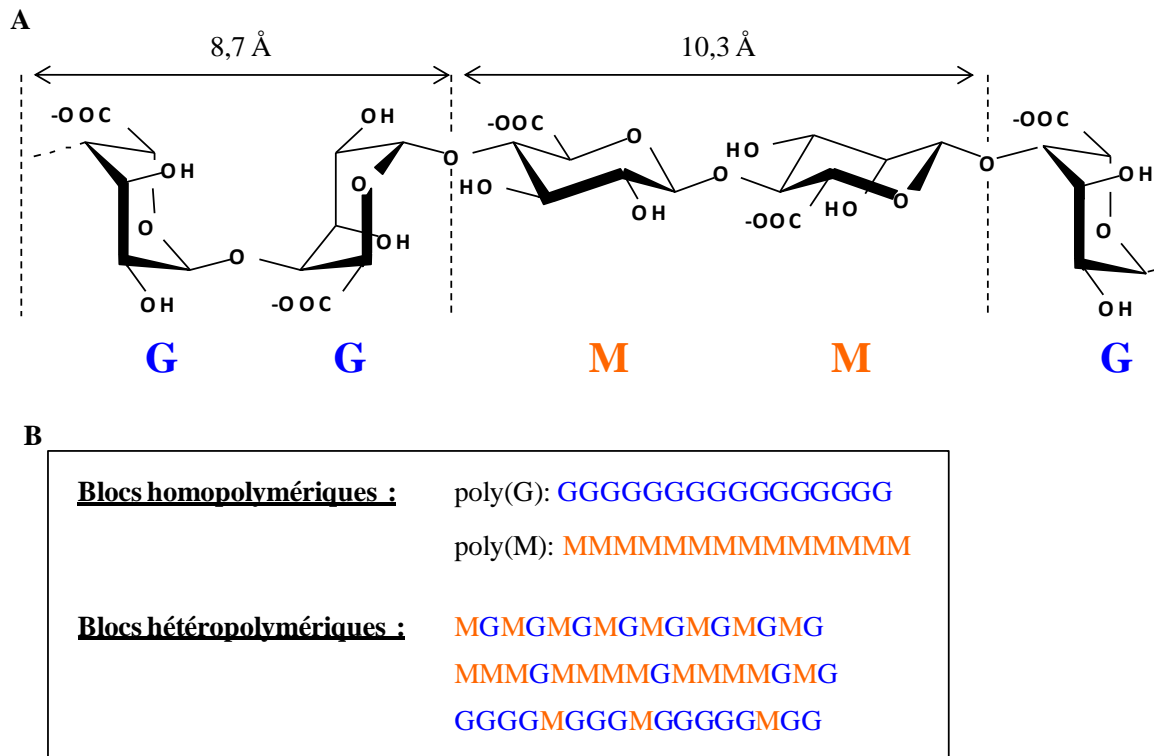
L'alginate est un polysaccharide constitué de l'enchaînement de deux types d'acides uroniques : l'acide  $\beta$ -1,4-D-mannuronique (M) et son C-5 épimère l'acide  $\alpha$ -1,4-L-guluronique (G). La seule différence entre les deux types de motifs est l'orientation du groupement carboxylique substitué en position 5. Il est au-dessus du plan de l'anneau saccharidique dans les motifs M et en-dessous dans les motifs G (Figure 1-10). La masse molaire d'un motif M ou G est de  $194 \text{ g.mol}^{-1}$ . L'alginate produit par certaines bactéries des familles *Pseudomonadaceae* et *Azotobacteriaceae* est par ailleurs acétylé sur les positions 2 et/ou 3 des motifs M.



**Figure 1-10 : Structure et conformation des constituants monosaccharidiques de l'alginate**

Dans le polymère, deux unités successives sont liées entre la position 1 de la première et la position 4 de la seconde (Figure 1-11). L'arrangement des deux types de motifs M et G permet de distinguer trois types d'organisation de l'alginate : des blocs homopolymériques de

motifs G [poly(G)], des blocs homopolymériques de motifs M ([poly(M)] et des blocs hétéropolymériques de motifs M et G. Ces derniers peuvent être constitués par l'alternance répétée de M et de G, ou par de courts blocs d'un type de motif séparés par un résidu de l'autre motif (Haug *et al.*, 1967).



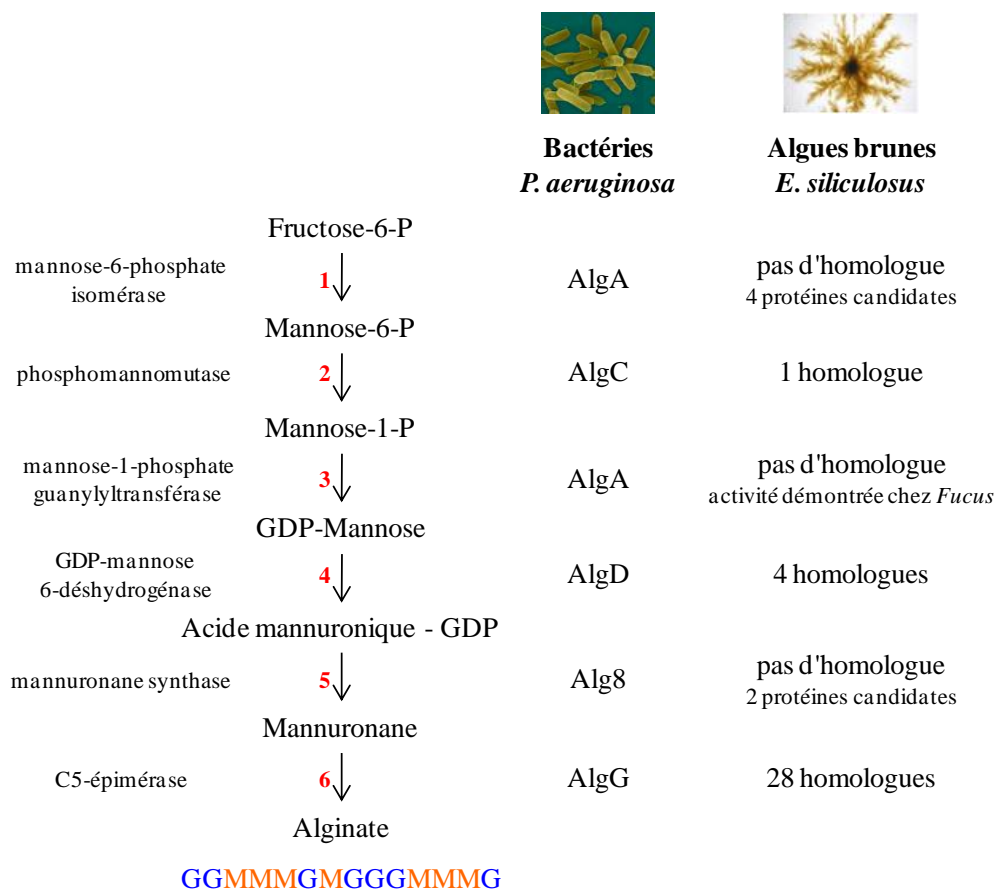
**Figure 1-11 : Organisation d'un polysaccharide d'alginate**

**A.** Liaisons possibles entre les deux types de motifs mannuronate (M) et guluronate (G). **B.** Les différents types de blocs composant une molécule d'alginate.

Les trois types de blocs peuvent être présents dans une même molécule d'alginate, chaque bloc comportant 20 à 30 résidus. La masse molaire d'une chaîne entière peut atteindre 800 kDa, soit un degré de polymérisation de plus de 4000 résidus (Kloareg and Quatrano 1988). Les blocs poly(G) forment des structures rigides et compactes, avec une distance entre deux résidus de 8,7 Å (Figure 1-11 A). Du fait de leur conformation différente, les blocs poly(M) forment des régions en "rubans allongés" plus souples avec une distance inter-résidus de 10,3 Å. Les régions hétéropolymériques constituées de petits blocs d'un motif forment des chaînes brisées, alors que celles où M et G alternent régulièrement adoptent une conformation en sinusoïde (Wong *et al.*, 2000). La variabilité structurale d'une chaîne d'alginate se joue donc à différents niveaux : le degré de polymérisation de la chaîne, l'abondance relative des deux types de motifs (ratio M/G) et leur enchaînement.

## II-4-b La biosynthèse des alginates

Les étapes de la biosynthèse de l'alginate et les enzymes impliquées sont bien connues chez deux espèces de Gammaprotéobactéries qui en produisent, *Pseudomonas aeruginosa* et *Azotobacter vinelandii* (Rehm and Valla 1997; Ramsey and Wozniak 2005). Le fructose-6-phosphate est à la base de la voie de synthèse (Figure 1-12). Il est converti en acide mannuronique-guanoside diphosphate (GDP-ManA) en quatre étapes catalysées par les enzymes AlgA (bifonctionnelle, mannose-6-phosphate isomérase et mannose-1-phosphate guanylyltransférase), AlgC (phosphomannomutase) et AlgD (GDP-mannose 6-déshydrogénase). Le nucléotide-ose GDP-ManA est la forme activée de l'acide mannuronique et le précurseur de l'alginate. L'enzyme Alg8 (mannuronane synthase) catalyse la polymérisation du GDP-ManA et synthétise donc des chaînes homopolymériques poly(M). Certains résidus M peuvent alors être convertis en motifs G par l'épimérisation du groupement carboxyl en position C5. Cette réaction qui a lieu au niveau du polymère est catalysée par la C5-épimérase AlgG. Suite à ce processus, on obtient donc une molécule d'alginate complexe dont le ratio M/G et l'enchaînement des deux motifs dépendent du mode d'action et de l'activité de la C5-épimérase.



**Figure 1-12 : La voie de biosynthèse de l'alginate chez les bactéries (*Pseudomonas aeruginosa*) et les algues brunes (*Ectocarpus siliculosus*)**  
d'après Michel *et al.* (2010)

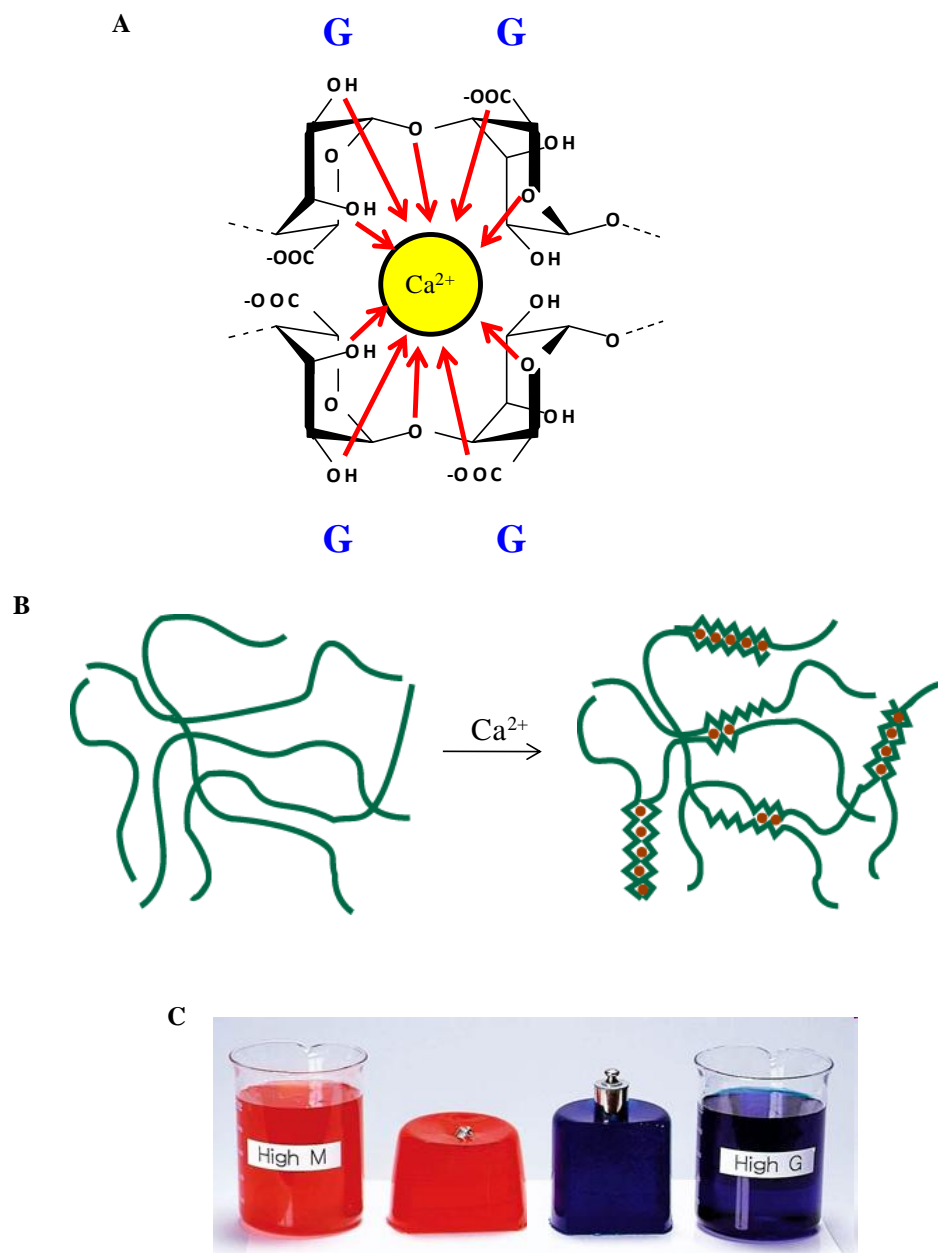
Il a été montré que la synthèse de l'alginate implique des réactions biochimiques équivalentes chez les algues brunes (Lin and Hassid 1966; Haug and Larsen 1969). Le récent séquençage du premier génome d'algue brune, celui d'*Ectocarpus siliculosus*, a permis d'apporter des informations sur les gènes potentiellement impliqués dans cette voie de biosynthèse (Cock *et al.*, 2010; Michel *et al.*, 2010). Des homologues des protéines bactériennes AlgC, AlgD et AlgG ont été trouvés. Notamment, la protéine Esi 016-0053 a été prédite comme une GDP-mannose déshydrogénase (Michel *et al.*, 2010). Son activité a récemment été confirmée grâce à l'expression hétérologue de la protéine chez *E. coli* (Tenhaken *et al.*, 2011). Aucun homologue de la mannuronane synthase Alg8 n'a pu être détecté dans le génome. Cependant, *E. siliculosus* possède deux protéines appartenant à la même famille que Alg8 (glycosyltransférase 2, GT2) et qui pourraient catalyser la polymérisation du GDP-ManA. De même, la protéine bifonctionnelle AlgA n'a pas d'homologue chez *E. siliculosus*. Les réactions 1 et 3 de la voie de biosynthèse (Figure 1-12) pourraient être catalysées par des enzymes différentes, monofonctionnelles. Ainsi, *E. siliculosus* possède 4 homologues de mannose-6-phosphate isomérase animales, qui convertissent le fructose-6-phosphate en mannose-6-phosphate et pourraient jouer le même rôle chez les algues. Par contre, aucune protéine candidate n'a pu être proposée pour catalyser l'activation du mannose-1-phosphate en GDP-mannose. Cette activité a cependant été détectée chez le genre *Fucus*, et doit donc être codée par un gène appartenant à une nouvelle famille (Michel *et al.*, 2010).

#### II-4-c Propriétés des alginates

Sous forme soluble, les alginates ont la propriété d'augmenter fortement la viscosité de solutions aqueuses. Ceci est dû à des contraintes empêchant la libre rotation autour des liaisons glycosidiques, qui augmentent selon le schéma suivant : MG < MM < GG. Les régions riches en blocs poly(M) et poly(G) sont donc plus rigides que celles où les deux types de motifs sont alternés. Une solution d'alginate riche en blocs homopolymériques poly(G) sera donc la plus visqueuse. D'autre part, l'alginate en solution peut être modélisé comme une succession de tiges raides (blocs homopolymériques) assemblées par des zones flexibles (blocs hétéropolymériques alternés). La solubilité de l'alginate dépend du pH de la solution car il détermine la présence de charges électrostatiques sur les résidus uroniques. Les constantes de dissociation ( $pK_a$ ) pour les motifs M et G sont de 3,38 et 3,65, respectivement.

Le  $pK_a$  du polymère est similaire. Une baisse brutale du pH en-dessous du  $pK_a$  provoque la précipitation de l'alginate (Haug and Larsen 1963).

Les sels d'acide alginique avec des ions monovalents (ex.  $Na^+$  et  $NH_4^+$ ) sont solubles. En revanche, l'alginate gélifie en présence de cations divalents ou multivalents, et en particulier de  $Ca^{2+}$ . Haug (1961) a démontré que l'affinité de l'alginate pour différents cations divalents (tels que  $Ca^{2+}$ ,  $Pb^{2+}$ ,  $Cd^{2+}$ ,  $Zn^{2+}$ ,  $Cu^{2+}$ , etc.) augmente avec la proportion de motifs G dans l'alginate. La plus grande affinité des motifs G pour les cations divalents s'explique par la conformation en "zigzag" adoptée par les blocs poly(G).



**Figure 1-13 : Les propriétés gélifiantes des alginates**

**A.** Liaison d'un ion  $Ca^{2+}$  dans la poche formée de motifs guluronates (G) selon le modèle des boîtes à œufs. D'après Donati et Paoletti (2009). **B.** Représentation schématique de la gélification de l'alginate en présence de calcium. D'après Christensen *et al.* (1990). **C.** Gels obtenus à partir d'alginate riche en motifs M (en rouge) ou G (en bleu) Photo : Kimica Corporation.

En présence de calcium (ou d'autres cations divalents de taille similaire), des connections inter-chaînes s'établissent entre des blocs poly(G) de différentes molécules d'alginate. L'alignement de deux blocs poly(G) crée des cavités où un ion  $\text{Ca}^{2+}$  peut venir se loger. Cette description du processus de gélification est connue sous le nom de modèle "en boîte à œufs" (*egg-box model*), dans lequel les œufs sont les cations divalents emprisonnés dans les boîtes constituées des motifs G (Figure 1-13A).

Le ratio M/G et l'enchaînement des différents types de motifs conditionnent donc les propriétés mécaniques des gels d'alginate. Les alginates riches en guluronate forment des gels durs, cassants et résistants à chaleur. Les alginates riches en mannuronate forment quant à eux des gels plus souples et élastiques (Figure 1-13 C).

#### II-4-d Variations biologiques de la composition en alginate

Comme décrit précédemment, la synthèse d'alginate dans les parois des algues brunes est un processus finement contrôlé. Le ratio M/G et la position des différents motifs sont contrôlés par l'action des C5-épimérasés. Ces paramètres conditionnent les performances mécaniques du polymère, et leurs variations reflètent des propriétés de paroi différentes (élasticité, résistance, etc.). Il n'est donc pas surprenant de trouver des variations biologiques de la composition des alginates.

En effet, la proportion d'alginate dans la paroi et le ratio M/G varient selon l'espèce d'algue brune considérée (Tableau 1-2). Cependant, ces variations sont difficiles à relier à la taxonomie (Kloareg and Quatrano 1988).

<i>Source</i>	$F_G$	$F_M$
<i>Laminaria japonica</i>	0.35	0.65
<i>Laminaria digitata</i>	0.41	0.59
<i>Laminaria hyperborea</i> , blade	0.55	0.45
<i>Laminaria hyperborea</i> , stipe	0.68	0.32
<i>Laminaria hyperborea</i> , outer cortex	0.75	0.25
<i>Lessonia nigrescens</i> <sup>a</sup>	0.38	0.62
<i>Ecklonia maxima</i>	0.45	0.55
<i>Macrocystis pyrifera</i>	0.39	0.61
<i>Durvillea antarctica</i>	0.29	0.71
<i>Ascophyllum nodosum</i> , fruiting body	0.10	0.90
<i>Ascophyllum nodosum</i> , old tissue	0.36	0.64

**Tableau 1-2 : Proportion de motifs M et G dans des alginates de différentes sources**

Proportions de motifs guluronates ( $F_G$ ) et mannuronates ( $F_M$ ) déterminées par spectroscopie RMN. D'après Draget *et al.* (2005).



Au sein d'un même individu, le ratio M/G de l'alginate change en fonction de la partie de l'algue. Ainsi, les frondes possèdent un alginate plus riche en mannuronate, et donc plus souple. Les stipes et les crampons, qui servent de point d'ancrage au substrat rocheux, possèdent un alginate plus riche en guluronate ce qui leur confère une grande rigidité et une meilleure résistance aux stress mécaniques (Haug *et al.*, 1974a). Les tissus âgés sont en général plus riches en motifs G, ce qui traduit l'action de C5-épimérase au cours du temps. Ceci est à relier avec les variations saisonnières qui peuvent être observées. Par exemple, chez *Undaria pinnatifida*, des suivis de la production d'alginate au cours du temps ont montré que celle-ci augmente de mars à juillet. L'alginate extrait des jeunes plantules collectées au mois de mars contient principalement des blocs poly(M), et la proportion de motifs G augmente au cours du printemps et de l'été au fur et à mesure que les algues grandissent (Skriptsova *et al.*, 2004). La composition de l'alginate répond également aux conditions environnementales. Ainsi, les algues qui poussent dans un milieu très battu ont généralement un contenu en alginate plus élevé que celles vivant dans des eaux plus calmes (Craigie *et al.*, 1984). D'autre part, il a été montré que la culture en présence de fortes concentrations en phosphate ou en métaux lourds influence la production d'alginate et augmente le taux de guluronate (Indergaard and Skjakbraek 1987; Andrade *et al.*, 2010).

#### II-4-e Les applications des alginates

Les propriétés physico-chimiques intéressantes des alginates ont trouvé de multiples applications dans des domaines comme l'industrie textile, agro-alimentaire ou pharmaceutique. La demande en alginate est importante et la production est de l'ordre de 30000 tonnes par an. Tous les alginates commerciaux sont extraits d'algues brunes, principalement des genres *Laminaria*, *Macrocystis*, *Lessonia* et *Ascophyllum* (Tableau 1-3).

Genre	Récolte (tonnes humides / an)	%
<i>Laminaria</i>	43500	34%
<i>Macrocystis</i>	35000	28%
<i>Lessonia</i>	20500	16%
<i>Ascophyllum</i>	20000	16%
<i>Durvillaea</i>	4500	3,5%
<i>Ecklonia</i>	3000	2,5%
TOTAL	126500	

**Tableau 1-3 : Masses d'algues récoltées dans le monde en 2001**  
D'après McHugh (2003).

La grande majorité de la biomasse algale utilisée est obtenue à partir de populations sauvages. La récolte peut encore se faire à la main pour certaines espèces (notamment *Ascophyllum*), même si la majeure partie est aujourd'hui réalisée mécaniquement sur des bateaux spécialement équipés. Ce prélèvement important dans les champs d'algues nécessite une gestion raisonnée des stocks afin de ne pas épuiser la ressource. Il n'existe que très peu de fermes de cultures d'algues brunes. On note cependant la culture de *Saccharina japonica* en Asie pour des utilisations alimentaires.

L'alginate est extrait de la paroi des algues brunes par un processus chimique, le but étant d'obtenir de l'alginate de sodium en poudre sèche. Le principe est de convertir tous les sels d'alginate présents dans la paroi en alginate de sodium, qui est soluble dans l'eau. Ceci est possible en ajoutant du carbonate de sodium à un broyat d'algues. Il existe ensuite deux méthodes pour extraire l'alginate de sodium de la solution aqueuse. Dans la première, on fait précipiter l'alginate par ajout d'acide et on élimine une partie de l'eau présente dans les "flocons" obtenus. Dans la deuxième méthode on utilise du chlorure de calcium pour obtenir des fibres d'alginate gélifiées. Un traitement à l'alcool et au carbonate de sodium permet de récupérer l'alginate de sodium sous forme de pâte qui peut alors être séchée (McHugh 2003). Le coût de production de l'alginate varie en fonction du degré de pureté souhaité. Des alginates peu raffinés, contenant encore des traces de débris d'algues, peuvent être obtenus pour 1 dollar US par kg. Ce coût peut monter jusqu'à 5 dollars par gramme pour des alginates ultra-purs utilisés dans des applications médicales (Draget *et al.*, 2005).

Les applications des alginates sont multiples. L'industrie agro-alimentaire consomme approximativement 50% de la production. L'alginate est utilisé comme additif alimentaire pour améliorer, modifier et stabiliser la texture des préparations (numéros E400 à E405). Pour éviter la précipitation de l'alginate de sodium en milieu acide, on utilise un de ces dérivés, le propylène glycol alginate (PGA). Les propriétés épaississantes de l'alginate sont utiles pour obtenir des sauces et sirops plus consistants. Les émulsions eau-huile, telles que la mayonnaise ou la vinaigrette industrielles, sont moins susceptibles de se séparer en deux phases si on y ajoute de l'alginate. L'alginate permet également de garder en suspension la pulpe dans les jus de fruits, et le chocolat dans les boissons au cacao. Il évite la formation de cristaux désagréables en bouche dans les sorbets et crèmes glacées. De petites quantités de PGA ajoutées à la bière donnent une belle mousse stable durant plus longtemps. Les propriétés gélifiantes de l'alginate sont aussi largement exploitées par l'industrie alimentaire.

Les premières cerises artificielles (des billes d'alginate de calcium parfumées et colorées) datent de 1946. Ce principe a depuis été repris et est à la base, par exemple, de la garniture des olives fourrées au piment, de pièces de viandes reconstituées, de desserts gélifiés, etc. Il inspire également les plus grands chefs de la "nouvelle cuisine" (Figure 1-14).



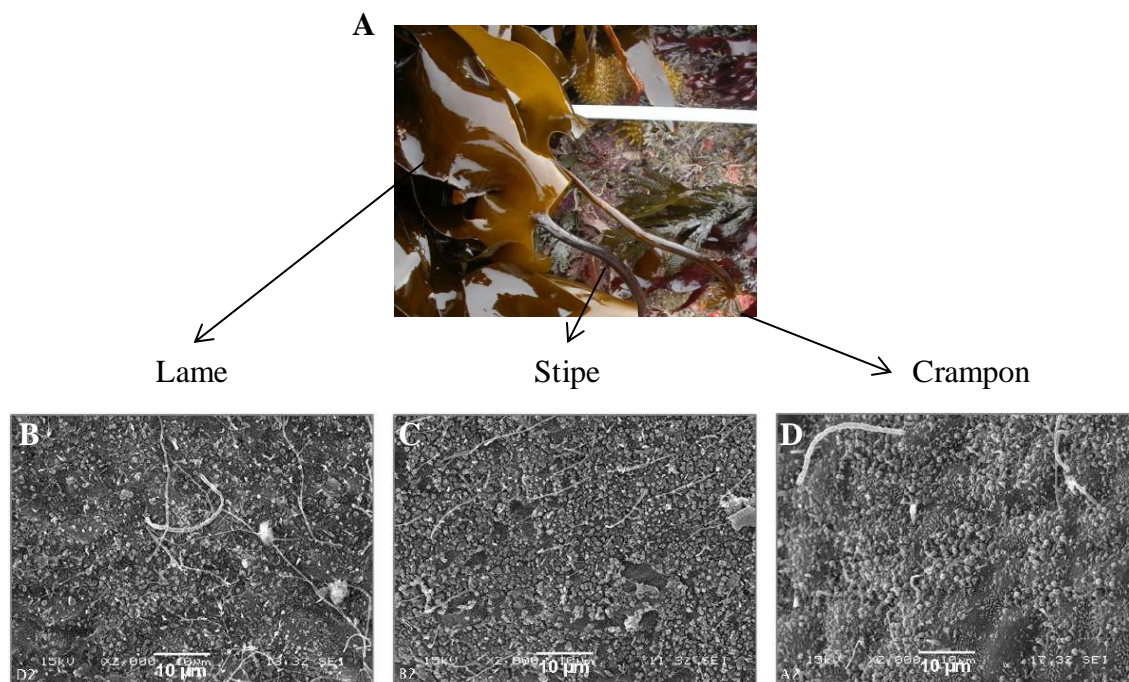
**Figure 1-14 : Exemples d'utilisations de l'alginate dans l'industrie**

L'alginate est d'autre part utilisé dans l'industrie textile et dans les imprimeries comme épaississant pour les encres. Sa forte affinité pour les cations divalents en fait un bon agent pour absorber les pollutions par des métaux lourds (Davis *et al.*, 2003). Son usage est répandu dans les cosmétiques (dentifrices, crèmes de beauté), toujours comme texturant. Dans le domaine médical, on exploite les fibres d'alginate de calcium en tant que pansement pour des blessures. Ces fibres peuvent être facilement absorbées par l'organisme, puisque le calcium est lentement échangé contre le sodium des fluides corporels, donnant un alginate soluble. Des substances actives comme des médicaments peuvent être encapsulées dans des billes d'alginate qui, une fois avalées, vont les libérer (Matricardi *et al.*, 2008). Des mélanges alginate / chitosane peuvent être utilisés pour améliorer les propriétés mécaniques des microsphères. De telles billes d'alginate peuvent aussi servir à délivrer de l'ADN lors de thérapies géniques (Pannier and Shea 2004) ou même des cellules vivantes, dans le cadre de

traitements de cancers (Hao *et al.*, 2005), de diabète (Basta *et al.*, 1995), et d'essais de transplantations neuronales pour soigner la maladie de Parkinson (Yasuhara and Date 2007).

## II-5 Rôle de la paroi des algues brunes dans l'interaction avec les bactéries

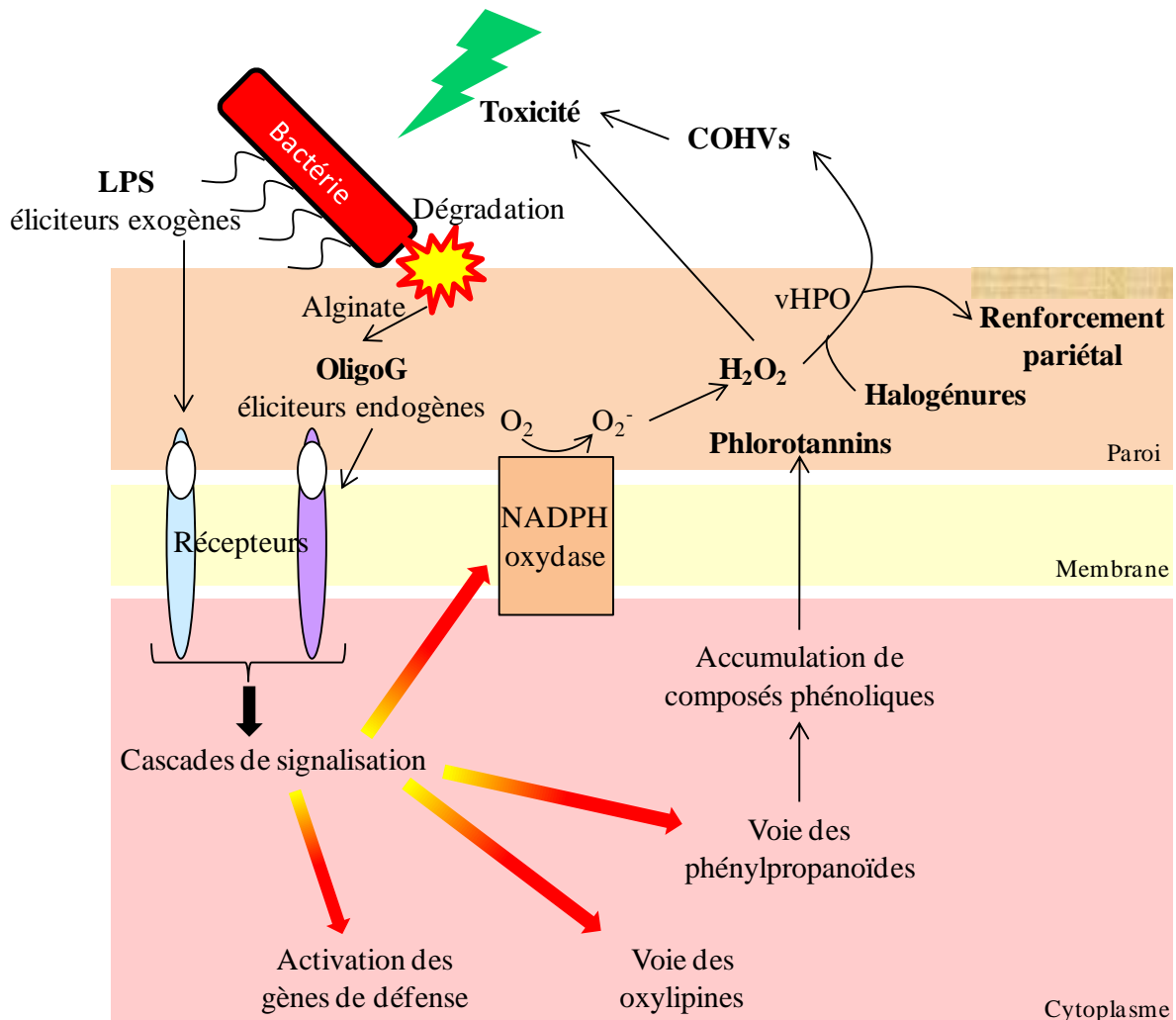
Comme il a déjà été mentionné précédemment, la paroi des algues est une zone d'interface avec l'environnement. Elle est colonisée par une communauté microbienne épiphyte abondante qui peut former des biofilms (Figure 1-15). Chez les algues brunes, il a été montré que la composition des populations microbiennes varie en fonction de la partie de l'algue (lame, stipe ou crampon) (Corre and Prieur 1990; Staufenberger *et al.*, 2008). Les bactéries présentes à la surface des algues brunes sont principalement des Gram négatifs, appartenant aux phyla des *Alphaproteobacteria*, *Gammaproteobacteria* et *Bacteroidetes*. Ces communautés bactériennes varient avec la saison, et il a été récemment montré que les *Planctomycetes* pouvaient être le groupe dominant sur les laminaires en été (Bengtsson and Ovreas 2010).



**Figure 1-15 : Communautés microbiennes associées à l'algue brune *Laminaria digitata***

A. Photo de *L. digitata* présentant les différentes parties de l'algue. B, C, D. Images en microscopie électronique à balayage de communautés microbiennes associées à la lame (B), au stipe (C) et au crampon (D). D'après Salaün (2009).

Les algues brunes sont capables de réguler le développement des bactéries épiphytes et de se défendre face aux agressions de bactéries pathogènes. La paroi joue un rôle majeur dans les phénomènes de défense impliqués, étant la première barrière physique entre les tissus de l'algue et les bactéries (Figure 1-16).



**Figure 1-16 : Schéma des réponses de défense des algues brunes face à une attaque bactérienne**  
Schéma simplifié inspiré de Cosse (2007). vHPO : Haloperoxydase à vanadium. COHV : Composés organiques halogénés volatils.

La première étape d'une réponse de défense est la reconnaissance de l'agresseur. Chez les algues du genre *Laminaria*, il a été montré que les lipopolysaccharides (LPS) bactériens peuvent être reconnus comme signaux d'une attaque (Küpper *et al.*, 2006). On les appelle des éliciteurs exogènes car ils ne proviennent pas de l'organisme attaqué. De manière particulièrement intéressante, il a été démontré que des fragments d'alginate provenant de la

digestion de la paroi agissent eux aussi comme des signaux déclenchant les réponses de défense chez les algues brunes (Küpper *et al.*, 2001; Küpper *et al.*, 2002). On parle alors d'éliciteurs endogènes. Cette reconnaissance est particulièrement spécifique, puisque les oligo-gulonates (blocs G) induisent une réponse beaucoup plus intense que les blocs M ou les blocs mixtes MG (Küpper *et al.*, 2001). La structure en boîte à œufs précédemment présentée pour les blocs G est similaire à celle adoptée par les oligosaccharides résultant de la dégradation de la pectine. Chez le tabac, il semble que cette structure soit responsable de l'activité des produits de dégradation de la pectine dans la réponse de défense aux bactéries phytopathogènes (Spiro *et al.*, 1998).

La perception de l'attaque déclenche des cascades de signalisation aboutissant à l'expression des réponses de défense (Cosse *et al.*, 2007). Chez les algues brunes, l'une des premières réactions est l'émission d'espèces actives de l'oxygène (EAO) comme le peroxyde d'hydrogène  $H_2O_2$ , ou l'anion superoxyde  $O_2^-$ . Cette émission très forte intervient dès les premières minutes après l'élicitation. Elle est transitoire, c'est pourquoi on parle de bouffée (ou *burst*) oxydative. Ces EAO ont une activité anti-microbienne et participent donc directement à limiter la progression des pathogènes. Par ailleurs,  $H_2O_2$  est le substrat d'enzymes nommées haloperoxydases dépendantes du vanadium (vHPO). Les vHPO catalysent l'oxydation d'halogénures (iodure, bromure ou chlorure) en un intermédiaire halogéné qui peut à son tour réagir avec des molécules organiques. L'action des vHPO conduit donc à la production d'iode moléculaire  $I_2$  et de composés organiques halogénés volatils (COHV), qui sont impliqués dans la défense chimique contre les agresseurs (Cosse *et al.*, 2007). D'autres voies de signalisation sont induites lors des réactions de défense. Après élicitation, la voie des phénylpropanoïdes conduit à l'accumulation de composés phénoliques (phlorotannins) dans les algues (Küpper *et al.*, 2002). Ces composés ont eux aussi une activité anti-microbienne. D'autre part, il a été montré que des vHPO peuvent catalyser le pontage oxydatif de molécules d'alginate avec les phlorotannins, ce qui conduit à un renforcement de la paroi et donc freine la dégradation (Bitton *et al.*, 2006; Salgado *et al.*, 2009). D'autres métabolites de défense sont également produits, tels que des oxylipines.

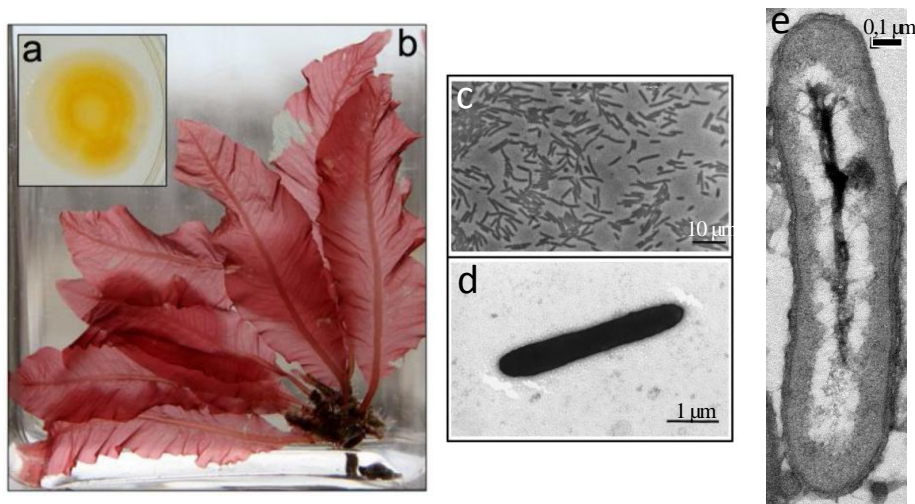
La description de la paroi des algues brunes a permis de mettre en évidence l'importance de l'alginate, pour son abondance, ses propriétés physico-chimiques, ses applications multiples et le rôle central qu'il semble jouer dans la reconnaissance des agresseurs. L'étude

de la dégradation de l'alginate par des bactéries devrait donc fournir des informations précieuses sur le recyclage de cette source de matière organique dans l'écosystème et les interactions algues / micro-organismes. Au cours de ma thèse, j'ai étudié le métabolisme de l'alginate chez *Zobellia galactanivorans*.

### III- *Zobellia galactanivorans* : un modèle d'interaction algues/bactéries

#### III-1 *Z. galactanivorans* et le genre *Zobellia*

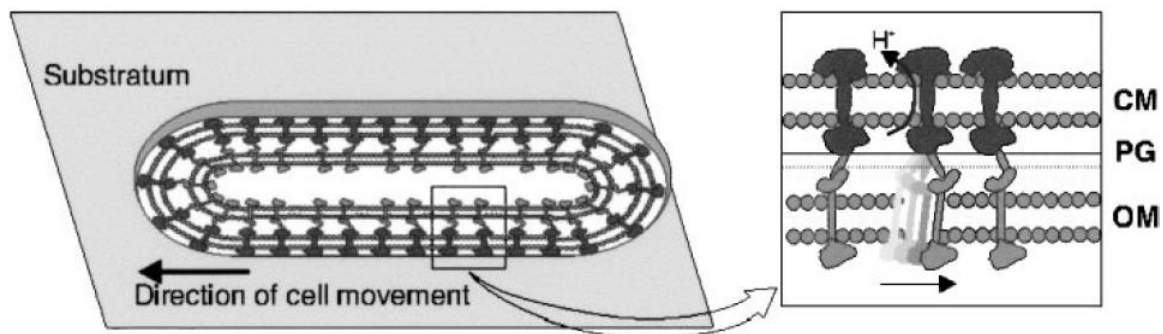
*Zobellia galactanivorans* (souche Dsij<sup>T</sup>) est une bactérie marine isolée en novembre 1988 dans la baie de Roscoff à partir de l'algue rouge *Delesseria sanguinea* (Potin *et al.*, 1991). Elle a tout d'abord été identifiée sous le nom de "*Cytophaga drobachiensis*" puis assignée au genre *Zobellia*, dont elle est l'espèce type (Barbeyron *et al.*, 2001). Ce genre appartient au phylum des *Bacteroidetes*, classe des *Flavobacteriia*, ordre des *Flavobacteriales*, famille des *Flavobacteriaceae*. Les membres de ce phylum ont colonisé de nombreuses niches écologiques et sont spécialisés dans la dégradation de la matière organique de haut poids moléculaire (cf. Chapitre 2).



**Figure 1-17 : La bactérie marine *Zobellia galactanivorans*, isolée de l'algue rouge *Delesseria sanguinea***  
**a.** Une colonie de *Z. galactanivorans* après 5 jours sur gel d'agar. **b.** L'algue rouge *D. sanguinea*. **c.** Photographie en microscopie par contraste de phase d'une culture liquide de *Z. galactanivorans*. **d, e.** Photographies en microscopie électronique à transmission. a, c, d : copiées de Barbeyron *et al.* (2001). b : ©Hehemann. e : ©Thomas.



*Z. galactanivorans* est une bactérie Gram-négative, en forme de longs bâtonnets (0,3-0,4  $\mu\text{m}$  de large sur 3,0-8,0  $\mu\text{m}$  de long) (Figure 1-17). Elle est hétérotrophe chimio-organotrophe, ce qui signifie qu'elle utilise des composés organiques (ex. glucose) comme source d'énergie. Son métabolisme est purement respiratoire, aérobie, avec l'oxygène comme accepteur final d'électrons. La couleur jaune-orange des colonies est due à la présence d'un pigment de type flexirubine. *Z. galactanivorans* ne possède pas de flagelle et donc n'est pas mobile en phase liquide. Par contre, dès qu'elle entre en contact avec une surface solide (ex. une lame de microscope), on observe facilement une mobilité par glissement de l'ordre de 1 à 4  $\mu\text{m}\cdot\text{s}^{-1}$ . Ceci explique la morphologie des colonies en cercles concentriques (Figure 1-17 a). Le mécanisme permettant cette mobilité chez les flavobactéries n'a pas encore été totalement éclairci. Il semble qu'il n'implique ni l'élongation/rétractation de pili de type IV (connus pour participer au glissement des myxobactéries) ni l'extrusion de polysaccharides qui propulserait la cellule (McBride 2001). En revanche, l'utilisation de découplants et de poisons respiratoires a montré qu'il est probable que la mobilité par glissement utilise un gradient de protons comme force motrice. Selon un des modèles proposés, des protéines de la membrane cytoplasmique utiliseraient cette force motrice pour propulser des protéines de la membrane externe le long du peptidoglycane. Ceci résulterait en un mouvement global de la membrane externe autour de la cellule, d'une manière similaire aux chenilles des roues d'un char (Figure 1-18).



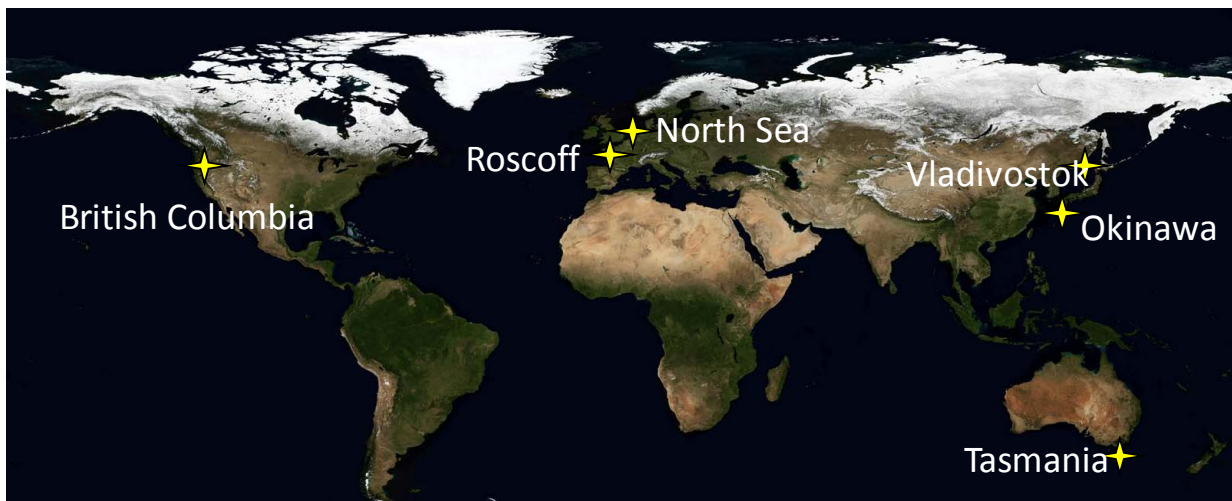
**Figure 1-18 : Modèle spéculatif de la mobilité par glissement chez les flavobactéries.**

Des protéines de la membrane cytoplasmique (CM) utiliseraient la force motrice d'un gradient de proton pour propulser des protéines de la membrane externe (OM) le long du peptidoglycane (PG). D'après McBride *et al.* (2001).



### III-2 *Zobellia* et les algues: "Je t'aime...moi non plus !"

Outre *Z. galactanivorans*, le genre *Zobellia* compte à ce jour 4 autres espèces décrites. Il s'agit de *Z. amurskyensis*, *Z. laminariae*, *Z. russellii* et *Z. uliginosa*. Dix souches non-décrites ont été assignées au genre *Zobellia* par similitude de séquences de l'ARNr 16S (NCBI, 10/04/11). Comme on peut le remarquer sur la Figure 1-19, le genre *Zobellia* semble inféodé aux zones littorales, dans différentes régions du globe (Europe, Canada, Asie, Australie). Ce genre apparaît donc comme un bon modèle de bactéries marines hétérotrophes en milieu côtier. Par ailleurs, les membres de ce genre ont été principalement retrouvés en étroite association avec des algues marines, qu'il s'agisse d'algues vertes (Matsuo *et al.*, 2003; Nedashkovskaya *et al.*, 2004), rouges (Potin *et al.*, 1991), brunes (Nedashkovskaya *et al.*, 2004) ou de dinoflagellés (Skerratt *et al.*, 2002).



**Figure 1-19 : Les sites d'isolement des membres du genre *Zobellia***

Des membres du genre *Zobellia* ont été détectés à Roscoff (Potin *et al.*, 1991), en Mer du Nord (Alonso *et al.*, 2007), au Canada (Schulze *et al.*, 2006), à Vladivostok (Nedashkovskaya *et al.*, 2004), au Japon (Matsuo *et al.*, 2003), en Tasmanie (Skerratt *et al.*, 2002)

Les interactions que les membres du genre *Zobellia* établissent avec les algues sont multiples, et varient de la symbiose à un comportement algicide. Ainsi, on sait que beaucoup d'algues vertes ne peuvent développer leur morphologie normale (avec un thalle foliacé) si elles sont cultivées en conditions axéniques. Par exemple, l'algue *Ulva lactuca* perd sa forme typique en "feuille de laitue" lorsqu'elle croît en milieu stérile et adopte une forme en filaments unisériés (Provasoli 1958). Il a été montré que certaines souches proches de *Z. uliginosa* sont capables d'induire une morphogénèse normale chez les algues *Monostroma oxyspermum*, *Ulva pertusa*, *U. conglobata* et *Enteromorpha intestinalis* (Matsuo *et al.*, 2003).

Un composé sécrété par les bactéries, la thallusine, a été identifié comme un puissant morphogène. A de très faibles concentrations (de l'ordre de  $10^{-15}$  à  $10^{-18}$  g.ml<sup>-1</sup>), il permet de restaurer le développement normal des thalles d'algues vertes (Matsuo *et al.*, 2005). Il semble que la thallusine doive être apportée en permanence par les bactéries épiphytes pour maintenir la forme classique de l'algue. Cette étroite interaction entre les algues vertes et la flore bactérienne peut être considérée comme de la symbiose. Le bénéfice pour l'algue consiste en une augmentation de la surface disponible pour la photosynthèse et les échanges de nutriments. L'avantage de cette interaction pour les bactéries reste moins évident, même si on peut imaginer que l'accroissement de la biomasse algale leur fournisse un habitat et des sources de matière organique plus abondantes.

A la différence de ces bactéries symbiotiques, d'autres membres du genre *Zobellia* ont un comportement purement pathogène. Ainsi, une espèce isolée en Tasmanie produit un exsudat capable de tuer des espèces de microalgues toxiques Rhaphidophytes et Dinoflagellés (Skerratt *et al.*, 2002). Il semble donc que les membres du genre *Zobellia* aient, au cours de l'évolution, développé différents degrés d'interaction avec les algues. Qu'en est-il pour *Z. galactanivorans* ?

### III-3 *Z. galactanivorans* : un dégradeur d'algues marines

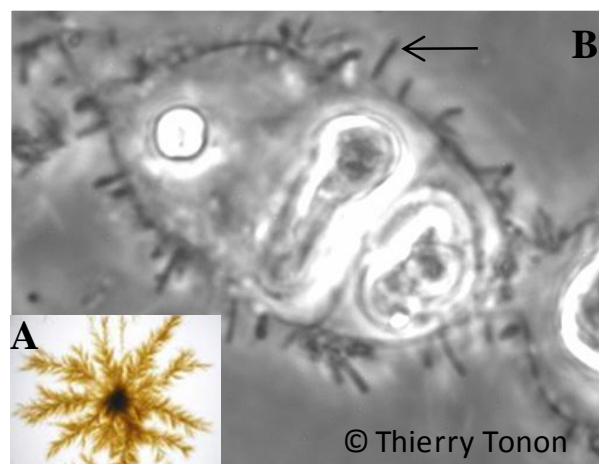
Comme il a déjà été mentionné, *Z. galactanivorans* a été isolée de l'algue rouge *D. sanguinea* pour sa capacité à dégrader les polysaccharides sulfatés pariétaux (agars et carraghénanes). Depuis lors, différentes études ont confirmé son grand potentiel de dégradation des algues marines et ont mis en évidence les liens étroits qu'elle peut établir avec ses hôtes végétaux.

*Z. galactanivorans* est capable d'utiliser de nombreux polysaccharides d'algues comme seule source de carbone (cf. Chapitre 3). Il s'agit aussi bien de polysaccharides d'algues rouges (agars, carraghénanes) que brunes (alginate, laminarine). Ceci signifie qu'elle possède toutes les enzymes nécessaires à la dégradation totale des différents substrats et à leur assimilation dans les voies métaboliques. Ce fort potentiel de dégradation a conduit à l'utilisation de *Z. galactanivorans* comme une source d'enzymes agissant sur ces polysaccharides. A ce jour, 11 polysaccharidases ont été caractérisées chez *Z. galactanivorans*, agissant toutes sur des composés d'algues rouges (Tableau 1-4).

Enzyme	Nom	Référence	Identifiant	PDB ID
iota-carraghénase	cgiA	(Barbeyron <i>et al.</i> , 2000)	Q9F284	-
iota-carraghénase	cgiA2	(Rebuffet <i>et al.</i> , 2010)	D9UAT1	-
iota-carraghénase	cgiA3	(Rebuffet <i>et al.</i> , 2010)	D9UAT2	-
kappa-carraghénase	cgkA	(Barbeyron <i>et al.</i> , 1998)	P43478	
$\beta$ -agarase	agaA	(Jam <i>et al.</i> , 2005)	Q9RGX9	1O4Y[A], 1URX[A]
$\beta$ -agarase	agaB	(Jam <i>et al.</i> , 2005)	Q9RGX8	1O4Z[A,B,C,D]
$\beta$ -agarase	agaC	(Jam <i>et al.</i> , 2005)	-	-
$\beta$ -agarase	agaD	(Hehemann <i>et al.</i> , 2010b)	D7GXX4	-
porphyranase	porA	(Hehemann <i>et al.</i> , 2010a)	D7GXX0	3ILF[A]
porphyranase	porB	(Hehemann <i>et al.</i> , 2010a)	D7GXF9	3JUJ[A,B]
anhydrogalactosidase	ahgA	(Rebuffet <i>et al.</i> , 2011)	F0V1E3	3P2N[A,B]

**Tableau 1-4 : Les polysaccharidases caractérisées de *Z. galactanivorans***

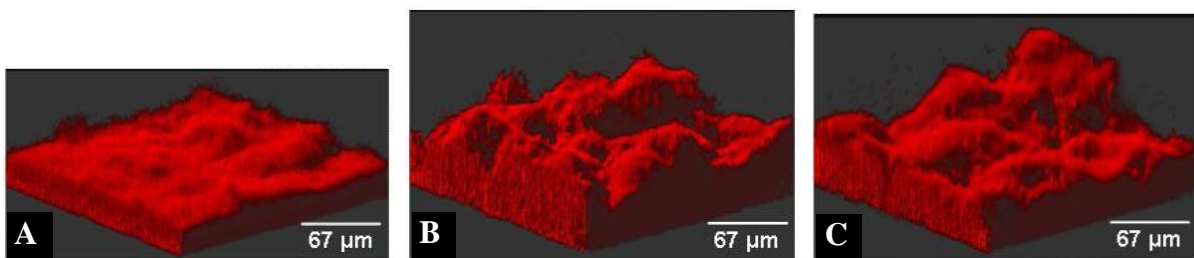
Ce travail de thèse est donc le premier à s'intéresser à la dégradation d'un polysaccharide d'algue brune, l'alginate, par *Z. galactanivorans*. Pourtant, des observations antérieures montrent que *Z. galactanivorans* est bien capable d'interagir étroitement avec ce type d'algues. Lors de la co-incubation de l'algue brune *Ectocarpus siliculosus* avec *Z. galactanivorans*, il a été constaté que les bactéries adhèrent fortement à la paroi (Figure 1-20) et provoquent l'apparition de protoplastes après quelques jours, indiquant une digestion de la paroi. Cette expérience a néanmoins été réalisée avec de fortes concentrations de *Z. galactanivorans* (Tonon et Michel, communication personnelle).



**Figure 1-20 : Interaction entre *Z. galactanivorans* et l'algue brune *Ectocarpus siliculosus***

A. Photographie de l'algue brune filamenteuse *E. siliculosus*. B. Photographie au microscope d'un filament d'*E. siliculosus* (grosse cellule) après incubation avec *Z. galactanivorans* (flèche noire). Tonon et Michel (communication personnelle).

Par ailleurs, on sait qu'en présence d'un support solide (par exemple la paroi d'une algue), *Z. galactanivorans* est capable d'adhérer et de former un biofilm dense et épais. Lors d'une précédente thèse dans l'Unité, Stéphanie Salaün a testé l'effet de composés iodés et d'exsudats de l'algue brune *Laminaria digitata* sur les capacités d'adhérence et de croissance en biofilm chez *Z. galactanivorans* (Salaün 2009). Comme il a été décrit dans la partie précédente, de tels composés sont émis par l'algue pour se défendre face à une agression, et on s'attendrait donc logiquement à ce qu'ils inhibent la croissance du biofilm. De manière très surprenante, les résultats ont montré que l'ajout de diiodométhane ( $\text{CH}_2\text{I}_2$ ) dans le milieu favorise l'adhérence de *Z. galactanivorans* sur un support solide inerte, et ne perturbe pas (voire augmente) la formation du biofilm. De même, l'ajout dans le milieu d'exsudats d'algues témoins ou préalablement traitées pour induire les réactions de défense augmente considérablement la formation du biofilm (Figure 1-21). Ces résultats suggèrent que *Z. galactanivorans* soit capable de résister aux réactions de défense des algues brunes et de coloniser leur paroi. Une fois fixées, les bactéries pourraient entamer des processus de dégradation des composés pariétaux tels que l'alginate. D'autre part, comme il a déjà été mentionné plus haut, les débris d'algues brunes constituent une part importante de la matière organique particulière dans les eaux côtières. Ces particules, chargées en polysaccharides, constituent également des supports pour l'implantation de bactéries telles que *Z. galactanivorans* sous forme de biofilms.



**Figure 1-21 : Biofilms de *Z. galactanivorans* en présence d'exsudats de l'algue brune *L. digitata***  
 Visualisation en 3D de biofilms formés en 24 h dans une chambre à flux contenant un milieu contrôle (A), un milieu enrichi en exsudats de *L. digitata* préalablement traitée (C) ou non (B) pour induire les réactions de défense. L'ensemble des bactéries a été coloré au Syto 61 Red (marqueur des acides nucléiques, fluorescent) avant observation au microscope confocal à balayage laser. Images copiées de Salaün (2009).

L'intérêt pour le potentiel d'enzymes de dégradation que représente *Z. galactanivorans* a permis d'obtenir le financement du séquençage de son génome par le Génoscope. Il comporte 5.5 Mb et 4738 gènes ont été prédits (Barbeyron, Thomas *et al.*, in prep). L'annotation experte a confirmé le potentiel de cette bactérie en tant que dégradeur de polysaccharides. En effet, *Z. galactanivorans* possède 123 glycoside hydrolases, 12

polysaccharide lyases et 18 carbohydate estérases. Par ailleurs, 72 gènes codant des sulfatases ont été annotés, ce qui pourrait refléter l'importance de la dégradation de polysaccharides sulfatés. Collectivement, environ 4,5% du génome seraient consacrés à la dégradation de polysaccharides. Le nombre important de CAZymes (*Carbohydrate Active enZymes*) prédites place *Z. galactanivorans* parmi les espèces les mieux équipées au sein du phylum des *Bacteroidetes* (cf. Chapitre 2). De nombreux systèmes impliqués dans la perception de l'environnement ont été trouvés. Il s'agit notamment de 118 récepteurs TonB dépendants dont l'implication a été montrée dans des locus d'utilisation de polysaccharides (cf. Chapitre 2) et de 68 systèmes de transduction du signal à 1 ou 2 composants. De manière intéressante, de nombreux clusters de gènes dont la fonction peut être reliée au catabolisme des sucres ont été trouvés dans le génome. Ceci suggère la présence d'opérons de dégradation, permettant une co-régulation de tous les gènes impliqués dans une voie.

Par ailleurs, l'hypothèse selon laquelle *Z. galactanivorans* serait capable de résister aux défenses des algues est corroborée par la présence de nombreux systèmes de détoxification et de résistance au stress oxydant. Pour détoxifier les espèces activées de l'oxygène, elle possède des enzymes cytoplasmiques (2 catalases, 1 peroxydase, une superoxyde dismutase à manganèse, 9 peroxyrédoxines) et périplasmiques ou fixées à la membrane externe (une azurine, 3 peroxydases à cytochrome c, une superoxyde dismutase [Cu-Zn] dépendante et une peroxyrédoxine). Elle pourrait également résister aux espèces activées de l'azote telles que l'oxyde nitrique NO grâce à plusieurs systèmes de réduction de ce composé en N<sub>2</sub>. Enfin, la présence de sulfoxyde réductases pourrait favoriser la réparation des dommages oxydatifs causés sur les protéines.

Ainsi, même si *Z. galactanivorans* a été isolée d'une algue rouge, il semble qu'elle soit également bien armée pour utiliser la matière organique provenant des algues brunes. Cet aspect n'a pourtant encore jamais été abordé auparavant. Ces potentielles capacités de résistance aux défenses des algues et de dégradation des composés pariétaux en font un bon modèle pour l'étude des interactions entre algues et bactéries.

## IV- La dégradation de l'alginate

La dégradation de l'alginate est effectuée par des enzymes spécifiques, les alginate lyases. Avant de présenter dans le détail ces polysaccharidases, j'introduirai brièvement les organismes connus pour dégrader l'alginate et je discuterai de la fonction biologique de cette dégradation.

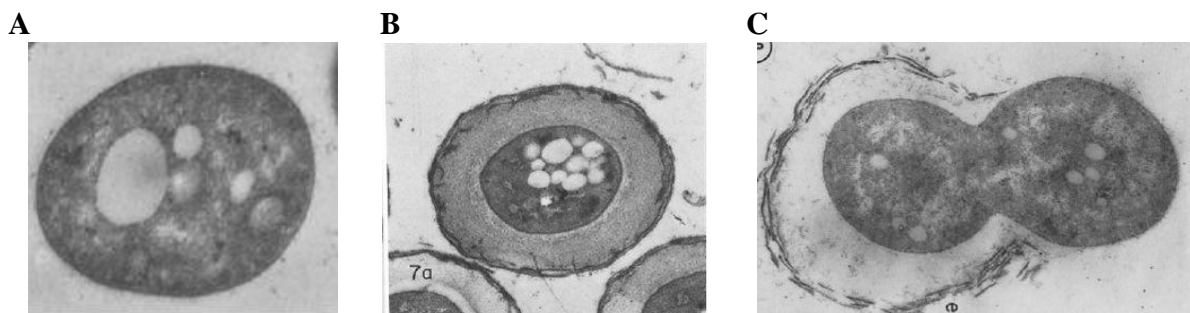
### IV-1 Les organismes dégradeurs

Deux catégories d'organismes peuvent dégrader l'alginate : ceux qui en produisent et ceux qui le consomment.

#### IV-1-a La dégradation de l'alginate chez les organismes producteurs

Deux types d'organismes principalement produisent de l'alginate : des bactéries et des algues brunes. Il faut noter également que de l'alginate a été détecté chez certaines algues rouges calcaires de la famille des Corallinaceae (Okazaki *et al.*, 1982). On retrouve des alginate lyases chez les organismes qui produisent ce polysaccharide.

A ce jour, les bactéries productrices d'alginate appartiennent exclusivement aux genres *Pseudomonas* et *Azotobacter*, qui le sécrètent sous forme d'exopolysaccharide (Sabra *et al.*, 2001). Chez *Azotobacter*, l'alginate est un composé majoritaire du ciste : il forme une couche extracellulaire qui protège les cellules dans un environnement défavorable (ex. dessiccation) (Figure 1-22). Les cystes sont capables de survivre pendant des années dans des sols desséchés et de "germer" lorsque les conditions redeviennent favorables (Wyss *et al.*, 1961). Il a été montré que l'activité alginate lyase joue un rôle important dans ces processus de germination (Gimmestad *et al.*, 2009).



**Figure 1-22 : Encapsulation et germination chez *A. vinelandii***

Microscopie électronique (x 25000). **A.** Cellule avant l'encapsulation. **B.** Cyste encapsulé dans une couche notamment constituée d'alginate. **C.** Cyste après germination. La capsule est dégradée. D'après Wyss *et al.* (1961).

La bactérie pathogène *Pseudomonas aeruginosa* qui infecte les malades de la mucoviscidose (fibrose kystique) excrète également de l'alginate. Celui sert à protéger les cellules bactériennes des agents anti-microbiens (système immunitaire, antibiotiques) et favorise l'adhésion des cellules et l'implantation dans les tissus. Dans ce cas, l'action des alginate lyases permet de fluidifier la matrice de polysaccharides et libère les bactéries qui peuvent alors coloniser d'autres sites d'infection (Boyd and Chakrabarty 1994).

Enfin, des alginate lyases ont été isolées à partir des algues *Laminaria digitata* (Madgwick *et al.*, 1973), *Pelvetia canaliculata* (Madgwick *et al.*, 1978) et *Undaria pinnatifida* (Watanabe and Nishizawa 1982). Des activités alginate lyases ont aussi été détectées dans des extraits bruts préparés à partir d'autres espèces (Shiraiwa *et al.*, 1975). Les auteurs ont noté que l'activité est plus forte lors des périodes où les algues sont fertiles et dans les parties âgées. Aucun gène d'alginate lyase n'a pu être détecté dans le génome de l'algue brune *Ectocarpus siliculosus*, le seul séquencé à ce jour (Cock *et al.*, 2010). Les alginate lyases d'algues sont donc vraisemblablement trop divergentes pour trouver des homologies de séquences avec les protéines bactériennes : elles pourraient constituer une (des) nouvelle(s) famille(s) (Michel *et al.*, 2010). La fonction biologique de ces enzymes serait de remodeler la paroi au cours de la croissance et de la différenciation cellulaire.

#### **IV-1-b La dégradation de l'alginate chez les organismes non-producteurs**

Chez les organismes ne produisant pas d'alginate, ce polysaccharide peut la plupart du temps être utilisé comme seule source de carbone, ce qui démontre l'existence d'une voie catabolique complète. Des activités de dégradation de l'alginate ont été retrouvées chez de multiples organismes procaryotes et eucaryotes (Wong *et al.*, 2000). Comme on peut s'y attendre, on compte parmi eux de nombreuses espèces marines, qui ont accès à l'abondance d'alginate dans la paroi des algues brunes et les détritiques en suspension. Il s'agit surtout de bactéries marines, principalement des phyla *Gammaproteobacteria* et *Bacteroidetes*. Chez la plupart de ces bactéries, la production des alginate lyases est induite par la présence du substrat dans le milieu (par exemple Stevens and Levin 1977; Doubet and Quatrano 1984; Weiner *et al.*, 2008). Cependant, chez certaines espèces, l'expression d'une alginate lyase peut être constitutive (Boyen *et al.*, 1990a; Sawabe *et al.*, 1992). Il est intéressant de remarquer que des bactéries alginolytiques ont été retrouvées dans des sédiments à 250 m de profondeur, au

large des côtes japonaises (Kobayashi *et al.*, 2009; Uchimura *et al.*, 2010). Ceci pourrait indiquer que la dégradation de l'alginate continue même lorsque la matière organique atteint le fond des océans.

Concernant les eucaryotes, des champignons marins isolés d'algues en décomposition peuvent dégrader l'alginate, tout comme des invertébrés marins herbivores se nourrissant d'algues brunes. Ainsi, des alginate lyases ont pu notamment être purifiées à partir des genres *Haliotis* (orveau), *Aplysia* (lièvre de mer), *Turbo* et *Littorina* (Elyakova and Favorov 1974; Muramatsu and Egawa 1982; Boyen *et al.*, 1990b). Cependant, un doute subsiste quand à l'origine de ces enzymes. Une hypothèse alternative serait qu'elles soient produites par des bactéries symbiontes des mollusques marins étudiés. Des bactéries alginolytiques ont en effet été isolées à partir de fluides gastriques de différents invertébrés marins tels que l'orveau et l'oursin (Sawabe *et al.*, 1995; Sawabe *et al.*, 1998). Le fait de trouver des bactéries alginolytiques dans l'intestin d'organismes friands d'algues brunes soulève une question : qu'en est-il chez l'Homme ? Cette question peut sembler incongrue, mais on sait que les bactéries intestinales (le "microbiome") participent activement à la dégradation de polysaccharides végétaux résistants aux enzymes digestives (cf. Chapitre 2). Par ailleurs, comme je l'ai mentionné précédemment, l'alginate est largement utilisé dans l'industrie agro-alimentaire et donc nous en consommons régulièrement. Des études sur la flore intestinale humaine ont montré une légère dégradation de l'alginate par les bactéries intestinales. Sur 25 espèces fréquemment rencontrées dans l'intestin (représentant 300 souches différentes), seules 10 souches de l'espèce *Bacteroides ovatus* ont été montrées comme capables de dégrader l'alginate (Michel and Macfarlane 1996). Seule une dégradation partielle de l'alginate a pu être obtenue *in vitro* en utilisant des populations mixtes de bactéries intestinales. Il semble donc que la plupart des bactéries intestinales n'aient pas la capacité de dégrader l'alginate alimentaire, et que cette source de matière soit perdue à la fois pour l'hôte humain et la majorité des bactéries résidentes du tractus digestif. De manière intéressante, il a été montré que les produits de dégradation de l'alginate accélèrent la croissance de bifidobactéries, un autre groupe de bactéries intestinales considérées comme bénéfique pour la santé (Akiyama *et al.*, 1992).

Des bactéries Gram-négatives et Gram-positives dégradant l'alginate ont aussi été retrouvées dans les sols. Nakagawa *et al.* (1998) ont montré que l'arrosage du sol avec une solution d'alginate de sodium pendant trois mois permet d'augmenter la population de bactéries alginolytiques et facilite leur isolement (6 souches isolées dans les sols traités, zéro dans les sols contrôles). Dans leur environnement naturel, ces bactéries alginolytiques

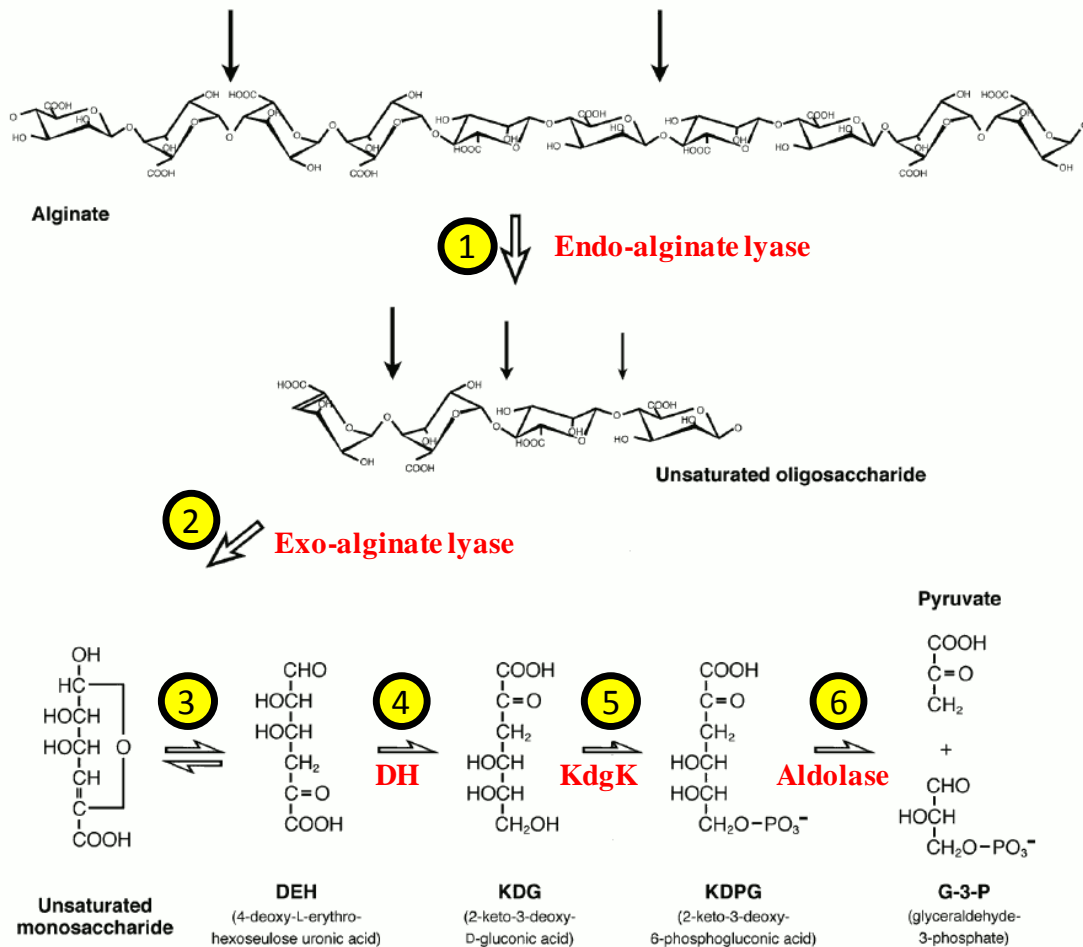


pourraient utiliser l'alginate acétylé excrété par d'autres bactéries (Hansen *et al.*, 1984; Nakagawa *et al.*, 1998).

Enfin, certains virus produisent des alginate lyases. C'est le cas de bactériophages spécifiques pour *Azotobacter* et *Pseudomonas* (Bartell *et al.*, 1966). Les enzymes aident les phages à pénétrer dans la matrice d'exopolysaccharides sécrétés par ces bactéries. Une lyase extraite du virus de la microalgue *Chlorella* (dont elle dégrade les résidus glucuronates présents dans la paroi) est aussi active sur l'alginate (Ogura *et al.*, 2009).

## IV-2 La voie catabolique de l'alginate

Dès les années 1960, des études sur une bactérie du genre *Pseudomonas* ont permis de proposer une voie catabolique complète de l'alginate, de la dégradation du polymère de départ jusqu'à l'assimilation dans les voies du métabolisme central (Preiss and Ashwell 1962a; Preiss and Ashwell 1962b). Les réactions impliquées semblent conservées chez les bactéries, puisque la voie est la même chez *Sphingomonas* sp. A1 (Takase *et al.*, 2010). La voie catabolique complète est représentée sur la Figure 1-23.



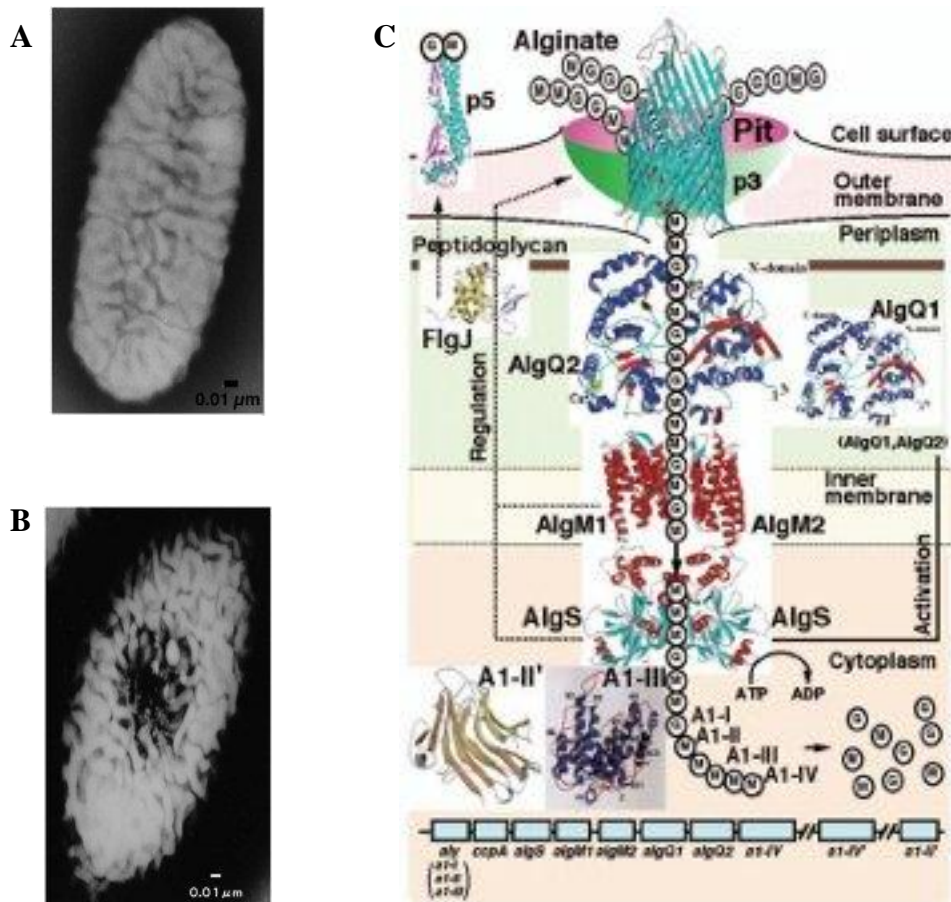
**Figure 1-23 : Voie catabolique de l'alginate chez les bactéries.**

DH : déshydrogénase. KdgK : cétodésoxygluconate kinase. Image modifiée d'après Takase *et al.* (2010).

La première étape est la dépolymérisation de l'alginate par les alginate lyases (réactions 1 et 2), soit en milieu de chaîne (mode d'action endo) soit en bout de chaîne (mode d'action exo). Ces enzymes seront présentées dans le détail dans la partie suivante. Le produit final est un monosaccharide insaturé qui est converti de manière non-enzymatique en acide uronique 4-désoxy-L-erythro-5-hexoseulose (DEH) (réaction 3). Une déshydrogénase prend alors en charge le DEH pour le convertir en acide cétodésoxygluconique (réaction 4), lui-même phosphorylé par une kinase KdgK (réaction 5). Cette voie catabolique aboutit à la formation d'acide cétodésoxyphosphogluconique (KDPG), un intermédiaire de la voie d'Entner-Doudoroff (Entner and Doudoroff 1952). L'ensemble de ces réactions permet l'assimilation de l'alginate dans le métabolisme bactérien et donc son utilisation comme source de carbone et d'énergie. Pour utiliser l'alginate comme substrat, les bactéries doivent posséder cette voie complète. Cependant, on peut envisager que certaines espèces n'effectuent que les dernières réactions et utilisent les produits de la dégradation de l'alginate effectuée par d'autres bactéries. C'est par exemple le cas pour la souche C58 d'*Agrobacterium tumefaciens*, qui ne pousse pas sur l'alginate intact, mais peut utiliser des produits de dégradation de plus faible degré de polymérisation (Ochiai *et al.*, 2006).

Dans la plupart des cas, les premières étapes de la dégradation de l'alginate ont lieu à l'extérieur de la cellule et sont catalysées par des enzymes sécrétées ou adressées à la membrane externe. En effet, les polysaccharides sont trop gros pour pouvoir pénétrer dans le périplasme *via* des transporteurs classiques (taille maximale ~ 600 Da). La bactérie *Sphingomonas* sp.A1 a cependant développé un système efficace d'import du polysaccharide directement dans le cytoplasme par un "supercanal" (Figure 1-24). Ce système est composé d'un trou dans la membrane externe (parfois comparé à une bouche), de récepteurs TonB dépendants sur la membrane externe, de protéines périplasmiques qui fixent le substrat et d'un ABC-transporteur lié à la membrane cytoplasmique (Hashimoto *et al.*, 2010). Des récepteurs à alginate sont présents à la surface des cellules et assurent la perception du signal, nécessaire à la mise en place de ce système. Une fois que le polymère rentre dans le cytoplasme, il est dégradé par les alginates lyases cytoplasmiques en monosaccharides (Hashimoto *et al.*, 2000; Yoon *et al.*, 2000). Les gènes qui codent ce système sont regroupés dans un cluster. Il apparaît comme une adaptation à la dégradation de l'alginate dans un environnement compétitif. Les produits de dégradation ne sont jamais libres dans le milieu extérieur, et donc ne peuvent pas être détournés par d'autres micro-organismes. Un cluster de gènes homologue à celui de

*Sphingomonas* sp.A1 a été détecté chez *Agrobacterium tumefaciens* (Ochiai et al., 2006). Cependant, ce type de "supersystème" de dégradation semble peu répandu dans la nature.



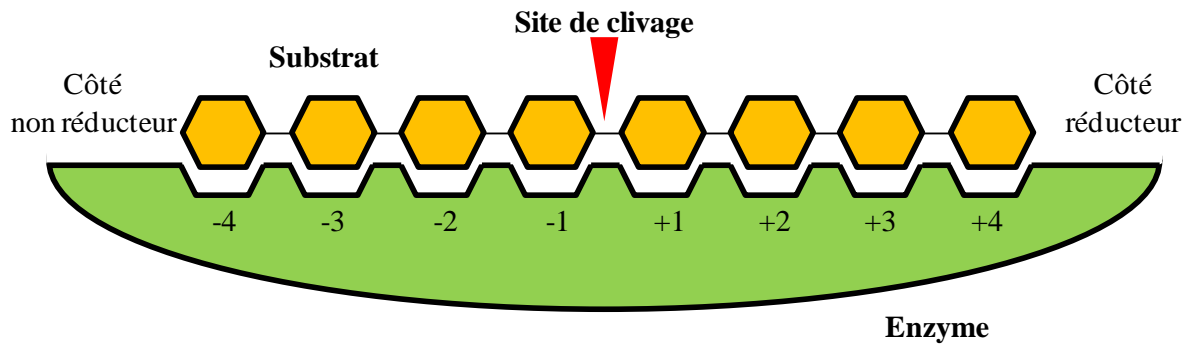
**Figure 1-24 : Système d'import et de dégradation de l'alginate chez *Sphingomonas* sp.A1**  
A, B. Microscopie électronique de cellules cultivées sans alginate (A) ou avec (B). La présence d'alginate induit la formation d'un pore dans la membrane, dont la structure est présentée sur le schéma C. D'après Hashimoto *et al.* (2010).

### IV-3 Les alginate lyases

#### IV-3-a Propriétés enzymatiques

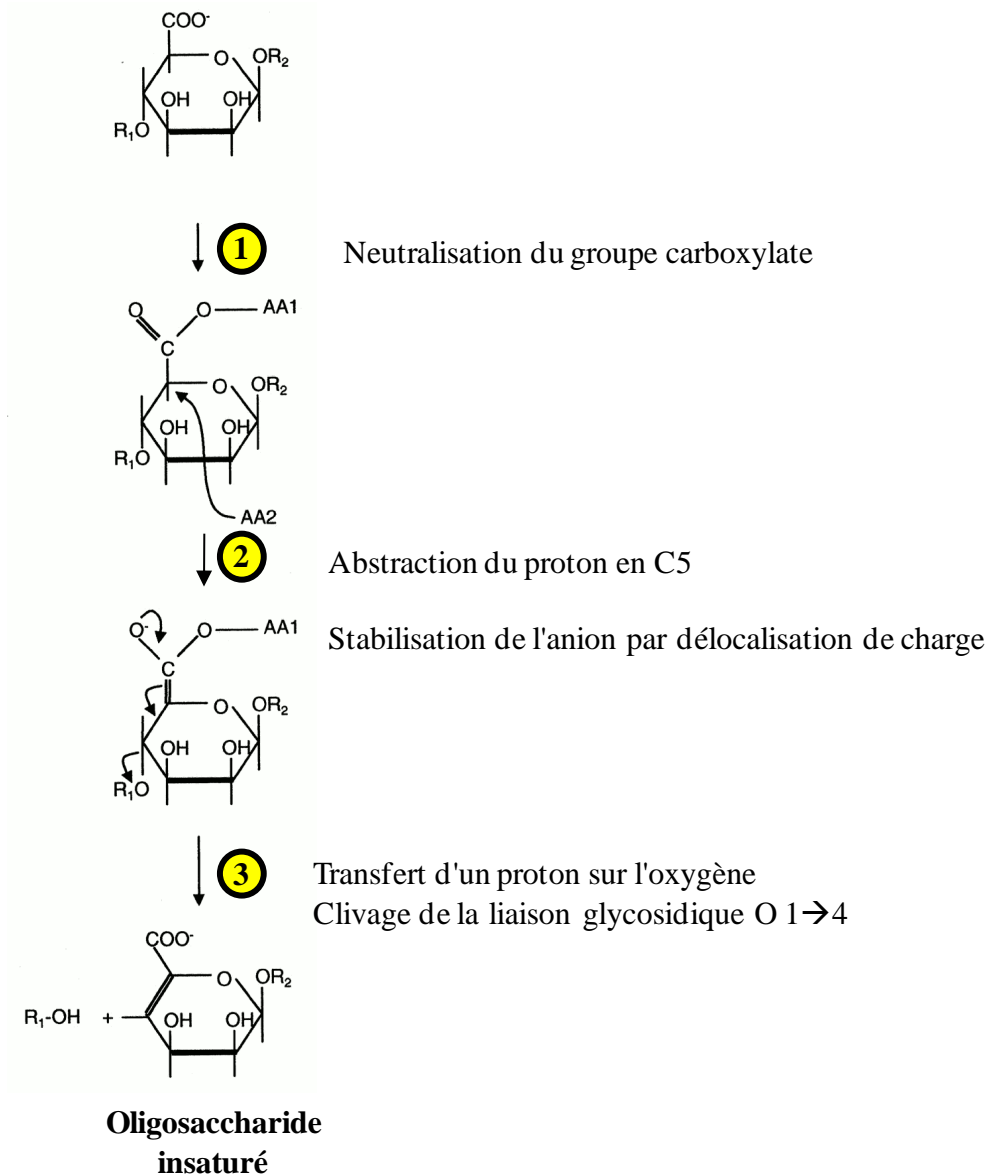
Les polysaccharides lyases (PL, EC 4.2.2.-) constituent un groupe d'enzymes qui clivent les chaînes de polysaccharides contenant des acides uroniques par un mécanisme de  $\beta$ -élimination. Afin de modéliser l'action de ces enzymes, on peut adapter la nomenclature utilisée pour les glycoside hydrolases (Davies *et al.*, 1997). Ainsi, on définit au sein de la poche catalytique de ces enzymes des sous-sites de liaison au substrat (Figure 1-25). Par définition, le clivage de la chaîne a lieu entre le sous-site - 1 et le sous-site + 1. Les sous-sites

négatifs sont ceux du côté de la nouvelle extrémité réductrice, et les sous-sites positifs sont du côté de la nouvelle extrémité non-réductrice. Le clivage par une polysaccharidase génère un résidu uronique insaturé au sous-site +1 et une nouvelle extrémité réductrice au sous-site -1.



**Figure 1-25 : Nomenclature des sous-sites dans une polysaccharidase.**  
D'après Davies *et al.* (1997).

Le mécanisme d'action d'une alginate lyase ( $\beta$ -élimination) proposé pour la première fois par Gacesa (1987) est représenté sur la Figure 1-26 page suivante. Il n'implique pas l'hydrolyse par une molécule d'eau. La première étape est la neutralisation de la charge négative portée par le groupement carboxylate en C<sub>6</sub> du résidu situé dans le sous-site +1 de l'enzyme. Pour cela, un pont salin est formé avec la chaîne latérale d'un acide aminé chargé positivement (AA1 dans la Figure 1-26) ou un cation divalent. Il y a ensuite abstraction du proton en C<sub>5</sub> par la chaîne latérale d'un résidu agissant comme une base (AA2). Des acides aminés tels qu'un acide aspartique, un acide glutamique, une histidine, une lysine ou une cystéine ont été suggérés pour intervenir à cette étape. L'anion énolate créé est alors stabilisé par délocalisation de charge. Enfin, un proton est transféré sur l'oxygène de la liaison glycosidique. Cette étape est facilitée par l'apport d'un proton par un autre acide aminé catalytique. De manière alternative, le proton pourrait aussi être fourni directement par le solvant. Ceci aboutit à la formation d'une double liaison entre les carbones C<sub>4</sub> et C<sub>5</sub> du résidu dans le sous-site +1 et à l'élimination de la liaison glycosidique O 1 $\rightarrow$ 4. Que l'on parte d'un motif guluronate ou mannuronate, le motif obtenu à l'extrémité non-réductrice de l'oligosaccharide est le même : l'acide uronique 4-désoxy-L-érythro-hex-4-enopyranosyluronique. Par convention, ce motif est noté delta ( $\Delta$ ).



**Figure 1-26 : Mécanisme de  $\beta$ -élimination d'une alginate lyase**

Schéma modifié d'après Svanem *et al.* (2001)

La conjugaison entre la double liaison  $\text{C}_4=\text{C}_5$  formée et le groupement carboxyl en  $\text{C}_6$  confère aux produits de la réaction un pic d'absorption en UV à 230-235 nm (Preiss and Ashwell 1962a). La réaction peut donc facilement être suivie au spectrophotomètre. Une autre méthode couramment utilisée est le test colorimétrique à l'acide thiobarbiturique (TBA), qui réagit avec les sucres insaturés pour donner un composé absorbant à 548 nm.

Comme je l'ai déjà mentionné, les alginate lyases peuvent cliver le polysaccharide soit en milieu de chaîne (mode d'action endolytique), soit en bout de chaîne (mode exolytique). La

grande majorité des alginate lyases caractérisées possède un mode d'action endolytique (Wong *et al.*, 2000). Quelques exo-alginate lyases ont été décrites, par exemple HdAlex chez *Haliotis discus hannai* (Suzuki *et al.*, 2006), A1-IV chez *Sphingomonas* sp. A1 (Hashimoto *et al.*, 2000) ou Atu3025 chez *Agrobacterium tumefaciens* (Ochiai *et al.*, 2006). De manière intéressante, certaines enzymes peuvent changer de mode d'action, par exemple en fonction du pH (Ogura *et al.*, 2009).

La diversité de structure de l'alginate, due à l'existence des deux types de motifs M et G et aux différents enchaînements possibles, se reflète dans la spécificité de substrat des alginate lyases. En effet, certaines clivent spécifiquement entre deux résidus G : on parle alors de guluronate lyases (EC 4.2.2.11). Il existe aussi des mannuronate lyases (EC 4.2.2.3) qui attaquent la liaison glycosidique entre deux motifs M. Certaines enzymes décrites comme plus spécifiques pour un type de motif peuvent tout de même présenter une activité résiduelle sur des homopolymères de l'autre motif (Wong *et al.*, 2000).

La température optimale de la plupart des alginate lyases de bactéries marines est comprise entre 25°C et 50°C. Leur pH optimum varie généralement entre 7,0 et 8,5. La présence de cations ( $\text{Na}^+$ ,  $\text{Ca}^{2+}$  par exemple) peut être nécessaire voire indispensable à l'activité. Les cations pourraient jouer un rôle dans la stabilisation de l'intermédiaire énolate formé au cours de la réaction (Figure 1-26).

#### **IV-3-b Classification et détermination structurale**

Comme les autres enzymes actives sur les sucres (Carbohydrate-Active enZymes, CAZymes), les polysaccharides lyases sont recensées dans la base de données CAZy (<http://www.cazy.org/>, Cantarel *et al.*, 2009). En avril 2011, 21 familles de PL sont recensées dans CAZy (PL1 à PL22, la famille PL19 ayant été reclassée dans les glycoside hydrolases). Cette classification hiérarchique est basée sur des similarités de séquences protéiques. Par conséquent, les familles reflètent plus un mécanisme d'action et un repliement commun qu'une spécificité de substrat. Pour tenter de faciliter la prédiction fonctionnelle des nouvelles PL trouvées dans les génomes, les familles de PL ont récemment été subdivisées en sous-familles (Lombard *et al.*, 2010). Les auteurs ont ainsi créé 41 sous-familles, et ont montré que 90% d'entre elles sont monospécifiques, c'est-à-dire contiennent des enzymes avec la même spécificité de substrat. La classification en sous-familles permet donc de prédire avec une

bonne confiance la spécificité de nouvelles enzymes non caractérisées. Cette prédiction doit toutefois être confirmée par des analyses biochimiques.

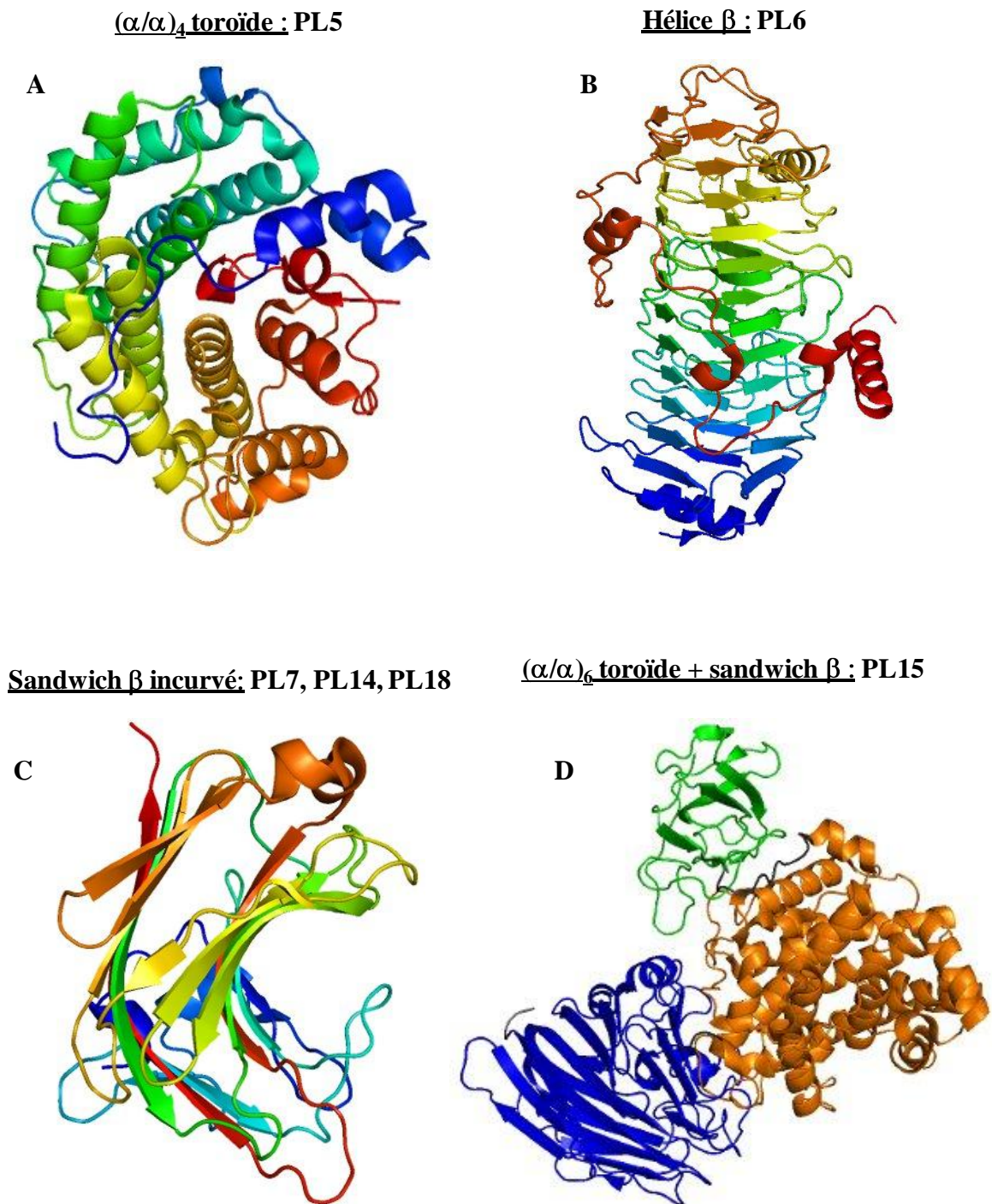
Les alginate lyases se répartissent dans 7 familles CAZy : PL5 (45 protéines), PL6 (36 protéines), PL7 (101 protéines), PL14 (49 protéines), PL15 (15 protéines), PL17 (47 protéines) et PL18 (6 protéines). A ce jour, seules des activités alginate lyases ont été démontrées pour ces familles, sauf dans PL6 qui contient aussi des chondroitinases. La majorité des alginate lyases caractérisées sont classées dans les familles PL5 et PL7 (Tableau 1-5).

Famille	Repliement	Acides aminés catalytiques			Sous-famille	Alginate lyases caractérisées	Structures disponibles	Spécificité
		Neutraliseur	Base	Acide				
PL5	$(\alpha/\alpha)_4$ toroïde	Arg/Asn	Tyr	Tyr		10	1	
					SF1	9	0	MM
					N.C.	1	1	MM
PL6	Hélice $\beta$	$\text{Ca}^{2+}$	Lys	Arg		1	0	
					N.C.	1	0	MM
PL7	$\beta$ -Jelly roll	Gln	His	Tyr		16	3	
					SF1	1	1	MG
					SF2	0	0	-
					SF3	2	1	GG
					SF4	0	0	-
					SF5	4	0	GG
N.C.	9	1	MM, GG, MG					
PL14	$\beta$ -Jelly roll	N.D.	N.D.	N.D.		4	1	
					SF1	1	1	glucuronate
					SF2	0	0	-
					SF3	2	0	MM
N.C.	1	0	MM					
PL18	$\beta$ -Jelly roll	Gln	His	Tyr	N.C.	4	1	MM, GG
PL15	tonneau $(\alpha/\alpha)_6$ + sandwich $\beta$	Arg	His	Tyr		4	1	
					SF1	3	1	-
					SF2	0	0	-
N.C.	1	0	MM					
PL17	N.D.	N.D.	N.D.	N.D.		1	0	
					SF1	0	0	-
					SF2	1	0	MM

Tableau construit d'après les données recensées dans CAZy (<http://www.cazy.org/>) en avril 2011. D'après Lombard *et al.* (2010) et Garron *et al.* (2010)

**Tableau 1-5 : Classification et caractéristiques des alginate lyases**





**Figure 1-27 : Les 4 types de repliement connus dans les familles d'alginate lyases**

A. A1-III de *Sphingomonas* sp.A1 (PDB 1HV6, Yoon *et al.*, 2001). B. Chondroitinase B de *P. heparinus* (1DBG, Huang *et al.*, 1999). C. A1-II' de *Sphingomonas* sp.A1 (2CWS, Yamasaki *et al.*, 2005) D. Atu3025 d'*A. tumefaciens* (3AFL, Ochiai *et al.*, 2010). Images générées dans PyMOL.



Dix alginate lyases ont été caractérisées biochimiquement dans la famille PL5. Elles ont toutes une préférence pour les blocs de motifs M. La structure de l'alginate lyase A1-III de *Sphingomonas* sp. A1 a été résolue (Yoon *et al.*, 2001). Cette protéine adopte un repliement en toroïde incomplet ( $\alpha/\alpha$ )<sub>4</sub>. (Garron and Cygler 2010) Elle est constituée d'une répétition de 4 paires d'hélices  $\alpha$  antiparallèles, qui forment des épingles (*hairpins*) (Figure 1-27 A). De courtes boucles séparent deux hélices d'une même épingle, alors qu'on trouve des connections plus longues entre deux épingles successives. Ces épingles sont arrangées en un demi-tonneau et forment une gorge dans laquelle se trouve le site actif. Les rôles d'acide et de base ont été attribués au même acide aminé, Tyr246, qui interviendrait donc deux fois dans le mécanisme catalytique. Le groupe carboxyl du sucre dans le sous-site +1 peut être neutralisé par l'arginine Arg239 et possiblement par Asn191 (Yoon *et al.*, 2001).

Une alginate lyase a été caractérisée dans la famille PL6. Il s'agit d'AlyP de *Pseudomonas* sp. OS-ALG-9 (Maki *et al.*, 1993). La structure de cette protéine n'est pas connue, mais celle d'une chondroitinase B de *Pedobacter heparinus* appartenant également à la famille PL6 a été déterminée. Du fait des similarités de séquences, on peut penser que les alginate lyases de la famille PL6 adoptent le même repliement. La chondroitinase possède une structure en hélice  $\beta$  droite (Huang *et al.*, 1999). Elle est constituée de trois feuillets  $\beta$ , qui forment une hélice à 11 tours (Figure 1-27 B). Les feuillets  $\beta$  PB1 et PB2 sont quasiment antiparallèles, et PB3 pratiquement perpendiculaire à PB2. Dans cette chondroitinase, un ion calcium joue le rôle de neutraliseur de la charge négative du groupement carboxylate (Michel *et al.*, 2004). Le résidu basique qui abstrait le proton en C<sub>5</sub> est une lysine (Lys250) et une arginine (Arg271) joue le rôle du donneur de proton (Tableau 1-5).

Les enzymes des familles PL7, PL14 et PL18 partagent le même repliement en sandwich  $\beta$  incurvé ( $\beta$  *jelly roll*). PL7 est la famille d'alginate lyases la mieux caractérisée, puisque l'activité a été démontrée pour 16 enzymes, dont trois de structures connues (Yamasaki *et al.*, 2004; Osawa *et al.*, 2005; Yamasaki *et al.*, 2005; Ogura *et al.*, 2008). La famille PL14 contient 4 alginate lyases caractérisées. La seule dont la structure est connue, l'enzyme virale vAL-1, clive les liaisons entre les résidus glucuronates du composé pariétal d'une micro-algue verte, mais présente une activité alginate lyase à pH basique (Ogura *et al.*, 2009). La famille PL18 comporte quant à elle 4 alginate lyases caractérisées, dont une structurellement. Le repliement en  $\beta$  *jelly roll* est constitué de deux feuillets  $\beta$  antiparallèles, présentant une courbure approchant un angle de 90° en leur milieu (Garron and Cygler 2010)

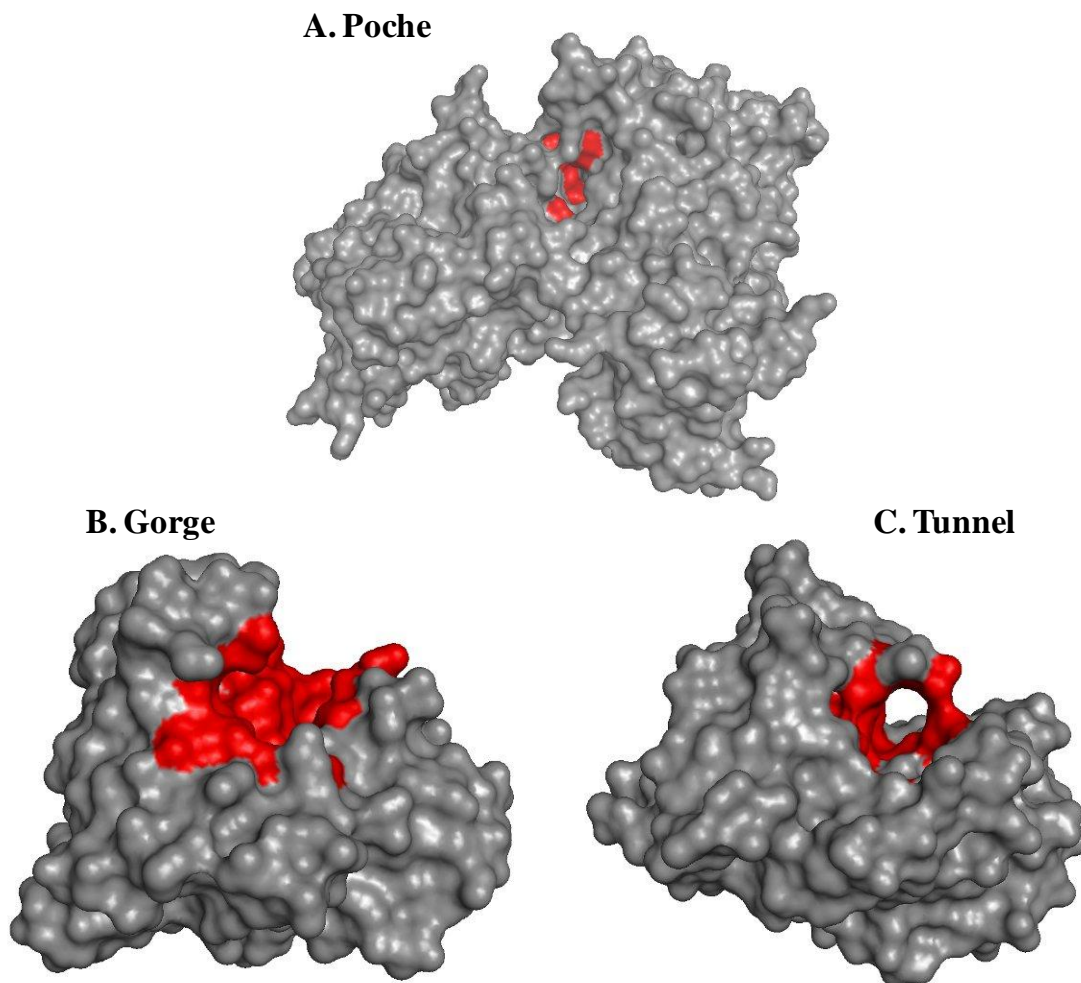
(Figure 1-27 C). Ceci crée une gorge dans laquelle le substrat se fixe et où a lieu la catalyse. Deux boucles flexibles peuvent interagir entre elles par des liaisons hydrogène entre deux résidus asparagine et forment ainsi une structure de "couvercle" au-dessus de la gorge. Par mutagenèse dirigée, Ogura *et al.* (2008) ont montré que la flexibilité de ces deux boucles est essentielle à la liaison du substrat dans la gorge. Les résidus catalytiques ont été déterminés et sont identiques pour les enzymes des familles PL7 et PL18. Il s'agit d'une glutamine qui neutralise la charge du groupement carboxyl, d'une histidine jouant le rôle de base et d'une tyrosine donneuse de proton. Aucun de ces résidus n'est conservé dans la famille PL14. Par mutagenèse dirigée, des acides aminés importants pour l'activité et conservés au sein de la famille ont été mis en évidence (Ogura *et al.*, 2009). Cependant, les acides aminés catalytiques n'ont pas encore été déterminés précisément.

Quatre alginate lyases ont été caractérisées dans la famille PL15. La structure de l'une d'entre elles est résolue, celle d'Atu3025 d'*Agrobacterium tumefaciens* (Ochiai *et al.*, 2010). Elle adopte un repliement en trois domaines (Figure 1-27 D): un petit domaine N terminal en feuillet  $\beta$ , un domaine  $(\alpha/\alpha)_6$  toroïde central et un domaine C terminal en sandwich  $\beta$ . Les acides aminés catalytiques sont l'arginine R199 (neutraliseur), l'histidine H311 (base) et la tyrosine Y365 (acide).

Enfin, on ne compte qu'une seule alginate lyase caractérisée dans la famille PL17. Il s'agit de AlyII de *Pseudomonas* sp. OS-ALG-9 (Kraiwanapong *et al.*, 1999). Cette famille de polysaccharide lyases reste en grande partie méconnue. Aucun acide aminé catalytique n'a été proposé pour le moment, et aucune structure n'a été obtenue.

Indépendamment de la famille structurale à laquelle elles appartiennent, les enzymes agissant sur des polysaccharides peuvent adopter trois types de topologie différents: en forme de poche, de sillon ou de tunnel (Davies and Henrissat 1995). Dans la topologie en "poche", le site catalytique se trouve enfoui au fond d'une dépression de la surface moléculaire. Cette topologie est optimale pour la reconnaissance de l'une ou l'autre des extrémités d'une chaîne de polysaccharide. On la retrouve donc chez les enzymes agissant de manière exolytique. C'est par exemple le cas de l'alginate lyase Atu3025 d'*Agrobacterium tumefaciens* (Ochiai *et al.*, 2010) (Figure 1-28A). La topologie ouverte en "gorge" ou "sillon" est adaptée à la fixation aléatoire le long de la chaîne du polysaccharide et est souvent retrouvée dans les endo-enzymes. Cette topologie est notamment celle de l'alginate lyase ALY-1 de *Corynebacterium* sp. (Osawa *et al.*, 2005) (Figure 1-28B). La topologie en "tunnel" est dérivée de la précédente.

Les protéines qui l'adoptent possèdent de longues boucles qui peuvent venir refermer la gorge de fixation au substrat. Ceci permet aux enzymes de relâcher le produit de la réaction tout en restant fermement ancrées au polysaccharide, créant les conditions nécessaires à la processivité. Les boucles couvrant le tunnel peuvent être flexibles, ce qui autorise une ouverture du site de fixation pour une attaque initiale en milieu de chaîne. Notamment, l'alginate lyase A1-II' de *Sphingomonas* sp.A1 adopte ce type de topologie (Ogura *et al.*, 2008) (Figure 1-28C).



**Figure 1-28 : Exemple de topologies adoptées par les polysaccharidases**

**A.** Topologie en poche de l'alginate lyase Atu3025 d'*Agrobacterium tumefaciens* (PDB n° 3AFL). **B.** Topologie en gorge ouverte de l'alginate lyase ALY-1 de *Corynebacterium* sp. (PDB n° 1UAI). **C.** Topologie en tunnel de l'alginate lyase A1-II' de *Sphingomonas* sp. A1 (PDB n° 2ZAB). Les sites catalytiques et de fixation au substrat sont colorés en rouge. Figures générées dans PyMOL.

Outre leurs domaines catalytiques, les polysaccharidases peuvent posséder des modules structuraux auxiliaires appelés *Carbohydrate Binding Modules* (CBM). Un CBM est défini comme une séquence contiguë d'acides aminés au sein d'une enzyme active sur les sucres, qui se replie de manière indépendante et possède une activité de liaison aux sucres (Boraston *et al.*, 2004). Dans la plupart des cas, un CBM fixe spécifiquement le substrat du domaine catalytique auquel il est associé. Dans le cas des familles polyspécifiques CBM6 et CBM32, il a été montré que les CBM ont co-évolué avec leur domaine catalytique vers une même spécificité de substrat (Abbott *et al.*, 2008; Michel *et al.*, 2009). Comme les modules catalytiques, les CBM sont également recensés dans la base de données CAZY (<http://www.cazy.org/>, Cantarel *et al.*, 2009). A ce jour, deux familles de CBM ont été prédites dans des séquences d'alginate lyases. Il s'agit des familles CBM16 (dans AlyA de *Pseudoalteromonas atlantica*, alyPEEC de *Pseudoalteromonas* sp. IAM14594 et Aly-SJ02 de *Pseudoalteromonas* sp. SM0524) et CBM32 (dans AlyPI de *Pseudoalteromonas* sp. CY24). Cependant, la capacité de fixation de l'alginate n'a pas encore été démontrée pour ces modules.

#### IV-3-c Les applications des alginate lyases

Les alginate lyases ont trouvé des applications dans des domaines nombreux et variés. Elles sont utilisées comme outils enzymatiques pour élucider la structure fine des alginates. En effet, en analysant les produits de dégradation obtenus avec une enzyme de spécificité et de mode d'action connus, on peut déduire des informations sur la structure du polysaccharide de départ (Fujibaya *et al.*, 1970; Min *et al.*, 1977; Boyd and Turvey 1978; Ostgaard 1993). Les alginate lyases sont également des outils de choix pour dégrader la paroi des algues brunes en vue de la préparation de protoplastes, des cellules nues et viables. Des protoplastes ont pu être isolés chez une trentaine d'espèces de Phaeophyceae et utilisés notamment pour des études sur la synthèse de la paroi et la régénération d'un individu (Reddy *et al.*, 2008). Ils pourraient également être des cibles intéressantes pour la transformation génétique qui n'est à ce jour pas possible chez les algues brunes.

L'usage des alginate lyases s'est aussi répandu pour la production d'oligosaccharides qui peuvent présenter différents effets physiologiques. En biologie végétale, il a été montré que des oligoalginates insaturés favorisent la germination et la croissance racinaire chez

différentes espèces de plantes terrestres (Murata *et al.*, 1993; Natsume *et al.*, 1994). Ils sont également capables d'éliciter des réponses de défense et pourraient donc être utilisés pour la protection des cultures (Akimoto *et al.*, 1999; An *et al.*, 2009). Les oligoalginates ont aussi un effet sur les cellules animales. Ils induisent par exemple la prolifération de kératinocytes *in vitro* en présence du facteur de croissance EGF (Kawada *et al.*, 1997). Ils stimulent la production de cytokines chez la souris (Yamamoto *et al.*, 2007) et sur des cellules humaines (Iwamoto *et al.*, 2003). Des activités anti-tumorales ont pu être mises en évidence pour des dérivés oligoalginates sulfatés (Hu *et al.*, 2004).

Une autre application thérapeutique des alginate lyases est envisagée dans le traitement de la mucoviscidose. Comme je l'ai déjà évoqué, les poumons de patients atteints de fibrose kystique sont infectés par une souche de *Pseudomonas aeruginosa* qui produit de l'alginate comme exopolysaccharide. En association avec de l'ADN excrété, il forme un mucus dense qui favorise l'adhésion des bactéries et les protège des cellules du système immunitaire et des antibiotiques (May *et al.*, 1991). L'utilisation d'alginate lyases pour dégrader cette couche muqueuse pourrait rendre les bactéries plus accessibles et donc faciliter l'action des macrophages (Eftekhari and Speert 1988) et des antibiotiques (Hatch and Schiller 1998; Alkawash *et al.*, 2006). Cette approche est encore en développement et comporte plusieurs limitations. Tout d'abord, l'alginate lyase utilisée doit être active sur l'alginate bactérien, qui est acétylé. Elle doit être tolérée par l'organisme et ne pas induire de réaction immunitaire. Sur ce point, il semble que la modification chimique des enzymes par du polyéthylène glycol soit une piste encourageante (Lamppa *et al.*, 2011). Enfin, la dégradation du biofilm pourrait favoriser la dissémination de *P. aeruginosa* vers d'autres sites d'infection.

## V- Présentation du travail

Ce projet de thèse s'inscrit dans le contexte d'étude des interactions entre algues et procaryotes et d'utilisation des ressources en composés organiques disponibles dans les écosystèmes marins. Il se propose d'étudier le métabolisme des polysaccharides pariétaux des algues par une bactérie marine à fort potentiel de dégradation, *Zobellia galactanivorans*. Ce projet est centré sur l'étude du métabolisme de l'alginate, un polysaccharide majeur de la paroi des algues brunes (Phaeophyceae). J'ai cherché à mieux comprendre les étapes de la perception et de la prise en charge du substrat par la bactérie. Je me suis particulièrement intéressé aux régulations métaboliques que cela engendre et aux différents types d'activités enzymatiques impliquées. L'objectif global est de comprendre les mécanismes impliqués dans la détection de ce substrat ainsi que les régulations mises en jeu, et d'étudier les différentes enzymes de dégradation (Figure 1-29). C'est un projet pluri-disciplinaire, alliant des approches de transcriptomique, de biochimie et de génomique structurale.

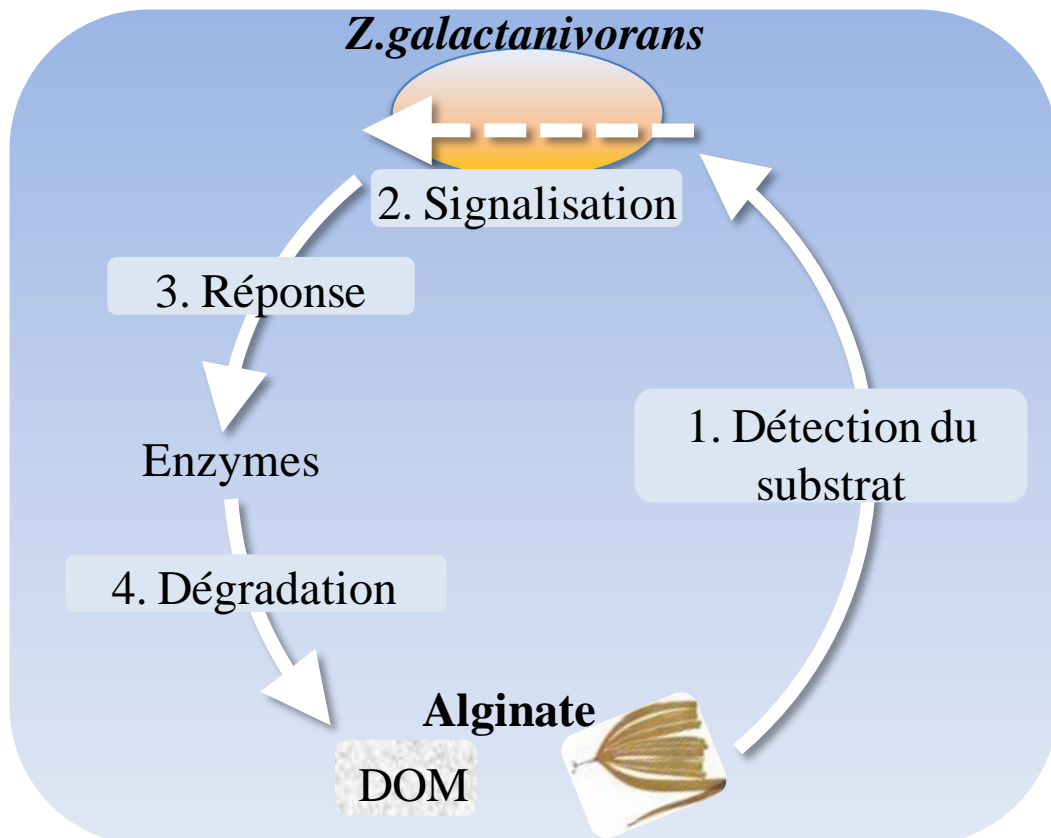


Figure 1-29 : Schéma présentant l'objectif de ce travail de thèse

Le travail présenté dans ce manuscrit est organisé de la façon suivante. Le **Chapitre 2** présente une synthèse bibliographique de l'importance des *Bacteroidetes* environnementaux, intestinaux et pathogènes pour la dégradation de la matière organique. Le **Chapitre 3** concerne la détermination de gènes de normalisation, qui pose les bases de l'étude transcriptomique de l'expression des gènes chez *Z. galactanivorans*. Le **Chapitre 4** est une étude bio-informatique et transcriptomique du système complet de dégradation de l'alginate par cette bactérie. J'étudie l'expression des différents gènes impliqués dans cette voie et mets en évidence pour la première fois l'existence d'opérons alginolytiques, conservés au sein des phyla des *Bacteroidetes* et *Proteobacteria*. Cette étude prometteuse a été élargie à l'ensemble du génome, par une approche transcriptomique par puces à ADN, dont les résultats sont présentés dans le **Chapitre 5**. Elle révèle les voies de transduction du signal potentiellement impliquées dans la dégradation de polysaccharides d'algues par *Z. galactanivorans* et dévoile de surprenants mécanismes d'inductions croisées.

Afin de caractériser biochimiquement les différents acteurs impliqués dans la dégradation de l'alginate, j'ai choisi d'exprimer un certain nombre d'entre eux de manière hétérologue chez *E. coli*. La mise en œuvre et les résultats de cette stratégie de clonage et d'expression à moyen débit sont présentés dans le **Chapitre 6**. La dernière partie du travail, présentée dans le **Chapitre 7**, a été la caractérisation phylogénétique, biochimique et cristallographique des alginate lyases de la famille PL7 de *Z. galactanivorans*.

En intégrant les résultats provenant des différentes approches, il est possible de dresser une vue globale des mécanismes nécessaires à la détection, l'import et l'utilisation de l'alginate par *Z. galactanivorans*. Ce travail suggère de nouvelles hypothèses quant aux interactions entre les bactéries hétérotrophes, la matière organique et les macroalgues. Il ouvre des perspectives qui sont discutées dans le **Chapitre 8**.

Chapitre 2  
*Bacteroidetes* intestinaux  
et environnementaux :  
*"The food connection"*

---





## Présentation de la revue

Les membres du phylum des *Bacteroidetes*, auquel appartient *Z. galactanivorans*, ont colonisé pratiquement toutes les niches écologiques. Dans l'environnement, ils ont une distribution quasi-ubiquitaire. Comme je l'ai déjà mentionné dans l'introduction, ils sont un des composants majeurs des communautés microbiennes en association avec les macroalgues. On les retrouve par ailleurs abondamment dans les sols, les océans et les eaux douces. D'autre part, les *Bacteroidetes* représentent une part importante du microbiome des animaux, particulièrement dans le tractus gastro-intestinal. Leur comportement peut varier de la symbiose avec l'hôte à la pathogénicité. Dans ces différents habitats, les *Bacteroidetes* sont reconnus comme étant spécialistes de la dégradation de la matière organique de haut poids moléculaire, comme les polysaccharides et les protéines.

Dans ce contexte, la revue présentée ci-après et publiée en 2011 dans le journal "*Frontiers in Microbiology*" propose un bilan des connaissances sur le rôle et les mécanismes de la dégradation des polysaccharides par les *Bacteroidetes*. Quel que soit leur environnement, la fonction biologique des *Bacteroidetes* semble liée à leur grande capacité à utiliser les biopolymères d'origine végétale ou animale. La description de nouveaux taxons dans ce phylum résulte souvent du criblage d'échantillons pour découvrir des activités enzymatiques originales avec de nouvelles applications potentielles. Cela a par exemple été le cas pour *Z. galactanivorans*, isolée pour sa capacité à dégrader les polysaccharides d'algues rouges (Potin *et al.*, 1991). L'ère de la génomique a permis le séquençage de nombreuses espèces au sein du phylum, la première d'entre elles étant *Bacteroides thetaiotaomicron*, un symbionte du système digestif de l'Homme. L'annotation des génomes confirme qu'ils codent un nombre important d'enzymes agissant sur les polysaccharides (CAZymes). Par exemple, le génome de *B. thetaiotaomicron* contient 44 glycoside hydrolases et polysaccharide lyases par mégabase. En plus de ces CAZymes, le séquençage des génomes révèle l'existence de deux familles de gènes paralogues largement représentées chez les *Bacteroidetes* : il s'agit des gènes codant les protéines homologues à SusC et SusD chez *B. thetaiotaomicron*. Ces protéines de la membrane externe sont impliquées dans la fixation et la dégradation de polysaccharides, comme il a été montré pour la première fois dans le cas de l'utilisation de l'amidon. Les protéines SusD sont uniques au phylum des *Bacteroidetes*. Au sein des génomes, les paires de gènes *susC*- et *susD-like* sont fréquemment retrouvées dans des

clusters en association avec des CAZymes et des régulateurs de l'expression génétique. Ces clusters sont connus sous le nom de *Polysaccharide Utilization Locus* (PUL). L'invention de protéines *SusD-like* et le développement de tels PUL pourraient en partie expliquer la spécialisation du phylum des *Bacteroidetes* dans la dégradation de polysaccharides. De plus, les génomes des *Bacteroidetes* sont très plastiques et fréquemment réorganisés par des phénomènes comme des duplications ou des transferts latéraux de gènes. Dans la dernière partie de cette revue, des liens potentiels entre symbiontes intestinaux et espèces environnementales sont discutés. Il apparaît que les *Bacteroides* intestinaux peuvent incorporer de nouveaux gènes provenant de bactéries de l'environnement ingérées en même temps que la nourriture. Ces transferts latéraux expliquent l'acquisition de nouvelles capacités cataboliques par le microbiome. Par ailleurs, la question de l'origine et de l'évolution des *Bacteroidetes* en tant que symbiontes des animaux est posée.



# Environmental and gut *Bacteroidetes*: the food connection

François Thomas<sup>1,2</sup>, Jan-Hendrik Hehemann<sup>1,2†</sup>, Etienne Rebuffet<sup>1,2†</sup>, Mirjam Czjzek<sup>1,2</sup> and Gurvan Michel<sup>1,2\*</sup>

<sup>1</sup> UMR 7139, Marine Plants and Biomolecules, Station Biologique de Roscoff, UPMC University Paris 6, Roscoff, France

<sup>2</sup> UMR 7139, CNRS, Marine Plants and Biomolecules, Station Biologique de Roscoff, Roscoff, France

## Edited by:

Peter J. Turnbaugh, Harvard University, USA

## Reviewed by:

Deborah Threadgill, North Carolina State University, USA  
Alain Stintzi, Ottawa Institute of Systems Biology, Canada

## \*Correspondence:

Gurvan Michel, UMR 7139, CNRS/UPMC, Marine Plants and Biomolecules, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France.  
e-mail: gurvan@sb-roscoff.fr

## †Current address:

Jan-Hendrik Hehemann, Department of Biochemistry and Microbiology, University of Victoria, PO Box 3055 STN CSC, Victoria, BC, Canada V8W 3P6;  
Etienne Rebuffet, Department of Chemistry, Biochemistry and Biophysics, University of Gothenburg, Box 462, SE-405 30 Göteborg, Sweden.

Members of the diverse bacterial phylum *Bacteroidetes* have colonized virtually all types of habitats on Earth. They are among the major members of the microbiota of animals, especially in the gastrointestinal tract, can act as pathogens and are frequently found in soils, oceans and freshwater. In these contrasting ecological niches, *Bacteroidetes* are increasingly regarded as specialists for the degradation of high molecular weight organic matter, i.e., proteins and carbohydrates. This review presents the current knowledge on the role and mechanisms of polysaccharide degradation by *Bacteroidetes* in their respective habitats. The recent sequencing of *Bacteroidetes* genomes confirms the presence of numerous carbohydrate-active enzymes covering a large spectrum of substrates from plant, algal, and animal origin. Comparative genomics reveal specific Polysaccharide Utilization Loci shared between distantly related members of the phylum, either in environmental or gut-associated species. Moreover, *Bacteroidetes* genomes appear to be highly plastic and frequently reorganized through genetic rearrangements, gene duplications and lateral gene transfers (LGT), a feature that could have driven their adaptation to distinct ecological niches. Evidence is accumulating that the nature of the diet shapes the composition of the intestinal microbiota. We address the potential links between gut and environmental bacteria through food consumption. LGT can provide gut bacteria with original sets of utensils to degrade otherwise refractory substrates found in the diet. A more complete understanding of the genetic gateways between food-associated environmental species and intestinal microbial communities sheds new light on the origin and evolution of *Bacteroidetes* as animals' symbionts. It also raises the question as to how the consumption of increasingly hygienic and processed food deprives our microbiota from useful environmental genes and possibly affects our health.

**Keywords:** *Bacteroidetes*, adaptation to environmental niches, microbiota

## THE PHYLUM BACTEROIDETES

The phylum *Bacteroidetes* is a very diverse bacterial phylum, the name of which changed several times over the past years. It is also known as the *Cytophaga–Flexibacter–Bacteroides* (CFB) group, an appellation that reflects the diversity of organisms found in this phylogenetic group (Woese, 1987; Woese et al., 1990). According to the Bergey's Manual of Systematic Bacteriology (Bergey's, 2011), the *Bacteroidetes* phylum comprises four classes: *Bacteroidia*, *Flavobacteria*, *Sphingobacteria*, and *Cytophagia*, representing around 7000 different species (NCBI, October 2010). The largest class is the *Flavobacteria*, grouping together around four times more species than the three others (Table 1). These bacteria are all Gram negative, cover a mixture of physiological types, from strictly anaerobic *Bacteroides* to strictly aerobic *Flavobacteria*. They are non-motile, flagellated, or move by gliding.

Members of the phylum *Bacteroidetes* have colonized many different ecological niches, including soil, ocean, freshwater, and the gastrointestinal tract (GIT) of animals, where they display various biological functions. In particular, they are well known degraders of polymeric organic matter. This review describes current knowledge on the role and mechanisms of polysaccharide degradation by *Bacteroidetes* in their respective habitats. We emphasize the features shared by members of the phylum that allow this

functional specialization in various environments. We address the links between these different microbial communities through food consumption, which raise the question of the evolution of gut microbes.

## BACTEROIDETES IN THE NORMAL MICROBIOTA OF ANIMALS

Microbes that live in and on humans (known as microbiota) can represent up to 100 trillion cells, 10 times more than the eukaryotic stem and somatic cells (Ley et al., 2006b; Turnbaugh et al., 2007). The large majority of these microbes reside in our GIT, and belong either to the *Firmicutes* or *Bacteroidetes* phyla (Marchesi, 2010). In fact, these two bacterial phyla account for >98% of the 16S rRNA sequences detected in the gut microbiota of mammals (Ley et al., 2006b). *Bacteroidetes* have colonized all the different parts of the GIT, despite the different conditions they have to face in terms of, e.g., pH, nutrients, and oxygen availability. Due to their easier accessibility, the mouth and colon microbiota have been the most studied. The large intestine is the most colonized compartment of the GIT, with bacterial densities reaching  $10^{11}$ – $10^{12}$  cells/ml (Whitman et al., 1998). Members of the *Bacteroidetes* phylum are well known colonizers of the colon. They account for about 50% of the 16S rRNA sequences detected from healthy human mucosal tissues (Eckburg et al., 2005). Among this phylum, members of the genus *Bacteroides*

**Table 1 | Census of the phylogenetic divisions inside the *Bacteroidetes* phylum (NCBI, October 2010).**

Class	No. of families	No. of genera	No. of species
<i>Bacteroidia</i>	5	28	858
<i>Flavobacteria</i>	3	110	3583
<i>Sphingobacteria</i>	3	29	787
<i>Cytophagia</i>	3	47	765
Unclassified strains	–	–	996

are the most abundantly represented in the fecal microbiota (Moore and Holdeman, 1974; Sghir et al., 2000). *Bacteroidetes* have also been found in the normal microbiota of the oral cavity, either in the saliva or dental plaque (Keijsers et al., 2008; Nasidze et al., 2009a,b). In the throat, Andersson et al. (2008) found that *Bacteroidetes* account for ~20% of the reads using barcoding pyrosequencing, and identified *Prevotella* genus as the second most represented after *Streptococcus*. The same proportions were found in the distal esophagus (Pei et al., 2004). Due to its low pH, the stomach is a harsh ecological niche, quite different from the other GIT compartments. However, even if *Helicobacter pylori* represents a major part of the stomach microbiota, *Bacteroidetes* were still found to account for 10–20% of the bacteria (Bik et al., 2006), including members of the genera *Prevotella*, *Capnocytophaga*, *Bergeyella*, *Porphyromonas*, and *Tannerella*. In fact, *Prevotella* was the third most abundant genus in the stomach microbiota, after *Helicobacter* and *Streptococcus*.

*Bacteroidetes* are also encountered in the microbiota of other mammals, such as mice (Dubos et al., 1965; Savage et al., 1968), dogs (Middelbos et al., 2010), pigs (Leser et al., 2002), and ruminants (Tajima et al., 1999; Leng et al., 2010). They appear in the GIT of domesticated and wild birds, such as chickens (Zhu et al., 2002), turkeys (Scupham et al., 2008), goose (Lu et al., 2009), and ostriches (Matsui et al., 2010). They have been demonstrated as dominant in the microbiota of echinoderms (Balakirev et al., 2008), millipedes (Knapp et al., 2010), and the last compartment of termite gut (Schmitt-Wagner et al., 2003).

The interaction between *Bacteroidetes* and their animal host is now known to be mutualism rather than commensalism since the fitness of both partners is increased (Backhed et al., 2005). Several studies have shown the implication of *Bacteroidetes* for the normal development of the GIT. For example, it has been shown that germ-free animals display an altered GIT compared to conventional ones, leading to changes in morphology and function (see for review, Abrams, 1983). Other contributions of *Bacteroidetes* to the health of their host include interactions with the immune system for the activation of T-cell mediated responses (Mazmanian et al., 2008; Wen et al., 2008), and limitation of the GIT colonization by potential pathogenic bacteria (Mazmanian, 2008). Gut *Bacteroidetes* generally produce butyrate, an end product of colonic fermentation which is thought to have antineoplastic properties and thus plays a role in maintaining a healthy gut (Kim and Milner, 2007). They are also involved in bile acid metabolism and transformation of toxic and/or mutagenic compounds (Smith et al., 2006).

However, their main and most studied biological function as symbionts is the degradation of biopolymers in the large intestine, and in particular polysaccharides. Carbohydrates represent

the bulk of normal human and animal diets, thus forming a main source of nutrients for both the host and the microbiota. Mammals can easily absorb simple sugars (e.g., glucose, galactose) via active transport in the proximal small intestine (Ferraris, 2001) or degrade some disaccharides (e.g., lactose, maltose, sucrose) to monomers. They also possess enzymes which allow the degradation of starch to glucose. However, they are generally under-equipped for the degradation of other complex polysaccharides, which are resistant to the action of digestive enzymes and reach the colon relatively intact. The microbiota and especially *Bacteroidetes* are believed to complement eukaryotic genomes with degradation enzymes targeting resistant dietary polymers, such as plant cell wall compounds (e.g., cellulose, pectin, and xylan). *Bacteroidetes* also degrade host-derived carbohydrates, primarily coming from GIT secretions, such as *N*-glycans found in mucins or chondroitin sulfates (Salysers et al., 1977). With this respect, one well studied example is *Bacteroides thetaiotaomicron*, a prominent member of the human microbiota that has been shown to principally forage on host mucin *O*-glycans. This has an impact on their colonization and probably represents an important evolved component of their adaptation to the gut habitat (Martens et al., 2008). The bacteria-mediated fermentation of these food, and host-derived polysaccharides in the colon leads to the release of volatile, short-chain fatty acids (mainly acetate, propionate, and butyrate) that are reabsorbed by the host. Therefore these intestinal bacteria help the host to gain energy from otherwise refractory carbohydrate sources. In herbivorous animals such as ruminants, the diet consists largely of plant cell wall compounds resistant to the host's digestive enzymes. Thus, assimilation of short-chain fatty acids produced by microbial fermentation of polysaccharides can support more than 50% of the total caloric supply (Carroll and Hungate, 1954). In omnivorous mammals, notably humans, this extra-source of energy accounts for 7–10% of the daily allowance (Smith and Bryant, 1979; Hooper et al., 2002). In rats, it has been shown that germ-free animals excrete 87% more calories within the feces than their normal counterparts, and need to ingest 30% more food to maintain their body weight (Wostmann et al., 1983). Hence, the presence of the intestinal microbiota is required for the optimal uptake of energy from the diet.

#### ENVIRONMENTAL BACTEROIDETES

Members of the phylum *Bacteroidetes* have colonized virtually all types of environments encountered on Earth. This versatility is reflected by the diversity of sources from which cultivated strains have been isolated, such as soil, activated sludge, decaying plant material, compost, freshwater, and marine samples, algae, dairy products, and diseased animals (Bernardet and Nakagawa, 2006; Reichenbach, 2006). While the GIT microbiota is mainly composed of species from the *Bacteroidia* class, environmental *Bacteroidetes* belong primarily to the *Flavobacteria*, *Cytophagia*, and *Sphingobacteria* classes.

Over the past years, the diversity of environmental bacterial communities has been re-evaluated due to the use of culture-independent techniques. *Bacteroidetes* are increasingly recognized as an important compartment of the bacterioplankton in marine environments (Jooste and Hugo, 1999; Kirchman, 2002), especially in pelagic oceans. They are highly abundant as shown by fluorescent *in situ* hybridization (FISH) experiments using specific probes

on seawater samples. They account for as much as half of the identified cells potentially identified and outnumber alternative species (Glöckner et al., 1999; Simon et al., 1999; Cottrell and Kirchman, 2000). Using 16S rRNA gene clone libraries from nine different sampling sites (including coastal and oceanic waters at 5 m-depth in temperate, tropical, and polar regions), Pommier et al. (2007) found that *Bacteroidetes* and *Proteobacteria* were dominant. Marine *Bacteroidetes* are also often demonstrated as living on marine inert or living surfaces, either in sediments (Llobet-Brossa et al., 1998; Julies et al., 2010), in biofilms (Edwards et al., 2010), in hydrothermal vents (Sievert et al., 2000), associated with corals (Frias-Lopez et al., 2002; Rohwer et al., 2002) or on the surface of macroalgae (Beleneva and Zhukova, 2006; Staufenberg et al., 2008; Salaun et al., 2010) and angiosperms (Crump and Koch, 2008).

*Bacteroidetes* also colonize freshwater environments to a significant extent. Using FISH, they were found in all the water samples collected in lakes from different locations (Austria, Switzerland, Germany, and Siberia), with a median abundance ranging from 2 to 12% which tends to increase with depth (Glöckner et al., 1999). *Bacteroidetes* accounted for 40–60% of the detectable bacteria in river samples from Spain (Simek et al., 2001). An integrative analysis of the available bacterial 16S rDNA sequences recovered from freshwater lakes and rivers over the past 15 years concluded that these environments host a specific planktonic bacterial community, different from the adjacent terrestrial habitats and sediments (Zwart et al., 2002). *Bacteroidetes* were amongst the major bacterial lineages in this community, together with *Proteobacteria*, *Actinobacteria*, and *Verrucomicrobia*.

Soil-associated bacterial communities also comprise members of the *Bacteroidetes* phylum. Using different techniques they have been detected in soil samples from various locations, including cultivated fields (Borneman et al., 1996; Martinez-Alonso et al., 2010), greenhouse soils (Kim et al., 2006) and unexploited areas (Buckley and Schmidt, 2003; Zhou et al., 2009). Recently, a study including soils from 88 different places revealed a positive correlation between the pH of the substrate and the relative abundance of *Bacteroidetes* (Lauber et al., 2009), ranging from 1.7% at low pH (<4) to 17% in basic soils (pH > 8). The surface of leaves, or phyllosphere, is another location of *Bacteroidetes* colonization in terrestrial environments. In a recent study, they accounted for 21.5% of the OTUs sequenced in samples from 56 tree species. Indeed, the most common group of bacteria found on the leaves was *Sphingobacteriales* which represented 21.3% of all sequences (Redford et al., 2010). Interestingly, *Bacteroidetes* tend to be more abundant on conifers (gymnosperms), reaching 70% of the sequences detected on the leaves, and less abundant on leaves from angiosperms.

Environmental *Bacteroidetes* are thought to be specialized in the degradation of complex organic matter in the biosphere, especially in the form of polysaccharides and proteins (Church, 2008). As a group, they are very versatile in the range of biopolymers they can use as carbon and energy source, e.g., plant, algal, or animal compounds. The description of new taxa often results from screening of environmental samples to discover original enzymatic activities with potential biotechnological applications (Barbeyron et al., 2001, 2008; Pankratov et al., 2006; Lee et al., 2010).

Several lines of evidence designate *Bacteroidetes* as crucial degraders of complex organic matter in the environment and not only in laboratory, many of them resulting from research on aquatic ecosystems (Kirchman, 2002). In the ocean, particulate and dissolved organic matter tends to aggregate to form particulate detritus (>0.5 mm) known as marine snow (Chin et al., 1998). These sinking particles are hotspots of organic carbon and thus play a key role in the export of matter from the euphotic surface to the sediment of the deep ocean (Azam and Malfatti, 2007). Marine snow hosts a considerable bacterial population, with densities reaching  $10^6$ – $10^8$  cells per aggregate (Alldredge et al., 1986). Numerous studies characterizing these particles-associated populations have revealed that *Bacteroidetes* account for a large part (Rath et al., 1998; Ploug et al., 1999) and can even be the most abundant phylogenetic type detected (DeLong et al., 1993). Similar results have been found for estuarine particles, where *Cytophaga* spp. are amongst the dominant species in the bacterial assemblage (Crump et al., 1999). This enrichment of *Bacteroidetes* can be linked to their ability (i) to efficiently degrade complex and recalcitrant biopolymers sequestered in the particulate organic detritus and (ii) to colonize surfaces (Kirchman, 2002). Similarly, the immersion of insoluble cellulose for one month in the Irish Sea resulted in the colonization of the fibers by a specific bacterial biofilm, dominated by *Bacteroidetes* and *Gammaproteobacteria* (Edwards et al., 2010). In lakes, Glöckner et al. (1999) found that *Bacteroidetes* relative abundance in the water column increased with the depth of sampling. They proposed that this was due to the accumulation of slowly degradable, sinking macromolecules, a substrate that may have selected efficient degraders like *Bacteroidetes*.

Another evidence for consumption of high molecular weight (HMW) compounds by *Bacteroidetes* comes from several experiments where the bacterial community was followed after an increase of input of organic matter. Coastal and deep-sea environments typically receive seasonal pulses of organic matter after phytoplankton bloom events (Goody, 2002; Kim et al., 2009). During such blooms of diatoms (Riemann et al., 2000) and dinoflagellates (Fandino et al., 2001), the population of *Bacteroidetes* increased significantly in a short time scale, concomitant with increasing activity of extracellular hydrolytic enzymes. Concordant with these results, a recent study of the communities of water masses from the North Atlantic Ocean revealed a specific enrichment of *Flavobacteria* in the space directly surrounding nanophytoplankton (Gomez-Pereira et al., 2010). In a freshwater lake, the virus-induced lysis of filamentous cyanobacteria led to the appearance of newly emerging bacterial species (van Hannen et al., 1999). The majority of these new microbial populations belonged to *Cytophagales* and *Actinomycetes*.

In alternative experiments, seasonal pulses were mimicked in laboratory conditions by an artificial increase of the organic carbon input. Gihring et al. (2009) simulated a phytodetritus deposition event by adding heat-killed *Spirulina* cyanobacteria on coastal sediments, resulting in an increase in the proportion of *Bacteroidetes* that was limited to the surface of the sediment (from 0 to 4 mm depth). This suggests a role of *Bacteroidetes* in the phytodetritus degradation in aerobic conditions (Gihring et al., 2009). The addition of exogenous protein to Californian coastal seawater mesocosms also triggered a change in the bacterial community, with an increase of the occurrence of several



*Bacteroidetes* species (Pinhassi et al., 1999). Finally, by combining microautoradiography and FISH (MICRO-FISH), Cottrell and Kirchman (2000) compared the uptake of biopolymers (protein and chitin) and their constituent monomers (amino acids and *N*-acetylglucosamine, respectively) by natural bacterial assemblages from coastal waters of Delaware. This approach showed an over-representation of *Bacteroidetes* in the portion of bacteria consuming polymers, whereas Alphaproteobacteria dominated uptake of monomers (Cottrell and Kirchman, 2000).

Taken together, these results show that environmental *Bacteroidetes* specialize in the mineralization of HMW organic matter. Thus, this phylum represents a key compartment for carbon fluxes and budgets in ecosystems (Nagata, 2008).

### BACTEROIDETES AS PATHOGENS

As outlined above, *Bacteroidetes* are found in the normal bacterial community of many diverse habitats, where they play a beneficial role in the degradation of organic matter. On the contrary, some members of this phylum can have a strong pathogenic behavior toward different eukaryotic species. This is underlined by the number of *Bacteroidetes* strains isolated from clinical samples of human and animal origin (Hugo et al., 1999), including blood, urine, infected wounds, and feces. Some members of the *Bacteroides* genus, although belonging to the normal GIT microbiota, can cause opportunistic infections if the integrity of the intestinal mucosal barrier is broken (Smith et al., 2006). Indeed, the majority of species isolated from anaerobic infections falls into the *Bacteroides* genus, acting on different discrete sites such as peritoneal cavity, vagina, sinuses, skin, and heart (Wexler, 2007). These infections are usually polymicrobial, but *B. fragilis* and *B. thetaiotaomicron* are the most frequent species. *Bacteroides* are also associated with bacteremia, and *B. fragilis* is the most common blood isolate recovered from patients (Brook, 2010). Members of the genera *Porphyromonas*, *Prevotella*, and *Tannerella* are well known pathogens of the oral cavity, where they can notably cause periodontal disease and caries (Tanner et al., 1986; Shah and Collins, 1990; Gibson and Attardo Genco, 2006). The emergence of an infection seems to be linked to the assemblage of pathogens in bacterial consortia more than on the individual action of specific species (Jenkinson and Lamont, 2005).

The *Flavobacteria* class also contains opportunistic human pathogens, invading hosts with poor immune system (Bernardet and Nakagawa, 2006). One of them, *Elizabethkingia meningoseptica*, causes meningitis in infants (King, 1959; Ratner, 1984; Kim et al., 2005). *Myroides* has been reported in cases of surgery wound, urinary tract infections, septicemia, pneumonia, meningitis, fasciitis, ventriculitis, and nosocomial infections (Mammeri et al., 2002). The genera *Empedobacter*, *Bergeyella*, *Weeksella*, and *Capnocytophaga* also contain pathogenic species (Hugo et al., 2006a,b; Leadbetter, 2006).

*Bacteroidetes* infections can have devastating effects for farmed and wild fish (Austin and Austin, 2007), as reported a century ago and first observed in aquaria (Davis, 1922). *Flavobacterium psychrophilum*, *F. columnare*, *F. branchiophilum*, *Tenacibaculum maritimum*, and *Chryseobacterium scophthalmum* are some of the species with the most severe economic impact (Bernardet and Bowman, 2006).

Members of the genera *Riemerella*, *Ornithobacterium*, and *Coenonia* can cause septicemia and respiratory tract infections in birds (Segers et al., 1993; Vandamme et al., 1994, 1999).

In addition, *Bacteroidetes* can affect various plants and algae. *F. johnsoniae* is incriminated as an opportunistic pathogen in the decay of fresh plants and vegetables called “soft rot” (Liao and Wells, 1986). Members of the genera *Zobellia*, *Cellulophaga*, and *Kordia* have algicidal activities (Skerratt et al., 2002; Sohn et al., 2004), and can therefore control blooms of microscopic algae. The Anaaki disease, severely damaging the red macroalga *Porphyra yezeensis*, is caused by *Flavobacterium* sp. LAD-1 (Sunairi et al., 1995). Uyenco (1977) isolated strains of *Flavobacterium* associated with “ice-ice disease” in decaying red alga *Euclima striatum*, characterized by whitening and hardening of the seaweed’s tissues. Similarly, a *Cytophaga* sp. promotes the ice-ice disease in stressed branches of the farmed Rhodophytes *Kappaphycus alvarezii* and *Euclima denticulatum* (Largo et al., 1995). Members of the *Flavobacterium* genus can also infect brown macroalga such as *Undaria pinnatifida*, where they are associated with “spot-rotting” disease and “shot hole disease” (Neill et al., 2008).

The pathogenicity of *Bacteroidetes* is at least partly due to the production of polymer-degrading enzymes targeting host cellular components. Animal pathogens produce a range of hydrolytic enzymes such as hyaluronidase, chondroitin sulfatase, heparinase, glucosidases, and *N*-acetylglucosaminidase (Rudek and Haque, 1976; Duerden, 1994). These enzymes damage the extracellular matrix of animals and thus are important both for tissue colonization and nutrients supply. In a study of the fish pathogen *F. columnare*, a chondroitin lyase was found to be related to the virulence (Suomalainen et al., 2006). Plant and algal pathogens are able to degrade cell wall compounds such as pectins and agars. Indeed, phytopathogenic strains of *F. johnsoniae* secrete much more pectate lyase than their non pathogenic counterparts (Liao and Wells, 1986). *Bacteroidetes* also secrete various proteases, and this has been documented as an important virulence factor. Proteases secreted by different pathogenic *Bacteroides* species can destroy human brush border digestive enzymes (Riepe et al., 1980). Greiner and Mayrand (1987) found that the pathogenic strains of *Porphyromonas gingivalis* had a higher collagenolytic activity than the non-virulent strains. In animal models, *P. gingivalis* mutants lacking the ability to produce various proteases lost their virulence (Loesche, 1993; Fletcher et al., 1995). Whatever their environment, it appears that the biological function of *Bacteroidetes* either as symbiont, environmental, or pathogen bacteria is connected with their propensity to degrade complex biopolymers, i.e., polysaccharides and proteins.

### HABITAT EVOLVED ADAPTATION MEASURED BY THE VARIETY OF POLYSACCHARIDES THAT ARE DEGRADED

As a phylum, and especially due to their versatility in habitats, *Bacteroidetes* have access to an amazing diversity of carbon sources. Indeed, the chemical diversity of polysaccharides largely outnumbers the possibility for protein folds – it has been calculated that there are  $1.05 \times 10^{12}$  possible linear and branched forms of a single hexasaccharide (Laine, 1994). Moreover, these structural variations have been harnessed by living organisms to fulfill very different roles: e.g., structural, storage, specific signaling, specific recognition, host–pathogen interactions to name but a few

(Carpita and Gibeaut, 1993; Graham et al., 2000; Stahl and Bishop, 2000). Consequently, carbohydrates account for around 75% of the biomass on Earth, a natural resource that was not lost on competing organisms that were developing their own strategies to utilize this chemical energy for their own survival.

But even more importantly several classes of polysaccharides are niche specific. While cell walls are a characteristic feature of all plants, they are not exclusive to plants, with most bacterial and algal cells as well as all fungal cells also being surrounded by extracellular, macromolecular barriers (extracellular matrix or ECM). The macromolecular composition, however, is characteristically different among the major evolutionary lineages of the living world, linking specific life style or nutritional habits to specifically encountered biopolymers. A vivid example is provided by the polysaccharides of the marine environment that are typically and to a large majority sulfated (carrageenans and fucans) or highly ionic (alginates) and unique to this particular habitat (Michel et al., 2010a,b; Popper et al., 2011). In contrast, the basic polysaccharide components of plant cell walls are cellulose and hemicellulose (pectins, xylans, mannans, xyloglucans, etc.), whereas fungal cell walls primarily consist of chitin (Niklas, 2004). In metazoa, the ECM will predominantly consist of chondroitin or dermatan, which are essentially made of sulfated polysaccharides referred to as glycosaminoglycans (GAGs; Sugahara and Kitagawa, 2002) that are interconnected by fibrillar proteins (collagens). Other sources of carbohydrates in animals are glycosylation sites, such as mucin that contain a high proportion of sialic acid in addition to GAGs (Raman et al., 2005).

Reflecting this chemical diversity of the substrate, glycosidases, the enzymes responsible for the breakdown of di-, oligo-, and polysaccharides, as well as glycoconjugates, are ubiquitous through all domains of life (Turnbaugh et al., 2010). Carbohydrate processing enzymes (CAZymes), including glycosidases and glycosyltransferases (the enzymes which transfer saccharides to other saccharide moieties, small molecules, lipids, or proteins), constitute between 1 and 3% of the genome of most organisms (Davies et al., 2005). Noteworthy, the genomes of *Bacteroidetes* species have revealed that they are champions with respect to the diversity and number of CAZymes they contain, reflecting the molecular strategies evolved by this microbial community to differentiate, capture, and degrade complex glycans. Consequently and as a result of this ability to degrade host and plant glycans, cultured (environmental or gut) species are often used to isolate specific enzymes for polysaccharide degradation (Berg et al., 1980; Tierny et al., 1994; Bernardet and Nakagawa, 2006; Reichenbach, 2006).

## THE GENOMIC PERSPECTIVE

The first sequenced genome of a *Bacteroidetes* representative was published in 2002 for the human symbiont *B. thetaiotaomicron* (Xu et al., 2003). Since then, many sequencing projects have been conducted to increase the genomic knowledge on this phylum. To date, 33 *Bacteroidetes* genomes are complete, publicly available and published (Table 2). Many others are in a draft state or have not yet been published (total of 125 sequences censused on NCBI). During the last 10 years, sequencing efforts have indiscriminately concerned environmental, pathogen, and symbiotic/comensal species with the aim to better understand their biological functions, including their capacity to interact with their habitats.

This data allows comparing the different enzymatic capabilities of various genera, which sheds new light on the specialization of *Bacteroidetes* toward degradation of organic matter. A striking common feature revealed by this comparative genomic approach is the trend of *Bacteroidetes* genomes to encode many polymer-degrading enzymes, acting either on proteins or carbohydrates. The census of Carbohydrate-Active enzymes (CAZ Ymes) in the CAZY database<sup>1</sup> (Cantarel et al., 2009) eases the comparison of the number of glycosylhydrolases (GH) and polysaccharide lyases (PL) in an increasing number of sequenced species. In each of the four classes of the *Bacteroidetes* phylum, there are examples of CAZYme-enriched species. In *Bacteroidia*, the proteomes of *B. thetaiotaomicron*, *B. fragilis*, and *P. ruminicola* comprise 272, 137, and 130 GH and PL respectively (Xu et al., 2003; Kuwahara et al., 2004), much more than other members of the gut microbiota, or outside the *Bacteroidetes* phylum, such as *Clostridium perfringens* (57 GH and PL) and *Bifidobacterium longum* (49 GH). The same is true for *Flavobacteria* [e.g., *G. forsetii* with 48 GH/PL (Bauer et al., 2006), *Z. profunda* with 120 GH (Qin et al., 2010), *F. johnsoniae* with 152 GH/PL (McBride et al., 2009)], *Cytophagia* [e.g., *S. linguale* with 151 GH/PL (Lail et al., 2010), *D. fermentans* with 98 GH/PL (Lang et al., 2009), and *Sphingobacteria*, e.g., *C. pinensis* with 184 and *P. heparinus* with 163 GH/PL, respectively (Han et al., 2009; Del Rio et al., 2010)]. In some cases, the prediction of CAZYme-encoding genes in newly sequenced organisms can even unveil unexpected catabolic capabilities toward specific substrates (McBride et al., 2009). Moreover, the predicted enzymatic battery of a bacterial species will help characterize its natural habitat (i.e., the available substrates) and its ecological function in organic matter recycling. Recently, this feature has been used for genome and habitat comparison, linking the number and occurrence of specific CAZYme-families to the environmental niche (Pope et al., 2010; Purushe et al., 2010). With this respect, the presence of a vast majority of exo-acting enzymes in the genome of *B. thetaiotaomicron*, suggests that the organism is able to use the saccharide decorations appended to the backbone of structural polysaccharides and glycoproteins (Xu et al., 2003). Another recent study elegantly demonstrates that this particularly evident expansion in exo-GHs of family GH92 enzymes (23 members) is indeed related to the  $\alpha$ -mannosides present in the N-glycans of host and dietary glycoproteins (Zhu et al., 2010).

However, one has to keep in mind that a rough analysis of the number of degradation enzymes in general is insufficient; one needs to go down to the enzymatic sub-family to infer a putative metabolism. Indeed, the CAZYme classification based on sequence similarity<sup>1</sup> has the consequence that gene families group together enzymes with widely different substrate or product specificities (Henrissat, 1991). Therefore, to derive knowledge useful for subsequent functional predictions, phylogenetic analyses defining subgroups that contain biochemically characterized representatives are needed to perform unambiguous assignments (Turnbaugh et al., 2010).

In spite of the general trend of *Bacteroidetes* to possess numerous degradation enzymes, there are several noteworthy exceptions. The fish pathogen *F. psychrophilum* genome harbors only 13 proteases, i.e., 4.5 proteases per megabase (Mb) and 3 GH

<sup>1</sup><http://www.cazy.org>



**Table 2 | List of *Bacteroidetes* with a fully sequenced and published genome, and their respective habitats.**

Species	Genome size (CDS number)	Habitat (functional role)	Reference
<b>FLAVOBACTERIA CLASS</b>			
<i>Capnocytophaga ochracea</i> DSM 7271	2.6 Mb (2193)	Animals and humans (S/P)	Mavrommatis et al. (2009)
<i>Gramella forsetii</i> KT0803	3.8 Mb (3585)	North Sea surface water during phytoplankton bloom (E)	Bauer et al. (2006)
<i>Robiginitalea biformata</i> HTCC2501	3.5 Mb (3211)	Sargasso Sea (E)	Oh et al. (2009)
<i>Flavobacterium johnsoniae</i> UW101	6.1 Mb (5056)	Soil and freshwater (E)	McBride et al. (2009)
<i>Flavobacterium psychrophilum</i> JIP02/86	2.9 Mb (2432)	Fish (P)	Duchaud et al. (2007)
<i>Croceibacter atlanticus</i> HTCC2559	3.0 Mb (2715)	Sargasso Sea (E)	Oh et al. (2010)
<i>Zunongwangia profunda</i> SM-A87	5.1 Mb (4653)	Deep-sea sediments (E)	Qin et al. (2010)
<i>Candidatus Sulcia muelleri</i> GWSS	0.3 Mb (228)	Sap-feeding insects flora (S)	McCutcheon and Moran (2007)
<i>Maribacter</i> sp. HTCC2170	3.9 Mb (3411)	Surface sea water, Oregon coast (E)	Oh et al. (2011)
<i>Periplaneta americana</i>	0.6 Mb (581)	Cockroach endosymbiont (S)	Sabree et al. (2009)
<i>Blattella germanica</i>	0.6 Mb (586)	Cockroach endosymbiont (S)	Lopez-Sanchez et al. (2009)
<i>Polaribacter dokdonensis</i> MED152	3.0 Mb (2646)	Northwestern Mediterranean sea surface water (E)	Gonzalez et al. (2008)
<i>Dokdonia donghaensis</i> MED134	3.3 Mb (2284)	Korean sea water (E)	Gomez-Consarnau et al. (2007)
<i>Leeuwenhoekella blandensis</i> MED217	4.2 Mb (3735)	Mediterranean sea water (E)	Gomez-Consarnau et al. (2007)
<b>SPHINGOBACTERIA CLASS</b>			
<i>Chitinophaga pinensis</i> DSM 2588	9.1 Mb (7302)	Pine litter in Australia (E)	Del Rio et al. (2010)
<i>Pedobacter heparinus</i> DSM 2366	5.2 Mb (4287)	Soil (E)	Han et al. (2009)
<i>Salinibacter ruber</i> M8	3.6 Mb (3086)	Saltern crystallizer pond, Mallorca (E)	Pena et al. (2010)
<i>Salinibacter ruber</i> M31	3.6 Mb (2934)	Saturated thalassic brines (E)	Mongodin et al. (2005)
<i>Rhodothermus marinus</i> R-10	3.4 Mb (2914)	Submarine hot spring, Iceland (E)	Nolan et al. (2009)
<b>CYTOPHAGIA CLASS</b>			
<i>Dyadobacter fermentans</i> NS114	7.0 Mb (5804)	Stems from <i>Zea mays</i> (E)	Lang et al. (2009)
<i>Cytophaga hutchinsonii</i> ATCC 33406	4.4 Mb (3790)	Soil (E)	Xie et al. (2007)
<i>Spirosoma linguale</i> 1	8.5 Mb (7069)	Soil, freshwater (E)	Lail et al. (2010)
<b>BACTEROIDIA CLASS</b>			
<i>Candidatus Azobacteroides pseudotricononymphae</i> CfPt1-2	1.1 Mb (758)	Endosymbiont of termite gut protist (S)	Hongoh et al. (2008)
<i>Porphyromonas gingivalis</i> ATCC 33277	2.4 Mb (2090)	Deep periodontal pockets (P)	Naito et al. (2008)
<i>Porphyromonas gingivalis</i> W83	2.3 Mb (1990)	Deep periodontal pockets (P)	Nelson et al. (2003)
<i>Parabacteroides distasonis</i> ATCC 8503	4.8 Mb (3867)	Human gastrointestinal tract (S)	Xu et al. (2007)
<i>Bacteroides vulgatus</i> ATCC 8482	5.2 Mb (4088)	Human gastrointestinal tract (S)	Xu et al. (2007)
<i>Bacteroides fragilis</i> NCTC 9343	5.2 Mb (4274)	Human gastrointestinal tract (S/P)	Cerdeno-Tarraga et al. (2005)
<i>Bacteroides fragilis</i> YCH46	5.3 Mb (4578)	Human gastrointestinal tract (S/P)	Kuwahara et al. (2004)
<i>Bacteroides thetaiotaomicron</i> VPI-5482	6.3 Mb (4779)	Human gastrointestinal tract (S)	Xu et al. (2003)
<i>Prevotella ruminicola</i> 23	3.6 Mb (2763)	Bovine rumen (S)	Purushe et al. (2010)
<i>Prevotella bryantii</i> B <sub>4</sub>	3.6 Mb (2780)	Bovine rumen (S)	Purushe et al. (2010)
<b>UNCLASSIFIED</b>			
<i>Candidatus Amoebophilus asiaticus</i> 5a2	1.9 Mb (1557)	Obligate intracellular ameba symbiont (S)	Schmitz-Esser et al. (2010)

The functional role is given in brackets: S, symbiont; P, pathogen; E, environmental.

per Mb (Duchaud et al., 2007). This is far less than its closely related soil-associated cousin *F. johnsoniae*, which possesses 20.5 proteases and 23 GH per Mb and has a genome more than twice as large. The reduced number of degradation enzymes in the fish pathogen can be explained by its dedication to the infection of animal tissues. For such a species, the relative restricted diversity of substrates (compared to a soil-associated species) would diminish the need of multiple families of hydrolases. A second, similar example is *P. gingivalis* that has a relatively lower number

of hydrolases (only 24 GHs) compared to other members of the *Bacteroidia* class (Nelson et al., 2003). Again, the most plausible explanation is the high speciation of this species to dental plaque degradation that results in the restricted diversity of substrates utilized by this bacterium.

Some marine representatives of the *Flavobacteria* class seem to alternate between two life strategies depending on the abundance of carbon sources. This is notably the case of *Polaribacter dokdonensis* (strain MED152), *Leeuwenhoekella blandensis* (strain

MED217<sup>T</sup>), and *Dokdonia donghaensis* (strain MED134). On the one hand, commonly with other *Bacteroidetes*, they are very well equipped to attach to surfaces and depolymerize organic matter. For example, *P. dokdonensis* and *D. donghaensis* genomes encode many enzymes to degrade proteins (93 and 120 peptidases, respectively) and polysaccharides (30 and 22 GH, respectively; Gonzalez et al., 2008; Kirchman, 2008; Woyke et al., 2009). On the other hand, when polymeric substrates become scarce, these species switch to a free-living lifestyle adapted to carbon-poor environments. Genome analysis showed the presence of proteorhodopsin, a light-dependent H<sup>+</sup> pump that can drive ATP synthesis (Beja et al., 2000). This protein allows the phototrophic production of sufficient energy to maintain the population growth when the concentration of organic carbon decreases (Gomez-Consarnau et al., 2007). Additionally, *P. dokdonensis* is enriched in enzymes involved in anaerobic reactions, and assimilates CO<sub>2</sub> faster in light conditions than in the dark (Gonzalez et al., 2008). Altogether, these results suggest that marine *Bacteroidetes* may cope as well with feast and famine, and complement the understanding of their role in carbon cycles (DeLong and Beja, 2010).

Beyond the prediction of numerous polysaccharide-degrading enzymes, two large paralogous families of proteins have been found in *Bacteroidetes* genomes, which likely participate in polysaccharide uptake. These include homologs of the outer membrane proteins SusC and SusD from *B. thetaiotaomicron*, which are involved in starch utilization. Seminal work of the Salyers' group showed that this human gut symbiont degrades starch via a dedicated starch utilization system (Sus) with several proteins acting in coordination to sense, bind, and hydrolyze the substrate (Anderson and Salyers, 1989a,b; Shipman et al., 2000). The genes encoding these proteins cluster on the bacterial chromosome into typical polysaccharide utilization loci (PUL).

### POLYSACCHARIDE UTILIZATION LOCUS: FROM THE ARCHETYPIC SUS SYSTEM TO NEW SUBSTRATES

In *B. thetaiotaomicron*, the Sus locus comprises eight genes, *sus-RABCDEFG* (Martens et al., 2009). *SusR* is an inner membrane regulatory protein which activates the transcription of the other genes in the presence of maltose or starch (D'Elia and Salyers, 1996). *SusC*, *SusD*, *SusE*, and *SusF* are outer membrane proteins involved in the binding of the polysaccharide (Shipman et al., 2000). In fact, analyses of mutant strains have shown that *SusC* and *SusD* account together for 70% of the starch-binding capabilities of the wild type (Reeves et al., 1997). Surface-bound starch is hydrolyzed by the outer membrane  $\alpha$ -amylase *SusG*, which acts endolytically and releases oligosaccharides larger than maltotriose (Shipman et al., 1999; Martens et al., 2009). These degradation products are then channeled to the periplasm through the TonB-dependent receptor,  $\beta$ -barrel-type *SusC*, where they are further cleaved by the neopullulanase *SusA* and  $\alpha$ -glucosidase *SusB*. The atomic structures of several protein members have been resolved, namely the starch-binding *SusD* (Koropatkin et al., 2008), the  $\alpha$ -glucosidase *SusB* (Kitamura et al., 2008), and the  $\alpha$ -amylase *SusG* (Koropatkin and Smith, 2010). The Sus locus is organized into two transcriptional units under the control of *SusR*, one containing *susA* and the other

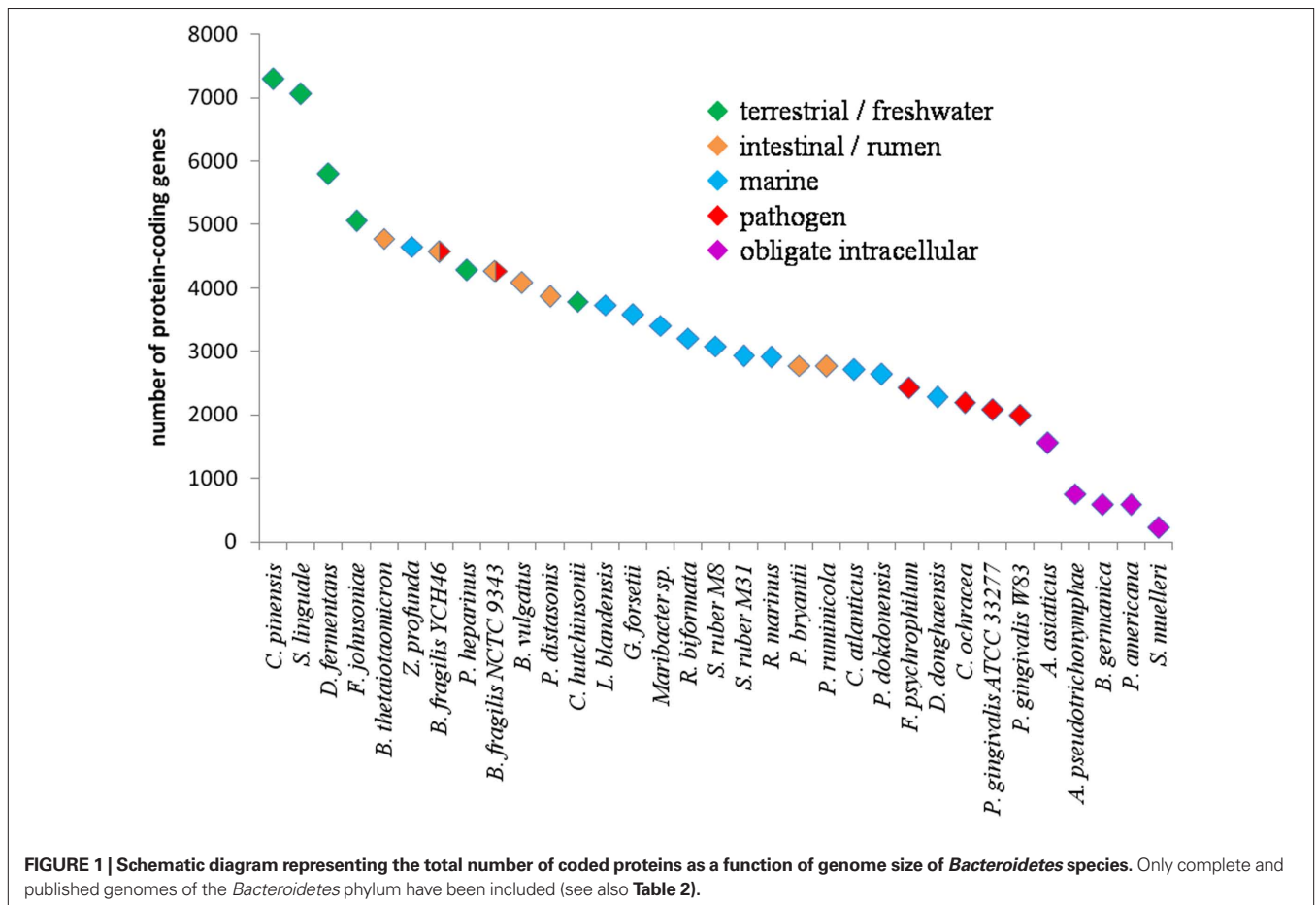
containing *susB* to *susG* (D'Elia and Salyers, 1996; Reeves et al., 1997). This allows a co-regulation of the PUL (Anderson and Salyers, 1989b).

The characterization of this first starch-specific PUL was followed by the discovery of numerous PULs in *Bacteroidetes*. *susC*-like and *susD*-like genes are strikingly frequent in *Bacteroidetes* genomes, often appearing in tandem and as the central units of substrate-specific PULs. *B. thetaiotaomicron* possesses 107 paralogs of *susC*, of which 101 are paired to a *susD*-like gene. 62 of these pairs are part of larger clusters, together with polysaccharide-degrading enzymes (Xu et al., 2003). In addition, some of these PULs comprise enzymes targeting glycan decorations, such as sulfatases or acetyl esterases. Thus, depending on the specificity of the predicted enzymes, one can infer the favorite substrate(s) of a given PUL. In total, PULs represent 18% of the genome of *B. thetaiotaomicron* (Martens et al., 2008). The closely related *B. fragilis*, *B. vulgatus*, and *Parabacteroides distasonis* also possess numerous PULs (Kuwahara et al., 2004; Xu et al., 2007). These PULs likely favor the success of *Bacteroides* spp. in the uptake of dietary and host-derived polysaccharides in the highly competitive gut habitat, and may explain their evolution as symbionts.

Interestingly, environmental species also harbor plenty of specific PULs. The annotation of the *F. johnsoniae* genome revealed 42 pairs of *susC*-like and *susD*-like genes, among which many were associated with CAZymes (McBride et al., 2009). The authors were notably able to predict PULs targeting starch (homologous the *B. thetaiotaomicron* Sus locus), chitin, and hemicelluloses. The genome of the marine *Flavobacteria G. forsetii* encodes 40 paralogs of *SusC*, and 14 clusters of *susCD*-like genes were detected (Bauer et al., 2006), often in the vicinity of CAZyme genes. This suggests that environmental *Bacteroidetes* as well as their gut-associated cousins use a unique and similar strategy to bind and degrade polymeric organic matter. Indeed, *SusD* homologs are only found in *Bacteroidetes* representatives. Thus, we speculate that the appearance of PULs including a *susD*-like gene in the ancestral *Bacteroidetes* could have at least partly driven the emergence of the phylum, and allowed its evolution as a group specialized in carbohydrate degradation.

### THE HIGH PLASTICITY OF BACTEROIDETES GENOMES

The size of *Bacteroidetes* genomes varies considerably between species (Table 2). Among published sequencing projects, *Chitinophaga pinensis* has the largest genome (9.1 Mb) whereas *Candidatus Sulcia muelleri* has the smallest (0.3 Mb). This great discrepancy can be at least partly associated with the different ecological niches colonized and the biological functions played by *Bacteroidetes* (Figure 1). Obligate intracellular symbionts have a more reduced genome size, due to their peculiar lifestyle. These species have evolved through successive inactivation and loss of genes, affecting virtually every cellular process. The possible causes of this genome reduction are multiple, including the unusual stability and metabolic richness of the cytoplasmic compartment they inhabit (McCutcheon and Moran, 2010). Pathogenic *Bacteroidetes*, such as *P. gingivalis* and *F. psychrophilum*, have a small genome typically around 2 Mb. This can be linked to the dedication of their metabolic capabilities toward the infection of specific sites. Living



in complex habitats and metabolizing a lot of different substrates, environmental, and intestinal species tend to have larger genomes (Figure 1), correlating with their broader catabolic capabilities. This is also the case of the opportunistic pathogen *B. fragilis*, which is part of the normal human gut microbiota but can cause infections at many other sites. To date, marine representatives harbor smaller genomes than their intestinal or terrestrial counterparts, but this will certainly progress as the number of complete genome sequences increases. For example, our group has recently annotated the genome of the marine *Flavobacteria Zobellia galactanivorans*. It comprises 5.5 Mb and encodes 4738 proteins, representing one of the largest genomes for a marine *Bacteroidetes* (Barbeyron et al., unpublished data). Interestingly, the proteorhodopsin-containing *P. dokdonensis* and *D. donghaensis* possess the smallest proteome among environmental species (2646 and 2284 proteins, respectively). Thus, the genome size fits the ecological niche. Big genomes increase the metabolic capacities, and hence broaden the spectrum of potential substrates for bacteria living in complex environments. In more stable habitats, bacteria tend to specialize toward specific functions and harbor smaller genomes. This raises the question of the nature of the ancestral *Bacteroidetes* genome. Indeed, the present variations in genome size could be due either to massive loss of genes from a large ancestral genome, or to successive acquisitions completing a small genome with genes representing a selective advantage.

In this respect, the sequencing era unraveled the plasticity of *Bacteroidetes* genomes, which evolved, and probably still evolve, through dynamic processes. The outcome of this plasticity reflects in the rapid deterioration of the global synteny between evolutionary-related species living in the same environment, as revealed for gut and rumen *Bacteroidetes* (Xu et al., 2007; Purushe et al., 2010). Their evolution is driven by highly frequent genetic rearrangements, gene duplications, and lateral gene transfers (LGT) between species. Genome analysis of two *B. fragilis* strains revealed extensive DNA inversions affecting the expression patterns of several genes (Kuwahara et al., 2004; Cerdeno-Tarraga et al., 2005). These events control the antigenic composition of bacterial surface structures and likely help *B. fragilis* evading the immune system and colonize novel sites. Notably, the expression of 20 SusC-like proteins, most of them coupled with SusD homologs, is regulated through DNA inversions (Kuwahara et al., 2004). This may participate in the cell adaptation to degrade specific polysaccharides found at the infection sites. *Bacteroidetes* evolution is also characterized by frequent gene duplications and further divergence in sequence and function, leading to considerably expanded paralogous groups. Notably, GH and SusC/SusD-like proteins are amongst the largest paralogous families in several sequenced *Bacteroidetes* (Xu et al., 2003, 2007; Bauer et al., 2006). In addition to this intra-strain plasticity, *Bacteroidetes* genomes evolve through inter-species exchange of genetic material (Thomas and Nielsen, 2005). Using a phyloge-

netic approach, Xu et al. (2007) showed that around 5.5% of the genes in gut *Bacteroidetes* genomes were laterally acquired from non gut-associated bacteria, among which glycosyltransferases (GT) were significantly over-represented. These LGT events could partly explain the niche specialization of different species. The authors suggest that acquisition of new genes from outside the gut brought novel metabolic pathways to intestinal *Bacteroidetes* and broaden the spectrum of digestible substrates. Furthermore, it has been shown that the convergence of GT and GH repertoires in gut *Bacteroidetes* sharing the same habitat is largely due to massive LGT rather than gene duplications (Lozupone et al., 2008). Conjugative LGT events are also demonstrated to be responsible for antibiotic resistance spreading in natural communities of gut *Bacteroides* (Shoemaker et al., 2001). The exchange of genetic material is not necessarily restricted to closely related species and can overcome phylogenetic barriers. Indeed, glyceraldehyde-3-phosphate dehydrogenase genes have been horizontally transferred from a  $\beta$ -proteobacteria to a *Bacteroidetes* (Figge et al., 1999). Another example is the transfer of genes between an Archaea and the hyperhalophilic *Sphingobacteria* *S. ruber* (Mongodin et al., 2005).

Taken together, recent analyses of *Bacteroidetes* genome sequences have shown that: (i) there is a gradation in the size of the genomes correlated with the functional specialization; (ii) genomes can undergo massive reorganizations; (iii) highly frequent LGT events allow spreading of novel metabolic capabilities inside *Bacteroidetes* populations. As already mentioned, intestinal *Bacteroidetes* are specialized in the degradation of plant-derived polymers, a feature shared with environmental relatives. In the last part of this review, we will discuss the potential connections between these two communities that are not *prima facie* obviously interacting.

### THE FOOD CONNECTION: TELL ME WHAT YOU EAT, I WILL TELL YOU WHAT YOUR BACTERIA CAN DO

Several studies have shown that the diet strongly influences the intestinal microbiota. Early research focused on the comparison of fecal microbes retrieved from individuals with different nutritional habits. Benno et al. (1986) showed significant variations in the cultivable microbiota of rural Japanese and urban Canadians. They proposed that this discrepancy relates to the contrasted diet of the two populations. Similarly, in a rat model transplanted with human microbiota, the consumption of resistant starch changed the bacterial composition compared to a sucrose diet (Silvi et al., 1999). Yet, the use of cultivable bacterial counts is a limited method (Amann et al., 1995) and no statistical difference was found when comparing the fecal microbiota of strictly vegetarians and individuals consuming a general diet (Goldberg et al., 1977). The development of culture-independent techniques to assess bacterial abundance and diversity helped testing the influence of diet on the GIT microbiota. In a mouse model reproducing the human intestinal microbiome, the bacterial community composition and the representation of metabolic pathways was strongly dependant on the nature of the diet (Turnbaugh et al., 2009). The proportion of *Bacteroidetes* representatives decreased drastically when animals were switched from a chow containing low levels of fat and high level of plant polysaccharides to a Western diet (high fat, high sugar). Feeding on the diet rich in plant polysaccharides resulted in an enriched set of pathways including *N*-glycan, glycosaminoglycan,

and starch degradation that are typical for *Bacteroidetes* (Turnbaugh et al., 2009). An independent study on a murine model showed that a high-fat diet was associated with a decrease in more than 30 lineages within the *Bacteroidetes* phylum, including in the *Bacteroidaceae*, *Prevotellaceae*, and *Rikenellaceae* families (Hildebrandt et al., 2009). Recent studies have investigated the diet impact on the human gut microbiota. The consumption of chemically modified resistant starch (RS4) instead of normal, digestible starch led to a shift in the bacterial community (Martinez et al., 2010). Even if results varied substantially between the 10 considered subjects, RS4 consumption was notably followed by enrichment in *Bacteroidetes*, among which *Parabacteroides distasonis* increased sevenfold. The observed changes were completely reversible within 1 week, demonstrating the high population dynamics (Martinez et al., 2010). In a recent comparison of the fecal microbiota of children from Burkina Faso, and Italy, De Filippo et al. (2010) showed a significant difference in the community composition. African children showed a higher proportion of *Bacteroidetes* (57 vs. 22%) and a lower proportion of *Firmicutes* (27 vs. 63%) than Europeans. The authors explained this difference by the higher dietary fiber content of the rural African food, mainly composed of cereals, legumes, and vegetables, which would favor the development of the polysaccharide-degrading *Bacteroidetes*. Interestingly, the genera *Prevotella*, *Xylanibacter*, *Cytophaga*, and *Paludibacter* were found exclusively in African microbiota. This is probably due to their increased fitness to grow on polysaccharides abundant in the Burkina Faso diet, such as xylan or cellulose (De Filippo et al., 2010). The control of gut microbiome composition by the diet quality likely denotes a selection of the population that optimally degrades the available substrates. In the case of *Bacteroidetes*, the selection criteria would primarily be based on the ability to digest complex polymers. Some species may have acquired specific catabolic pathways that others lack. This hypothesis has been recently tested in a mouse model. Germ-free mice were inoculated with two *Bacteroides thetaiotaomicron* and *B. caccae* strains of which the latter one can grow with inulin as carbon source. When the mice were fed with an inulin rich diet the ratio of the two species changed toward *B. caccae*. The ability to use inulin was associated with a GH32 absent in *B. thetaiotaomicron*. This clearly showed that diet selects species composition in the animal intestine (Sonnenburg et al., 2010).

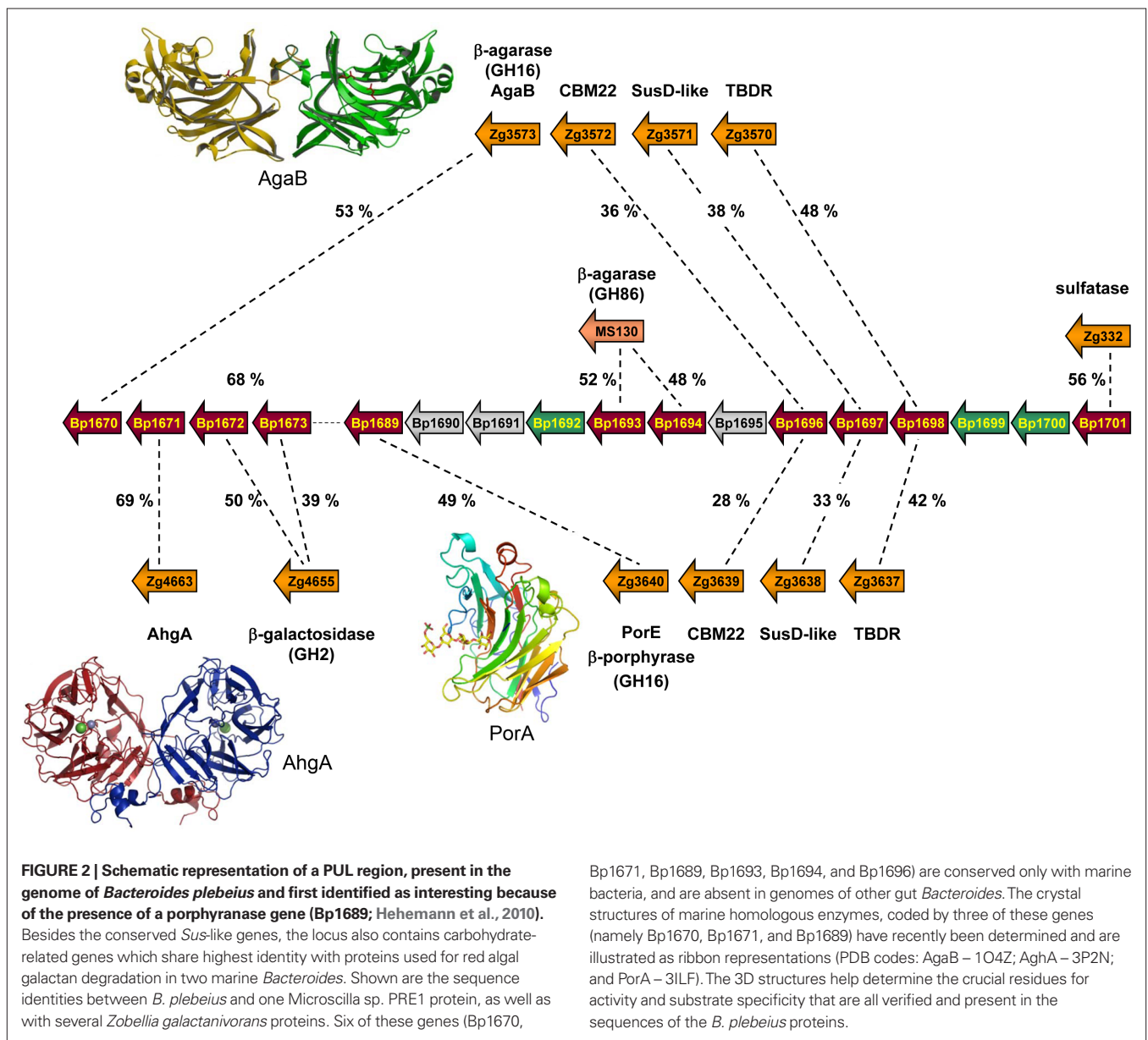
The question arises if different human populations with different diets contain specific food adaptations on the genetic level of their gut microbes. Recently, our group has shown that gut *Bacteroidetes* were able to get gene updates from environmental species to acquire novel functions (Hehemann et al., 2010; Rebuffet et al., 2011). Indeed, in the marine flavobacterium *Zobellia galactanivorans*, we have discovered and characterized the first porphyranases (Hehemann et al., 2010) as well as a 1,3- $\alpha$ -3,6-anhydro-1-galactosidase (Rebuffet et al., 2011). These enzymes are used by marine bacteria to degrade agarocolloids, sulfated galactans only found in the cell walls of red algae, such as *Porphyra* or *Gracilaria*. When using these new sequences as lead sequences to probe publicly available databases, homologs were identified not only in other marine bacteria, but surprisingly also in the human gut isolate *Bacteroides plebeius*. The genome of *B. plebeius* (DSM 17135) contains a porphyran/agar degradation locus, transferred from an ancestral marine *Bacteroidetes* (Hehemann et al., 2010; Rebuffet et al., 2011). This PUL was identified as a result of biochemical and



structural characterization of the first two porphyranases belonging to the family GH16 (PorA and PorB; Hehemann et al., 2010) and the 1,3- $\alpha$ -3,6-anhydro-L-galactosidase AhgA belonging to the family GH117 (Rebuffet et al., 2011). The PUL contains sequences coding for two putative  $\beta$ -agarases (GH86), one  $\beta$ -agarase (GH16), two  $\beta$ -galactosidases (GH2), a sulfatase, a carbohydrate-binding module, and a susD-like gene associated with its TonB-dependant receptor (Figure 2). Altogether these enzymes form a complete system of detection and degradation for porphyran and agar, which provides *B. plebeius* with the set of utensils to use these polysaccharides as carbon source.

Metagenomic data revealed that porphyranases and 1,3- $\alpha$ -3,6-anhydro-L-galactosidases are absent in North American and Danish population but present in Spanish and Japanese populations with proportions of 10 and 38% respectively (Hehemann et al., 2010;

Rebuffet et al., 2011). The biological rationale of “marine” enzymes in gut microbes could be linked to the high input of sea-derived products in the diet of these two populations. Indeed, the Japanese consume about 14.2 g seaweed per day and person (Fukuda et al., 2007), and the most popular seaweed is Nori (*Porphyra* spp.) used to make maki-sushi (Nisizawa et al., 1987). Similarly, Spain is the second largest consuming nation of seafood in the world (Manrique and Jensen, 2001). In both populations, contact between human-associated microbes and non-sterile seaweed or seafood could have created a favorable condition for a LGT from marine bacteria to human gut *Bacteroidetes*. Noteworthy, Nori is the only food that contains porphyran, which allowed associating the transfer of these genes to one special food source. This first evidence of a life style-associated adaptation of the genetic repertoire of the human gut microbiome could be detected due to the unique signature of



seaweed degrading CAZymes. Therefore marine CAZymes may be particularly good probes to reveal such adaptations in the human gut microbiome.

These examples of LGT show that gut bacteria are able to acquire new functions via transfer of a complete degradation pathway from food-associated environmental bacteria. Interestingly, the *Firmicutes Epulopiscium* sp., which lives in the gut of grazer surgeonfish (Clements and Bullivant, 1991), also possesses a putative agar degradation locus acquired from marine bacteria (Rebuffet et al., 2011). One can therefore extrapolate that similar events occurred in omnivorous and herbivorous animals. During the course of evolution, the metabolic repertoire of their gut microbes could have been influenced by contacts with food-associated environmental bacteria. Thus, acquisition of selectively advantageous genes by successive LGT events could explain how gut symbionts acquired CAZymes involved in green plant polysaccharide degradation.

In conclusion, some crucial criteria seem to have particularly favored the adaptation of *Bacteroidetes* to such contrasting environments, rendering their distribution/dominance close to ubiquitous. Among these, we note the presence of specific Sus-like PULs that are dedicated to polysaccharide degradation (Xu et al., 2003;

Flint et al., 2008) and their extreme genome flexibility (Xu et al., 2007) that assures the capacity to gain new functions by LGT from species living in different habitats. Consequently, the recently revealed LGT linking environmental and gut *Bacteroidetes* (Hehemann et al., 2010) raises the question as to how the ancestral members of the intestinal microbiota might have evolved from plant-degrading species, settled in the GIT of early herbivores, such as marine protists. Evidently, gut symbionts have co-evolved with their hosts (De Filippo et al., 2010), notably in response to the consumption of specific food, and this interdependence likely shaped and still shapes the lifestyle of human populations (Ley et al., 2006a). This “food connection” points toward the fact that recurrent contacts between environmental and gut microbes can have beneficial effects (Rook and Brunet, 2005). At the same time, it underlines the potential problem of our modern lifestyle and the consumption of hyper-hygienic, extensively processed food for human health, depriving us of the environmental reservoirs of microbial genes that allow adaptation by lateral transfer (De Filippo et al., 2010; Sonnenburg, 2010). On the other hand, global travel and trade are providing contact to new types of food and to diverse populations, perhaps giving access to new microbes harboring novel genes destined for integration into our microbiome.

## REFERENCES

- Abrams, G. D. (1983). “Impact of the intestinal microflora on intestinal structure and function,” in *Human Intestinal Microflora in Health and Disease*, ed. D. J. Hentges (New York, NY: Academic Press), 292–310.
- Allredge, A. L., Cole, J. J., and Caron, D. A. (1986). Production of heterotrophic bacteria inhabiting macroscopic organic aggregates (marine snow) from surface waters. *Limnol. Oceanogr.* 31, 68–78.
- Amann, R. L., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.
- Anderson, K. L., and Salyers, A. A. (1989a). Biochemical evidence that starch breakdown by *Bacteroides thetaio-taomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *J. Bacteriol.* 171, 3192–3198.
- Anderson, K. L., and Salyers, A. A. (1989b). Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaio-taomicron*. *J. Bacteriol.* 171, 3199–3204.
- Andersson, A. F., Lindberg, M., Jakobsson, H., Backhed, F., Nyren, P., and Engstrand, L. (2008). Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS ONE* 3, e2836. doi: 10.1371/journal.pone.0002836
- Austin, B., and Austin, D. A. (2007). *Bacterial Fish Pathogens: Disease of Farmed and Wild Fish*. Edinburgh: Springer Verlag.
- Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* 5, 782–791.
- 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.
- Balakirev, E. S., Pavlyuchkov, V. A., and Ayala, F. J. (2008). DNA variation and symbiotic associations in phenotypically diverse sea urchin *Strongylocentrotus intermedius*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16218–16223.
- Barbeyron, T., L’haridon, S., Corre, E., Kloareg, B., and Potin, P. (2001). *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of Flavobacteriaceae isolated from a red alga, and classification of. *Int. J. Syst. Evol. Microbiol.* 51, 985–997.
- Barbeyron, T., L’haridon, S., Michel, G., and Czjzek, M. (2008). *Mariniflexile fucanivorans* sp. nov., a marine member of the Flavobacteriaceae that degrades sulphated fucans from brown algae. *Int. J. Syst. Evol. Microbiol.* 58, 2107–2113.
- Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., Wurdemann, C. A., Quast, C., Kuhl, H., Knaust, F., Woebken, D., Bischof, K., Musmann, M., Choudhuri, J. V., Meyer, F., Reinhardt, R., Amann, R. I., and Glockner, F. O. (2006). Whole genome analysis of the marine Bacteroidetes “*Gramella forsetii*” reveals adaptations to degradation of polymeric organic matter. *Environ. Microbiol.* 8, 2201–2213.
- Beja, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., Jovanovich, S. B., Gates, C. M., Feldman, R. A., Spudich, J. L., Spudich, E. N., and Delong, E. F. (2000). Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289, 1902–1906.
- Beleneva, I. A., and Zhukova, N. V. (2006). Bacterial communities of brown and red algae from Peter the Great Bay, the Sea of Japan. *Mikrobiologiya* 75, 410–419.
- Benno, Y., Suzuki, K., Narisawa, K., Bruce, W. R., and Mitsuoka, T. (1986). Comparison of the fecal microflora in rural Japanese and urban Canadians. *Microbiol. Immunol.* 30, 521–532.
- Berg, J. O., Lindqvist, L., and Nord, C. E. (1980). Purification of glycoside hydrolases from *Bacteroides fragilis*. *Appl. Environ. Microbiol.* 40, 40–47.
- Bergey’s. (2011). *Bergey’s Manual of Systematic Bacteriology*. New York: Springer.
- Bernardet, J.-F., and Bowman, J. P. (2006). The genus *Flavobacterium*. *Prokaryotes* 7, 481–531.
- Bernardet, J.-F., and Nakagawa, Y. (2006). An introduction to the family of Flavobacteriaceae. *Prokaryotes* 7, 455–480.
- Bik, E. M., Eckburg, P. B., Gill, S. R., Nelson, K. E., Purdom, E. A., Francois, F., Perez-Perez, G., Blaser, M. J., and Relman, D. A. (2006). Molecular analysis of the bacterial microbiota in the human stomach. *Proc. Natl. Acad. Sci. U.S.A.* 103, 732–737.
- Borneman, J., Skroch, P. W., O’sullivan, K. M., Palus, J. A., Rumjanek, N. G., Jansen, J. L., Nienhuis, J., and Triplett, E. W. (1996). Molecular microbial diversity of an agricultural soil in Wisconsin. *Appl. Environ. Microbiol.* 62, 1935–1943.
- Brook, I. (2010). The role of anaerobic bacteria in bacteremia. *Anaerobe* 16, 183–189.
- Buckley, D. H., and Schmidt, T. M. (2003). Diversity and dynamics of microbial communities in soils from agro-ecosystems. *Environ. Microbiol.* 5, 441–452.
- 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–D238.
- Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30.
- Carroll, E. J., and Hungate, R. E. (1954). The magnitude of the microbial fermentation in the bovine rumen. *Appl. Microbiol.* 2, 205–214.
- Cerdeno-Tarraga, A. M., Patrick, S., Crossman, L. C., Blakely, G., Abratt, V., Lennard, N., Poxton, I., Duerden, B., Harris, B., Quail, M. A., Barron, A., Clark, L., Corton, C., Doggett, J., Holden, M. T., Larke, N., Line, A., Lord, A., Norbertczak, H., Ormond, D., Price, C., Rabinowitz, E., Woodward, J., Barrell, B., and Parkhill, J. (2005). Extensive DNA insertions in the *B. fragilis* genome control variable gene expression. *Science* 307, 1463–1465.

- Chin, W. C., Orellana, M. V., and Verdugo, P. (1998). Spontaneous assembly of marine dissolved organic matter into polymer gels. *Nature* 391, 568–572.
- Church, M. J. (2008). “Resource control of bacterial dynamics in the sea,” in *Microbial Ecology of the Oceans*, 2nd Edn, ed. D. L. Kirchman (Hoboken, NJ: Wiley & Sons), 335–382.
- Clements, K. D., and Bullivant, S. (1991). An unusual symbiont from the gut of surgeonfishes may be the largest known prokaryote. *J. Bacteriol.* 173, 5359–5362.
- Cottrell, M. T., and Kirchman, D. L. (2000). Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 66, 5116–5122.
- Crump, B. C., Armbrust, E. V., and Baross, J. A. (1999). Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia river, its estuary, and the adjacent coastal ocean. *Appl. Environ. Microbiol.* 65, 3192–3204.
- Crump, B. C., and Koch, E. W. (2008). Attached bacterial populations shared by four species of aquatic angiosperms. *Appl. Environ. Microbiol.* 74, 5948–5957.
- Davies, G. J., Gloster, T. M., and Henrissat, B. (2005). Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr. Opin. Struct. Biol.* 15, 637–645.
- Davis, H. S. (1922). “A new bacterial disease of fresh-water fishes,” in *Bulletin of the U.S. Bureau of Fisheries*, Washington, DC, 261–280.
- De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., Collini, S., Pieraccini, G., and Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14691–14696.
- D’Elia, J. N., and Salyers, A. A. (1996). Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* 178, 7180–7186.
- Del Rio, T. G., Abt, B., Spring, S., Lapidus, A., Nolan, M., Tice, H., Copeland, A., Cheng, J.-F., Chen, F., Hugenholtz, P., Kyrpides, N. C., Klenk, H.-P., and Lucas, S. (2010). Complete genome sequence of *Chitinophaga pinensis* type strain (UQM 2034T). *Stand. Genomic Sci.* 2, 87–95.
- DeLong, E. F., and Beja, O. (2010). The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS Biol.* 8, e1000359. doi: 10.1371/journal.pbio.1000359
- DeLong, E. F., Franks, D. G., and Alldredge, A. L. (1993). Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* 38, 924–934.
- Dubos, R., Schaedler, R. W., Costello, R., and Hoet, P. (1965). Indigenous, normal, and autochthonous flora of the gastrointestinal tract. *J. Exp. Med.* 122, 67.
- Duchaud, E., Boussaha, M., Loux, V., Bernardet, J. F., Michel, C., Kerouault, B., Mondot, S., Nicolas, P., Bossy, R., Caron, C., Bessieres, P., Gibrat, J. F., Claverol, S., Dumetz, F., Le Henaff, M., and Benmansour, A. (2007). Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*. *Nat. Biotechnol.* 25, 763–769.
- Duerden, B. I. (1994). Virulence factors in anaerobes. *Clin. Infect. Dis.* 18, 253–259.
- Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E., and Relman, D. A. (2005). Diversity of the human intestinal microbial flora. *Science* 308, 1635–1638.
- Edwards, J. L., Smith, D. L., Connolly, J., McDonald, J. E., Cox, M. J., Joint, I., Edwards, C., and McCarthy, A. J. (2010). Identification of carbohydrate metabolism genes in the metagenome of a marine biofilm community shown to be dominated by Gammaproteobacteria and Bacteroidetes. *Genes* 1, 371–384.
- Fandino, L. B., Riemann, L., Steward, G. F., Long, R. A., and Azam, F. (2001). Variations in bacterial community structure during a dinoflagellate bloom analyzed by DGGE and 16S rDNA sequencing. *Aquat. Microb. Ecol.* 23, 119–130.
- Ferraris, R. P. (2001). Dietary and developmental regulation of intestinal sugar transport. *Biochem. J.* 360, 265–276.
- Figge, R. M., Schubert, M., Brinkmann, H., and Cerff, R. (1999). Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra- and inter-kingdom gene transfer. *Mol. Biol. Evol.* 16, 429–440.
- Fletcher, H. M., Schenkein, H. A., Morgan, R. M., Bailey, K. A., Berry, C. R., and Macrina, F. L. (1995). Virulence of a *Porphyromonas gingivalis* W83 mutant defective in the prtH gene. *Infect. Immun.* 63, 1521–1528.
- Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R., and White, B. A. (2008). Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* 6, 121–131.
- Frias-Lopez, J., Zerkle, A. L., Bonheyo, G. T., and Fouke, B. W. (2002). Partitioning of bacterial communities between seawater and healthy, black band diseased, and dead coral surfaces. *Appl. Environ. Microbiol.* 68, 2214–2228.
- Fukuda, S., Saito, H., Nakaji, S., Yamada, M., Ebine, N., Tsushima, E., Oka, E., Kumeta, K., Tsukamoto, T., and Tokunaga, S. (2007). Pattern of dietary fiber intake among the Japanese general population. *Eur. J. Clin. Nutr.* 61, 99–103.
- Gibson, F. C., and Attardo Genco, C. (2006). The genus *Porphyromonas*. *Prokaryotes* 7, 428–454.
- Ghirring, T. M., Humphrys, M., Mills, H. J., Huettel, M., and Kostka, J. E. (2009). Identification of phytodetritus-degrading microbial communities in sublittoral Gulf of Mexico sands. *Limnol. Oceanogr.* 54, 1073–1083.
- Glöckner, F. O., Fuchs, B. M., and Amann, R. (1999). Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 65, 3721–3726.
- Goldberg, M. J., Smith, J. W., and Nichols, R. L. (1977). Comparison of the fecal microflora of seventh-day adventists with individuals consuming a general diet. Implications concerning colonic carcinoma. *Ann. Surg.* 186, 97–100.
- Gomez-Consarnau, L., Gonzalez, J. M., Coll-Llado, M., Gourdon, P., Pascher, T., Neutze, R., Pedros-Alio, C., and Pinhassi, J. (2007). Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature* 445, 210–213.
- Gomez-Pereira, P. R., Fuchs, B. M., Alonso, C., Oliver, M. J., Van Beusekom, J. E., and Amann, R. (2010). Distinct flavobacterial communities in contrasting water masses of the north Atlantic Ocean. *ISME J.* 4, 472–487.
- Gonzalez, J. M., Fernandez-Gomez, B., Fernandez-Guerra, A., Gomez-Consarnau, L., Sanchez, O., Coll-Llado, M., Del Campo, J., Escudero, L., Rodriguez-Martinez, R., Alonso-Saez, L., Latasa, M., Paulsen, I., Nedashkovskaya, O., Lekunberri, I., Pinhassi, J., and Pedros-Alio, C. (2008). Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proc. Natl. Acad. Sci. U.S.A.* 105, 8724–8729.
- Gooday, A. J. (2002). Biological responses to seasonally varying fluxes of organic matter to the ocean floor: a review. *J. Oceanogr.* 58, 305–332.
- Graham, L. E., Cook, M. E., and Busse, J. S. (2000). The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4535–4540.
- Greiner, D., and Mayrand, D. (1987). Selected characteristics of pathogenic and nonpathogenic strains of *Bacteroides gingivalis*. *J. Clin. Microbiol.* 25, 738–740.
- Han, C., Spring, S., Lapidus, A., Del Rio, T. G., Tice, H., Copeland, A., Cheng, J.-F., Lucas, S., Chen, F., Nolan, M., Bruce, D., Hugenholtz, P., Kyrpides, N. C., Klenk, H.-P., and Detter, J. C. (2009). Complete genome sequence of *Pedobacter heparinustype* strain (HIM 762-3T). *Stand. Genomic Sci.* 1, 54–62.
- Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010). Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* 464, 908–912.
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280, 309–316.
- Hildebrandt, M. A., Hoffmann, C., Sherrill-Mix, S. A., Keilbaugh, S. A., Hamady, M., Chen, Y. Y., Knight, R., Ahima, R. S., Bushman, F., and Wu, G. D. (2009). High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* 137, 1716–1724, e1711–1712.
- Hongoh, Y., Sharma, V. K., Prakash, T., Noda, S., Toh, H., Taylor, T. D., Kudo, I., Sakaki, Y., Toyoda, A., Hattori, M., and Ohkuma, M. (2008). Genome of an endosymbiont coupling N<sub>2</sub> fixation to cellulolysis within protist cells in termite gut. *Science* 322, 1108–1109.
- Hooper, L. V., Midtvedt, T., and Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu. Rev. Nutr.* 22, 283–307.
- Hugo, C. J., Bruun, B., and Jooste, P. J. (2006a). The genera *Bergeyella* and *Weeksella*. *Prokaryotes* 7, 532–538.
- Hugo, C. J., Bruun, B., and Jooste, P. J. (2006b). The genera *Empedobacter* and *Myroides*. *Prokaryotes* 7, 630–637.
- Hugo, C. J., Jooste, P. J., Segers, P., Vancanneyt, M., and Kersters, K. (1999). A polyphasic taxonomic study of *Chryseobacterium* strains isolated from dairy sources. *Syst. Appl. Microbiol.* 22, 586–595.
- Jenkinson, H. F., and Lamont, R. J. (2005). Oral microbial communities in sickness and in health. *Trends Microbiol.* 13, 589–595.
- Jooste, P. J., and Hugo, C. J. (1999). The taxonomy, ecology and cultivation of bacterial genera belonging to the family Flavobacteriaceae. *Int. J. Food Microbiol.* 53, 81–94.
- Julies, E. M., Fuchs, B. M., Arnosti, C., and Bruchert, V. (2010). Organic carbon degradation in anoxic



- organic-rich shelf sediments: biogeochemical rates and microbial abundance. *Geomicrobiol. J.* 27, 303–314.
- Keijsers, B. J., Zaura, E., Huse, S. M., Van Der Vossen, J. M., Schuren, F. H., Montijn, R. C., Ten Cate, J. M., and Crielaard, W. (2008). Pyrosequencing analysis of the oral microflora of healthy adults. *J. Dent. Res.* 87, 1016–1020.
- Kim, B. Y., Weon, H. Y., Cousin, S., Yoo, S. H., Kwon, S. W., Go, S. J., and Stackebrandt, E. (2006). *Flavobacterium daejeonense* sp. nov. and *Flavobacterium suncheonense* sp. nov., isolated from greenhouse soils in Korea. *Int. J. Syst. Evol. Microbiol.* 56, 1645–1649.
- Kim, H.-J., Miller, A. J., McGowan, J., and Carter, M. L. (2009). Coastal phytoplankton blooms in the Southern California Bight. *Prog. Oceanogr.* 82, 137–147.
- Kim, K. K., Kim, M. K., Lim, J. H., Park, H. Y., and Lee, S. T. (2005). Transfer of *Chryseobacterium meningosepticum* and *Chryseobacterium miricola* to *Elizabethkingia* gen. nov. as *Elizabethkingia meningoseptica* comb. nov. and *Elizabethkingia miricola* comb. nov. *Int. J. Syst. Evol. Microbiol.* 55, 1287–1293.
- Kim, Y. S., and Milner, J. A. (2007). Dietary modulation of colon cancer risk. *J. Nutr.* 137, 2576–2579.
- King, E. O. (1959). Studies on a group of previously unclassified bacteria associated with meningitis in infants. *Am. J. Clin. Pathol.* 31, 241–247.
- Kirchman, D. L. (2002). The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol. Ecol.* 39, 91–100.
- Kirchman, D. L. (2008). New light on an important microbe in the ocean. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8487–8488.
- 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.
- Knapp, B. A., Seeber, J., Rief, A., Meyer, E., and Insam, H. (2010). Bacterial community composition of the gut microbiota of *Cylindroiulus fulviceps* (diplopoda) as revealed by molecular fingerprinting and cloning. *Folia Microbiol. (Praha)* 55, 489–496.
- Koropatkin, N. M., Martens, E. C., Gordon, J. I., and Smith, T. J. (2008). Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* 16, 1105–1115.
- Koropatkin, N. M., and Smith, T. J. (2010). SusG: a unique cell-membrane-associated alpha-amylase from a prominent human gut symbiont targets complex starch molecules. *Structure* 18, 200–215.
- Kuwahara, T., Yamashita, A., Hirakawa, H., Nakayama, H., Toh, H., Okada, N., Kuhara, S., Hattori, M., Hayashi, T., and Ohnishi, Y. (2004). Genomic analysis of *Bacteroides fragilis* reveals extensive DNA inversions regulating cell surface adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14919–14924.
- Lail, K., Sikorski, J., Saunders, E., Lapidus, A., Del Rio, T. G., Copeland, A., Tice, H., Cheng, J.-F., Lucas, S., Kyrpides, N. C., Klenk, H.-P., and Chen, F. (2010). Complete genome sequence of *Spirosoma linguale* type strain (1T). *Stand. Genomic Sci.* 2, 176–185.
- Laine, R. A. (1994). A calculation of all possible oligosaccharide isomers both branched and linear yields  $1.05 \times 10^{12}$  structures for a reducing hexasaccharide: the isomer barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4, 759–767.
- Lang, E., Lapidus, A., Chertkov, O., Brettin, T., Detter, J. C., Han, C., Copeland, A., Del Rio, T. G., Nolan, M., Chen, F., Hugenholtz, P., Göker, M., Rohde, M., Kyrpides, N. C., and Klenk, H.-P. (2009). Complete genome sequence of *Dyadobacter fermentans* type strain (NS114T). *Stand. Genomic Sci.* 1, 133–140.
- Largo, D. B., Fukami, K., and Nishijima, T. (1995). Occasional pathogenic bacteria promoting ice-ice disease in the carrageenan-producing red algae *Kappaphycus alvarezii* and *Eucheuma denticulatum* (Solieriaceae, Gigartinales, Rhodophyta). *J. Appl. Phycol.* 7, 545–554.
- Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* 75, 5111–5120.
- Leadbetter, E. R. (2006). The genus *Capnocytophaga*. *Prokaryotes* 7, 709–711.
- Lee, M., Woo, S. G., Park, J., and Yoo, S. A. (2010). *Dyadobacter soli* sp. nov., a starch-degrading bacterium isolated from farm soil. *Int. J. Syst. Evol. Microbiol.* 60, 2577–2582.
- Leng, J., Xie, L., Zhu, R., Yang, S., Gou, X., Li, S., and Mao, H. (2010). Dominant bacterial communities in the rumen of Gayals (*Bos frontalis*), Yaks (*Bos grunniens*) and Yunnan Yellow Cattle (*Bos taurus*) revealed by denaturing gradient gel electrophoresis. *Mol. Biol. Rep.* PMID:21140219. [Epub ahead of print].
- Leser, T. D., Amenuvor, J. Z., Jensen, T. K., Lindecrona, R. H., Boye, M., and Møller, K. (2002). Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl. Environ. Microbiol.* 68, 673–690.
- Ley, R. E., Peterson, D. A., and Gordon, J. I. (2006a). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848.
- Ley, R. E., Turnbaugh, P. J., Klein, S., and Gordon, J. I. (2006b). Microbial ecology: human gut microbes associated with obesity. *Nature* 444, 1022–1023.
- Liao, C. H., and Wells, J. M. (1986). Properties of *Cytophaga johnsonae* strains causing spoilage of fresh produce at food markets. *Appl. Environ. Microbiol.* 52, 1261–1265.
- Llobet-Brossa, E., Rossello-Mora, R., and Amann, R. (1998). Microbial community composition of Wadden Sea sediments as revealed by fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 64, 2691–2696.
- Loesche, W. J. (1993). Bacterial mediators in periodontal disease. *Clin. Infect. Dis.* 16, 203–210.
- Lopez-Sanchez, M. J., Neef, A., Peretó, J., Patiño-Navarrete, R., Pignatelli, M., Latorre, A., and Moya, A. (2009). Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cockroach *Blattella germanica*. *PLoS Genet.* 5, e1000721. doi: 10.1371/journal.pgen.1000721
- Lozupone, C. A., Hamady, M., Cantarel, B. L., Coutinho, P. M., Henrissat, B., Gordon, J. I., and Knight, R. (2008). The convergence of carbohydrate active gene repertoires in human gut microbes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15076–15081.
- Lu, J., Santo Domingo, J. W., Hill, S., and Edge, T. A. (2009). Microbial diversity and host-specific sequences of Canada goose feces. *Appl. Environ. Microbiol.* 75, 5919–5926.
- Mammeri, H., Bellais, S., and Nordmann, P. (2002). Chromosome-encoded beta-lactamases TUS-1 and MUS-1 from *Myroides odoratus* and *Myroides odoratimimus* (formerly *Flavobacterium odoratum*), new members of the lineage of molecular subclass B1 metalloenzymes. *Antimicrob. Agents Chemother.* 46, 3561–3567.
- Manrique, J., and Jensen, H. H. (2001). Spanish household demand for seafood. *J. Agric. Econ.* 52, 23–37.
- Marchesi, J. R. (2010). Prokaryotic and eukaryotic diversity of the human gut. *Adv. Appl. Microbiol.* 72, 43–62.
- Martens, E. C., Chiang, H. C., and Gordon, J. I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* 4, 447–457.
- Martens, E. C., Koropatkin, N. M., Smith, T. J., and Gordon, J. I. (2009). Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *J. Biol. Chem.* 284, 24673–24677.
- Martinez, I., Kim, J., Duffy, P. R., Schlegel, V. L., and Walter, J. (2010). Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. *PLoS ONE* 5, e15046. doi: 10.1371/journal.pone.0015046
- Martinez-Alonso, M., Escolano, J., Montesinos, E., and Gaju, N. (2010). Diversity of the bacterial community in the surface soil of a pear orchard based on 16S rRNA gene analysis. *Int. Microbiol.* 13, 123–134.
- Matsui, H., Kato, Y., Chikaraishi, T., Moritani, M., Ban-Tokuda, T., and Wakita, M. (2010). Microbial diversity in ostrich ceca as revealed by 16S ribosomal RNA gene clone library and detection of novel Fibrobacter species. *Anaerobe* 16, 83–93.
- Mavrommatis, K., Gronow, S., Saunders, E., Land, M., Lapidus, A., Copeland, A., Del Rio, T. G., Nolan, M., Lucas, S., Chen, F., Tice, H., Cheng, J.-F., Bruce, D., Goodwin, L., Pitluck, S., Pati, A., Ivanova, N., Chen, A., Palaniappan, K., Chain, P., Hauser, L., Chang, Y.-J., Jeffries, C. D., Brettin, T., Detter, J. C., Han, C., Bristow, J., Göker, M., Rohde, M., Eisen, J. A., Markowitz, V., Kyrpides, N. C., Klenk, H.-P., and Hugenholtz, P. (2009). Complete genome sequence of *Capnocytophaga ochracea* type strain (VPI 2845T). *Stand. Genomic Sci.* 1, 101–109.
- Mazmanian, S. K. (2008). Capsular polysaccharides of symbiotic bacteria modulate immune responses during experimental colitis. *J. Pediatr. Gastroenterol. Nutr.* 46(Suppl. 1), E11–E12.
- Mazmanian, S. K., Round, J. L., and Kasper, D. L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* 453, 620–625.
- McBride, M. J., Xie, G., Martens, E. C., Lapidus, A., Henrissat, B., Rhodes, R. G., Goltsman, E., Wang, W., Xu, J., Hunnicutt, D. W., Staroscik, A. M., Hoover, T. R., Cheng, Y. Q., and Stein, J. L. (2009). Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl. Environ. Microbiol.* 75, 6864–6875.
- McCutcheon, J. P., and Moran, N. A. (2010). Functional convergence in reduced genomes of bacterial symbi-



- onts spanning 200 my of evolution. *Genome Biol. Evol.* 2, 708–718.
- Michel, G., Tonon, T., Scornet, D., Cock, J. M., and Kloareg, B. (2010a). The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytol.* 188, 82–97.
- Michel, G., Tonon, T., Scornet, D., Cock, J. M., and Kloareg, B. (2010b). Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in Eukaryotes. *New Phytol.* 188, 67–81.
- Middelbos, I. S., Vester Boler, B. M., Qu, A., White, B. A., Swanson, K. S., and Fahey, G. C. Jr. (2010). Phylogenetic characterization of fecal microbial communities of dogs fed diets with or without supplemental dietary fiber using 454 pyrosequencing. *PLoS ONE* 5, e9768. doi: 10.1371/journal.pone.0009768
- Mongodin, E. F., Nelson, K. E., Daugherty, S., Deboy, R. T., Wister, J., Khouri, H., Weidman, J., Walsh, D. A., Papke, R. T., Sanchez Perez, G., Sharma, A. K., Nesbo, C. L., Macleod, D., Bapteste, E., Doolittle, W. F., Charlebois, R. L., Legault, B., and Rodriguez-Valera, F. (2005). The genome of *Salinibacter ruber*: convergence and gene exchange among hyperhalophilic bacteria and archaea. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18147–18152.
- Moore, W. E., and Holdeman, L. V. (1974). Special problems associated with the isolation and identification of intestinal bacteria in fecal flora studies. *Am. J. Clin. Nutr.* 27, 1450–1455.
- Nagata, T. (2008). “Organic matter-bacteria interactions in seawater,” in *Microbial Ecology in the Oceans*, 2nd Edn, ed. D. L. Kirchman (Hoboken, NJ: Wiley & Sons), 207–242.
- Naito, M., Hirakawa, H., Yamashita, A., Ohara, N., Shoji, M., Yukitake, H., Nakayama, K., Toh, H., Yoshimura, E., Kuhara, S., Hattori, M., Hayashi, T., and Nakayama, K. (2008). Determination of the genome sequence of *Porphyromonas gingivalis* strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in *P. gingivalis*. *DNA Res.* 15, 215–225.
- Nasidze, I., Li, J., Quinque, D., Tang, K., and Stoneking, M. (2009a). Global diversity in the human salivary microbiome. *Genome Res.* 19, 636–643.
- Nasidze, I., Quinque, D., Li, J., Li, M., Tang, K., and Stoneking, M. (2009b). Comparative analysis of human saliva microbiome diversity by bar-coded pyrosequencing and cloning approaches. *Anal. Biochem.* 391, 64–68.
- Neill, K., Heesch, S., and Nelson, W. (2008). *Diseases, Pathogens and Parasites of Undaria pinnatifida*. Wellington: Ministry of Agriculture and Forestry.
- Nelson, K. E., Fleischmann, R. D., Deboy, R. T., Paulsen, I. T., Fouts, D. E., Eisen, J. A., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Gwinn, M., Haft, D. H., Kolonay, J. F., Nelson, W. C., Mason, T., Tallon, L., Gray, J., Granger, D., Tettelin, H., Dong, H., Galvin, J. L., Duncan, M. J., Dewhirst, F. E., and Fraser, C. M. (2003). Complete genome sequence of the oral pathogenic Bacterium *Porphyromonas gingivalis* strain W83. *J. Bacteriol.* 185, 5591–5601.
- Niklas, K. J. (2004). The cell walls that bind the tree of life. *Bioscience* 54, 831–841.
- Nisizawa, K., Noda, H., Kikuchi, R., and Watanabe, T. (1987). The main seaweed foods in Japan. *Hydrobiologia* 151/152, 5–29.
- Nolan, M., Tindall, B. J., Pomrenke, H., Lapidus, A., Copeland, A., Glavina Del Rio, T., Lucas, S., Chen, F., Tice, H., Cheng, J. F., Saunders, E., Han, C., Bruce, D., Goodwin, L., Chain, P., Pitluck, S., Ovchinnikova, G., Pati, A., Ivanova, N., Mavromatis, K., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. D., Brettin, T., Göker, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C., Klenk, H. P., and Detter, J. C. (2009). Complete genome sequence of *Rhodothermus marinus* type strain (R-10). *Stand. Genomic Sci.* 1, 283–290.
- Oh, H. M., Giovannoni, S. J., Lee, K., Ferreira, S., Johnson, J., and Cho, J. C. (2009). Complete genome sequence of *Robiginitalea biformata* HTCC2501. *J. Bacteriol.* 191, 7144–7145.
- Oh, H. M., Kang, I., Ferreira, S., Giovannoni, S. J., and Cho, J. C. (2010). Complete genome sequence of *Croceibacter atlanticus* HTCC2559(T). *J. Bacteriol.* 192, 4796–4797.
- Oh, H. M., Kang, I., Yang, S. J., Jang, Y., Vergin, K. L., Giovannoni, S. J., and Cho, J. C. (2011). Complete genome sequence of strain HTCC2170, a novel member of the genus *Maribacter* in the family Flavobacteriaceae. *J. Bacteriol.* 193, 303–304.
- Pankratov, T. A., Kulichevskaya, I. S., Liesack, W., and Dedysh, S. N. (2006). Isolation of aerobic, gliding, xylanolytic and laminarinolytic bacteria from acidic Sphagnum peatlands and emended description of *Chitinophaga arvensicola* Kampfer et al. 2006. *Int. J. Syst. Evol. Microbiol.* 56, 2761–2764.
- Pei, Z., Bini, E. J., Yang, L., Zhou, M., Francois, F., and Blaser, M. J. (2004). Bacterial biota in the human distal esophagus. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4250–4255.
- Pena, A., Teeling, H., Huerta-Cepas, J., Santos, F., Yarla, P., Brito-Echeverria, J., Lucio, M., Schmitt-Kopplin, P., Meseguer, I., Schenowitz, C., Dossat, C., Barbe, V., Dopazo, J., Rosselló-Mora, R., Schüller, M., Glöckner, F. O., Amann, R., Gabaldón, T., and Antón, J. (2010). Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter ruber* strains. *ISME J.* 4, 882–895.
- Pinhassi, J., Azam, F., Hemphälä, J., Long, R. A., Martinez, J., Zweifel, U. L., and Hagström, A. (1999). Coupling between bacterioplankton species composition, population dynamics, and organic matter degradation. *Aquat. Microb. Ecol.* 17, 13–26.
- Ploug, H., Grossart, H. P., Azam, F., and Jørgensen, B. B. (1999). Photosynthesis, respiration, and carbon turnover in sinking marine snow from surface waters of Southern California Bight: implications for the carbon cycle in the ocean. *Mar. Ecol. Prog. Ser.* 179, 1–11.
- Pommier, T., Canback, B., Riemann, L., Bostrom, K. H., Simu, K., Lundberg, P., Tunlid, A., and Hagstrom, A. (2007). Global patterns of diversity and community structure in marine bacterioplankton. *Mol. Ecol.* 16, 867–880.
- Pope, P. B., Denman, S. E., Jones, M., Tringe, S. G., Barry, K., Malfatti, S. A., Mchardy, A. C., Cheng, J. F., Hugenholtz, P., Mcsweeney, C. S., and Morrison, M. (2010). Adaptation to herbivory by the Tamar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14793–14798.
- Popper, Z. A., Michel, G., Hervé, C., Domozych, S. C., Willats, W. G. T., Tuohy, M. G., Kloareg, B., and Stengel, D. B. (2011). Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant Biol.* 62, 567–590.
- Purushe, J., Fouts, D. E., Morrison, M., White, B. A., Mackie, R. I., Coutinho, P. M., Henrissat, B., and Nelson, K. E. (2010). Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii*: insights into their environmental niche. *Microb. Ecol.* 60, 721–729.
- Qin, Q. L., Zhang, X. Y., Wang, X. M., Liu, G. M., Chen, X. L., Xie, B. B., Dang, H. Y., Zhou, B. C., Yu, J., and Zhang, Y. Z. (2010). The complete genome of *Zunongwangia profunda* SM-A87 reveals its adaptation to the deep-sea environment and ecological role in sedimentary organic nitrogen degradation. *BMC Genomics* 11, 247. doi: 10.1186/1471-2164-11-247
- Raman, R., Sasisekharan, V., and Sasisekharan, R. (2005). Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem. Biol.* 12, 267–277.
- Rath, J., Wu, K. Y., Herndl, G. J., and Delong, E. F. (1998). High phylogenetic diversity in a marine-snow-associated bacterial assemblage. *Aquat. Microb. Ecol.* 14, 261–269.
- Ratner, H. (1984). *Flavobacterium meningosepticum*. *Infect. Control* 5, 237–239.
- Rebuffet, E., Groisillier, A., Thompson, A., Jeudy, A., Barbeyron, T., Czjzek, M., and Michel, G. (2011). Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ. Microbiol.* 13, 1253–1270.
- Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y., and Fierer, N. (2010). The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ. Microbiol.* 12, 2885–2893.
- Reeves, A. R., Wang, G. R., and Salyers, A. A. (1997). Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* 179, 643–649.
- Reichenbach, H. (2006). The order cytophagales. *Prokaryotes* 7, 549–590.
- Riemann, L., Steward, G. F., and Azam, F. (2000). Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Appl. Environ. Microbiol.* 66, 578–587.
- Riepe, S. P., Goldstein, J., and Alpers, D. H. (1980). Effect of secreted Bacteroides proteases on human intestinal brush border hydrolases. *J. Clin. Invest.* 66, 314–322.
- Rohwer, F., Seguritan, V., Azam, F., and Knowlton, N. (2002). Diversity and distribution of coral-associated bacteria. *Mar. Ecol. Prog. Ser.* 243, 1–10.
- Rook, G. A., and Brunet, L. R. (2005). Microbes, immunoregulation, and the gut. *Gut* 54, 317–320.
- Rudek, W., and Haque, R. U. (1976). Extracellular enzymes of the genus *Bacteroides*. *J. Clin. Microbiol.* 4, 458–460.
- Sabree, Z. L., Kambampati, S., and Moran, N. A. (2009). Nitrogen recycling and nutritional provisioning by Blattabacterium, the cockroach endosymbiont. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19521–19526.
- Salaun, S., Kervarec, N., Potin, P., Haras, D., Piotto, M., and La Barre, S. (2010). Whole-cell spectroscopy is a convenient tool to assist molecular identification of cultivatable marine bacteria and to investigate their adaptive metabolism. *Talanta* 80, 1758–1770.
- Salyers, A. A., Vercellotti, J. R., West, S. E., and Wilkins, T. D. (1977). Fermentation of mucin and plant polysaccharides by strains of *Bacteroides*

- from the human colon. *Appl. Environ. Microbiol.* 33, 319–322.
- Savage, D. C., Dubos, R., and Schaedler, R. W. (1968). The gastrointestinal epithelium and its autochthonous bacterial flora. *J. Exp. Med.* 127, 67–76.
- Schmitt-Wagner, D., Friedrich, M. W., Wagner, B., and Brune, A. (2003). Phylogenetic diversity, abundance, and axial distribution of bacteria in the intestinal tract of two soil-feeding termites (*Cubitermes* spp.). *Appl. Environ. Microbiol.* 69, 6007–6017.
- Schmitz-Esser, S., Tischler, P., Arnold, R., Montanaro, J., Wagner, M., Rattei, T., and Horn, M. (2010). The genome of the amoeba symbiont “*Candidatus Amoebophilus asiaticus*” reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J. Bacteriol.* 192, 1045–1057.
- Scupham, A. J., Patton, T. G., Bent, E., and Bayles, D. O. (2008). Comparison of theecal microbiota of domestic and wild turkeys. *Microb. Ecol.* 56, 322–331.
- Segers, P., Mannheim, W., Vancanneyt, M., De Brandt, K., Hinz, K. H., Kersters, K., and Vandamme, P. (1993). *Riemerella anatipestifer* gene nov. comb. nov., the causative agent of septicemia anserum exsudativa, and its phylogenetic affiliation within the *Flavobacterium-Cytophaga* rRNA homology group. *Int. J. Syst. Bacteriol.* 43, 768–776.
- Sghir, A., Gramet, G., Suau, A., Rochet, V., Pochart, P., and Dore, J. (2000). Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization. *Appl. Environ. Microbiol.* 66, 2263–2266.
- Shah, H. N., and Collins, D. M. (1990). *Prevotella*, a new genus to include *Bacteroides melaninogenicus* and related species formerly classified in the genus *Bacteroides*. *Int. J. Syst. Bacteriol.* 40, 205–208.
- Shipman, J. A., Berleman, J. E., and Salyers, A. A. (2000). Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *J. Bacteriol.* 182, 5365–5372.
- Shipman, J. A., Cho, K. H., Siegel, H. A., and Salyers, A. A. (1999). Physiological characterization of SusG, an outer membrane protein essential for starch utilization by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* 181, 7206–7211.
- Shoemaker, N. B., Vlamakis, H., Hayes, K., and Salyers, A. A. (2001). Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Appl. Environ. Microbiol.* 67, 561–568.
- Sievert, S. M., Kuever, J., and Muyzer, G. (2000). Identification of 16S ribosomal DNA-defined bacterial populations at a shallow submarine hydrothermal vent near Milos Island (Greece). *Appl. Environ. Microbiol.* 66, 3102–3109.
- Silvi, S., Rumney, C. J., Cresci, A., and Rowland, I. R. (1999). Resistant starch modifies gut microflora and microbial metabolism in human flora-associated rats inoculated with faeces from Italian and UK donors. *J. Appl. Microbiol.* 86, 521–530.
- Simek, K., Armengol, J., Comerma, M., Garcia, J. C., Kojecka, P., Nedoma, J., and Hejzlar, J. (2001). Changes in the epilimnetic bacterial community composition, production, and protist-induced mortality along the longitudinal axis of a highly eutrophic reservoir. *Microb. Ecol.* 42, 359–371.
- Simon, M., Glockner, F. O., and Amann, R. (1999). Different community structure and temperature optima of heterotrophic picoplankton in various regions of the Southern Ocean. *Aquat. Microb. Ecol.* 18, 275–284.
- Skerratt, J. H., Bowman, J. P., Hallegraaff, G., James, S., and Nichols, P. D. (2002). Algalicidal bacteria associated with blooms of a toxic dinoflagellate in a temperate Australian estuary. *Mar. Ecol. Prog. Ser.* 244, 1–15.
- Smith, C. J., and Bryant, M. P. (1979). Introduction to metabolic activities of intestinal bacteria. *Am. J. Clin. Nutr.* 32, 149–157.
- Smith, C. J., Rocha, E. R., and Paster, B. J. (2006). The medically important *Bacteroides* spp. in health and disease. *Prokaryotes* 7, 381–427.
- Sohn, J. H., Lee, J. H., Yi, H., Chun, J., Bae, K. S., Ahn, T. Y., and Kim, S. J. (2004). *Kordia algicida* gen. nov., sp. nov., an algalicidal bacterium isolated from red tide. *Int. J. Syst. Evol. Microbiol.* 54, 675–680.
- Sonnenburg, E. D., Zheng, H., Joglekar, P., Higginbottom, S. K., Firbank, S. J., Bolam, D. N., and Sonnenburg, J. L. (2010). Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* 141, 1241–1252.
- Sonnenburg, J. L. (2010). Microbiology: genetic pot luck. *Nature* 464, 837–838.
- Stahl, E. A., and Bishop, J. G. (2000). Plant-pathogen arms races at the molecular level. *Curr. Opin. Plant Biol.* 3, 299–304.
- Staufenger, T., Thiel, V., Wiese, J., and Imhoff, J. F. (2008). Phylogenetic analysis of bacteria associated with *Laminaria saccharina*. *FEMS Microbiol. Ecol.* 64, 65–77.
- Sugahara, K., and Kitagawa, H. (2002). Heparin and heparan sulfate biosynthesis. *IUBMB Life* 54, 163–175.
- Sunairi, M., Tsuchiya, H., Tsuchiya, T., Omura, Y., Koyanagi, Y., Ozawa, M., Iwabuchi, N., Murooka, H., and Nakajima, M. (1995). Isolation of a bacterium that causes anaika disease of the red algae *Porphyra yezoensis*. *J. Appl. Microbiol.* 79, 225–229.
- Suomalainen, L.-R., Tiirola, M., and Valtonen, E. T. (2006). Chondroitin AC lyase activity is related to virulence of fish pathogenic *Flavobacterium columnare*. *J. Fish Dis.* 29, 701–770.
- Tajima, K., Aminov, R. I., Nagamine, T., Ogata, K., Nakamura, M., Matsui, H., and Benno, Y. (1999). Rumen bacterial diversity as determined by sequence analysis of 16S rDNA libraries. *FEMS Microbiol. Ecol.* 29, 159–169.
- Tanner, A. C. R., Listgarten, M. A., Ebersole, J. L., and Strzempko, M. N. (1986). *Bacteroides forsythus* sp. nov., a slow-growing, *Fusiform Bacteroides* sp. from the human oral cavity. *Int. J. Syst. Bacteriol.* 36, 213–221.
- Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721.
- Tiery, Y., Bechet, M., Joncquiert, J. C., Dubourguier, H. C., and Guillaume, J. B. (1994). Molecular cloning and expression in *Escherichia coli* of genes encoding pectate lyase and pectin methylesterase activities from *Bacteroides thetaiotaomicron*. *J. Appl. Bacteriol.* 76, 592–602.
- Turnbaugh, P. J., Henriessat, B., and Gordon, J. I. (2010). Viewing the human microbiome through three-dimensional glasses: integrating structural and functional studies to better define the properties of myriad carbohydrate-active enzymes. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 66, 1261–1264.
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007). The human microbiome project. *Nature* 449, 804–810.
- Turnbaugh, P. J., Ridaura, V. K., Faith, J. J., Rey, F. E., Knight, R., and Gordon, J. I. (2009). The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* 1, 6ra14.
- Uyenco, F. R. (1977). “Microbiological studies of diseased *Eucheuma* sp. and other seaweeds,” in *National Seaweeds Symposium*, Manila, Philippines.
- van Hanne, E. J., Zwart, G., Van Agterveld, M. P., Gons, H. J., Ebert, J., and Laanbroek, H. J. (1999). Changes in bacterial and eukaryotic community structure after mass lysis of filamentous cyanobacteria associated with viruses. *Appl. Environ. Microbiol.* 65, 795–801.
- Vandamme, P., Segers, P., Vancanneyt, M., Van Hove, K., Muters, R., Hommez, J., Dewhirst, F., Paster, B., Kersters, K., Falen, E., Devriese, L. A., Bisgaard, M., Hinz, K. H., and Mannheim, W. (1994). *Ornithobacterium rhinotracheale* gen. nov., sp. nov. isolated from the avian respiratory tract. *Int. J. Syst. Bacteriol.* 44, 24–37.
- Vandamme, P., Vancanneyt, M., Segers, P., Ryll, M., Köhler, B., Ludwig, W., and Hinz, K. H. (1999). *Coenonia anatina* gen. nov., sp. nov., a novel bacterium associated with respiratory disease in ducks and geese. *Int. J. Syst. Bacteriol.* 49, 867–874.
- Wen, L., Ley, R. E., Volchkov, P. Y., Stranges, P. B., Avanesyan, L., Stonebraker, A. C., Hu, C., Wong, F. S., Szot, G. L., Bluestone, J. A., Gordon, J. I., and Chervonsky, A. V. (2008). Innate immunity and intestinal microbiota in the development of type 1 diabetes. *Nature* 455, 1109–1113.
- Wexler, H. M. (2007). Bacteroides: the good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* 20, 593–621.
- Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6578–6583.
- Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* 51, 221–271.
- Woese, C. R., Yang, D., Mandelco, L., and Stetter, K. O. (1990). The flexibacter-flavobacter connection. *Syst. Appl. Microbiol.* 13, 161–165.
- 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.
- Woyke, T., Xie, G., Copeland, A., Gonzalez, J. M., Han, C., Kiss, H., Saw, J. H., Senin, P., Yang, C., Chatterji, S., Cheng, J. F., Eisen, J. A., Sieracki, M. E., and Stepanauskas, R. (2009). Assembling the marine metagenome, one cell at a time. *PLoS ONE* 4, e5299. doi:10.1371/journal.pone.0005299
- Xie, G., Bruce, D. C., Challacombe, J. F., Chertkov, O., Detter, J. C., Gilna, P., Han, C. S., Lucas, S., Misra, M., Myers, G. L., Richardson, P., Tapia, R., Thayer, N., Thompson, L. S., Brettin, T. S., Henriessat, B., Wilson, D. B., and McBride, M. J. (2007). Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. *Appl. Environ. Microbiol.* 73, 3536–3546.
- 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., Mahowald, M. A., Ley, R. E., Lozupone, C. A., Hamady, M., Martens, E. C., Henriessat, B., Coutinho, P. M., Minx, P., Latreille, P., Cordum,

- H., Van Brunt, A., Kim, K., Fulton, R. S., Fulton, L. A., Clifton, S. W., Wilson, R. K., Knight, R. D., and Gordon, J. I. (2007). Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biol.* 5, e156. doi: 10.1371/journal.pone.0005299
- Zhou, J. P., Huang, Y., and Mo, M. (2009). Phylogenetic analysis on the soil bacteria distributed in karst forest. *Braz. J. Microbiol.* 40, 827–837.
- Zhu, X. Y., Zhong, T., Pandya, Y., and Joerger, R. D. (2002). 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. *Appl. Environ. Microbiol.* 68, 124–137.
- Zhu, Y., Suits, M. D., Thompson, A. J., Chavan, S., Dinev, Z., Dumon, C., Smith, N., Moremen, K. W., Xiang, Y., Siriwardena, A., Williams, S. J., Gilbert, H. J., and Davies, G. J. (2010). Mechanistic insights into a Ca<sup>2+</sup>-dependent family of alpha-mannosidases in a human gut symbiont. *Nat. Chem. Biol.* 6, 125–132.
- Zwart, G., Crump, B. C., Kamst-Van Agterveld, M. P., Hagen, F., and Han, S. K. (2002). Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat. Microb. Ecol.* 28, 141–155.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 12 January 2011; accepted: 14 April 2011; published online: 30 May 2011.
- Citation: Thomas F, Hehemann J-H, Rebuffet E, Czejek M and Michel G (2011) *Environmental and gut Bacteroidetes: the food connection.* *Front. Microbio.* 2:93. doi: 10.3389/fmicb.2011.00093
- This article was submitted to *Frontiers in Cellular and Infection Microbiology*, a specialty of *Frontiers in Microbiology*. Copyright © 2011 Thomas, Hehemann, Rebuffet, Czejek and Michel. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.

Chapitre 3  
Evaluation des gènes de référence  
pour la RT-PCR quantitative  
chez *Z. galactanivorans*

---



## **Présentation de l'article**

L'annotation du génome de *Z. galactanivorans* a mis en évidence un certain nombre de gènes potentiellement impliqués dans la dégradation de l'alginate. Dans le cadre de ce projet de thèse, l'étude transcriptomique de l'expression de ces différents gènes a été entreprise. Depuis son établissement, la technique de PCR quantitative couplée à la rétro-transcription (RT-qPCR) est la méthode de choix pour évaluer précisément les niveaux d'expression d'un lot de gènes d'intérêt dans diverses conditions. Cependant, la validité des résultats dépend de l'utilisation de gènes de normalisation préalablement testés. L'expression de tels gènes de normalisation (ou de référence) doit rester relativement stable quelle que soit la condition que l'on souhaite analyser. Il est désormais communément admis que l'évaluation des gènes de référence doit être réalisée au préalable pour chaque organisme modèle et dans les conditions à tester.

L'article reporté ci-après présente les résultats de l'évaluation des gènes de référence à utiliser lors d'expériences de RT-qPCR sur *Z. galactanivorans*. Il a été publié en 2011 dans la revue "*Journal of Microbiological Methods*". Les bactéries ont été cultivées dans 11 conditions de croissance différentes, présentant un intérêt pour les recherches conduites au laboratoire. Elles incluent des variations de la source de carbone employée (glucose, alginate, laminarine, agar, porphyrane), de la température d'incubation (10°C, 20°C, 44°C) et l'application de stress oxydants (H<sub>2</sub>O<sub>2</sub> et KI). Le lot de gènes à tester a été choisi pour couvrir différentes classes fonctionnelles, afin d'éviter les risques de co-régulation. Quatorze gènes candidats ont été testés, appartenant à 10 voies métaboliques : réplication, transcription, traduction, métabolisme des acides aminés, des nucléotides et de l'acide folique, cycle de Krebs, biosynthèse du peptidoglycane, de la flavine mononucléotide et des composés aromatiques.

Après mise au point des extractions d'ARN et de la technique de RT-qPCR pour *Z. galactanivorans*, les niveaux de transcription de ces quatorze gènes ont été mesurés dans les 11 conditions en triplicats. La stabilité de l'expression des gènes candidats a été évaluée en utilisant deux programmes différents, geNorm et Normfinder. Les quatre gènes dont l'expression reste la plus stable dans les conditions testées ont été identifiés comme étant *dnaG*, *gmkA*, *glyA* et *icdA*. Lors de la normalisation de données de RT-qPCR, il est

recommandé d'utiliser un "facteur de normalisation" égal à la moyenne géométrique de l'expression de plusieurs gènes de référence, plutôt que l'expression d'un seul gène. Le nombre de gènes à inclure dans le facteur de normalisation a été déterminé en utilisant le programme geNorm. Cette étude montre qu'un facteur incluant les niveaux de transcription des trois gènes *glyA*, *icdA* et *gmkA* est approprié.

L'étude présentée dans cet article fournit le premier lot de gènes de normalisation validés expérimentalement pour une espèce de la classe *Flavobacteriia*. Elle apporte une procédure fiable pour la quantification des niveaux de transcription des gènes chez *Z. galactanivorans*, et ouvre la voie à des études transcriptomiques ultérieures.





## Evaluation of reference genes for real-time quantitative PCR in the marine flavobacterium *Zobellia galactanivorans*

François Thomas, Tristan Barbeyron, Gurvan Michel \*

UPMC University Paris 6, UMR 7139 Marine Plants and Biomolecules, Station Biologique de Roscoff, F-29682 Roscoff, Bretagne, France  
CNRS, UMR 7139 Marine Plants and Biomolecules, Station Biologique de Roscoff, F-29682 Roscoff, Bretagne, France

### ARTICLE INFO

#### Article history:

Received 10 September 2010  
Received in revised form 19 October 2010  
Accepted 19 October 2010  
Available online 31 October 2010

#### Keywords:

Algae  
Marine Bacteroidetes  
Marine polysaccharide  
Reference genes  
RT-qPCR  
*Zobellia*

### ABSTRACT

The marine bacteria *Zobellia galactanivorans* is an emerging model microorganism for the bioconversion of algal polysaccharides. The sequence analysis of its genome opens the way to in-depth gene expression analysis, such as reverse transcription quantitative PCR (RT-qPCR) studies. The selection and validation of reference genes are a mandatory first step for the accurate quantification of transcripts. We selected fourteen candidate reference genes belonging to distinct pathways, namely replication, transcription, translation, citric acid cycle, amino acid, nucleotide and dihydrofolate metabolisms, and peptidoglycan, FMN and aromatic compounds synthesis. We quantified their expression by RT-qPCR in various culture conditions corresponding to different temperatures, carbon sources or stresses. The applications geNorm and Normfinder allowed ranking the genes according to their stability and gave concordant results. We found that the geometric average of the expression of *glyA*, *icdA* and *gmkA* can be confidently used to normalize the transcript abundance of genes of interest. In conclusion, this work provides a reliable procedure for gene expression analysis in the flavobacterium *Z. galactanivorans* and a validated set of reference genes to be used in future transcriptomics approaches. The strategy developed could also be the starting point for similar studies in other members of the Flavobacteria class.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

The Flavobacteria class includes organisms that have colonized many terrestrial and aquatic ecosystems, and several members are associated to human or animal diseases (Jooste and Hugo, 1999; Duchaud et al., 2007). Flavobacteria are especially abundant on marine phytodetrital aggregates (Eilers et al., 2001; Bauer et al., 2006), in the phycosphere of phytoplanktons (Gomez-Pereira et al., 2010) and on the surface of macroalgae (Staufenberger et al., 2008). Accumulating evidences indicate that these heterotrophic bacteria play a key role in the mineralization of marine organic matter, and particularly in the degradation of algal polysaccharides (Kirchman, 2002; Gomez-Pereira et al., 2010). However, polysaccharides from marine algae are significantly different from those of land plants (Michel et al., 2010a,b). Few enzymes involved in the bioconversion of genuine marine polysaccharides are currently known and most of them constitute new protein families (Colin et al., 2006; Michel et al., 2006; Flament et al., 2007; Guibet et al., 2007). The marine

flavobacterium *Zobellia galactanivorans* has been isolated from the red alga *Delesseria sanguinea* for its capacity to degrade agars and carrageenans (Barbeyron et al., 2001; Michel et al., 2006). The sequencing of its complete genome has confirmed that this marine bacterium displays a huge potential for the degradation of algal polysaccharides (Barbeyron et al., unpublished). Using this genomic data, we have already characterized new glycoside hydrolases specific for iota-carrageenans (Rebuffet et al., 2010) and porphyrans (Hehemann et al., 2010). Interestingly, such glycoside hydrolase genes from marine flavobacteria have been transferred to Japanese gut microbiota, due to the traditional consumption of seaweeds in Japan (Hehemann et al., 2010). Moreover, the sequencing of *Zobellia* genome paves the way for new functional approaches such as gene expression analysis, to better understand the physiology of the bacteria in general, and to decipher the pathways for the biodegradation of algal polysaccharides in particular.

Since its establishment (Higuchi et al., 1992), real-time quantitative RT-PCR (RT-qPCR) has become the method of choice for quantifying the level of transcription on a targeted set of interest genes in many different conditions (Wong and Medrano, 2005; Nolan et al., 2006). This fast and practical technique combines the advantages of a wide accurate dynamic range with high specificity and sensitivity (Bustin et al., 2009). In order to produce reliable quantitative data, normalization is mandatory to compensate for inherent variability due

Abbreviations:  $C_q$ , quantification cycle (also known as  $C_t$  for threshold cycle); RT-qPCR, quantitative real-time polymerase chain reaction; GR, growth rate.

\* Corresponding author. Tel.: +33 298 29 23 30; fax: +33 298 29 23 24.

E-mail address: [gurvan@sb-roscoff.fr](mailto:gurvan@sb-roscoff.fr) (G. Michel).



to RNA extractions, cDNA synthesis, or PCR efficiency. Many normalization strategies have been tested (Vandesompele et al., 2002; Andersen et al., 2004; Huggett et al., 2005) and the use of reference genes is commonly accepted as the most appropriate method, but requires prior validation for each experimental setup and model (Dheda et al., 2004; Bustin et al., 2009). Reports on the identification of stably expressed genes are now available for a broad range of species representing many different lineages. It includes human cells (Vandesompele et al., 2002; Lee et al., 2009), animals (Ahn et al., 2008; de Boer et al., 2009), plants (Barsalobres-Cavallari et al., 2009; Paolacci et al., 2009), brown algae (Le Bail et al., 2008) and bacteria (Theis et al., 2007; Nieto et al., 2009; Ritz et al., 2009).

To allow the development of transcriptomic studies in *Z. galactanivorans*, we evaluated fourteen candidate reference genes from distinct metabolic pathways. RT-qPCR was used to quantify their expression in various culture conditions corresponding to different temperatures, carbon sources or stresses. This work defines a reliable protocol and normalization factor that should be used in future gene expression analysis on *Z. galactanivorans*, and can initiate similar approaches on other Flavobacteria.

## 2. Materials and methods

### 2.1. Bacterial strain and culture conditions

The type strain Dsij<sup>T</sup> of *Z. galactanivorans* was grown in ZoBell medium 2216E (ZoBell, 1941) at 20 °C and subsequently cultured under agitation in triplicates for each of the eleven tested conditions (Table 1). Growth was followed by the absorbance at 600 nm, and the growth rate was defined as the number of generations per hour in the log phase. The “carbon sources” treatments were conducted in a Marine Mineral Medium (MMM) composed for 1 L of 24.7 g NaCl, 6.3 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 4.6 g MgCl<sub>2</sub>·H<sub>2</sub>O, 2 g NH<sub>4</sub>Cl, 0.7 g KCl, 0.6 g CaCl<sub>2</sub>, 200 mg NaHCO<sub>3</sub>, 100 mg K<sub>2</sub>HPO<sub>4</sub>, 50 mg yeast extract, and 20 mg FeSO<sub>4</sub>·7H<sub>2</sub>O. The different carbon sources were added to a final concentration of 2 g L<sup>-1</sup>. The final medium is buffered at 50 mM Tris–HCl pH 8.0. For the “temperature” treatments, bacteria were incubated in ZoBell medium at 10, 20 or 44 °C. For treatments with KI, bacteria were grown at 20 °C in ZoBell medium containing 100 mM KI. For treatments with H<sub>2</sub>O<sub>2</sub>, bacteria were grown at 20 °C until the exponential phase and H<sub>2</sub>O<sub>2</sub> was added to the medium to a final concentration of 500 μM for 1 h.

**Table 1**

Culture conditions and measured growth rate (n = 3). MMM: Marine Mineral Medium. Growth rate was defined as the number of generations per hour. For experiments with 1 hour incubations with hydrogen peroxide, growth was followed before H<sub>2</sub>O<sub>2</sub> addition and thus no specific growth rate are given for these two conditions. See Material and Methods for experimental details.

Type of treatment	Abbreviation	Growth rate (h <sup>-1</sup> ) mean ± s.e.m.
<i>Carbon source</i>		
MMM + Glucose 2 g L <sup>-1</sup>	G	0.132 ± 0.004
MMM + Alginate 2 g L <sup>-1</sup>	A	0.086 ± 0.008
MMM + Laminarin 2 g L <sup>-1</sup>	L	0.098 ± 0.008
MMM + Agar 2 g L <sup>-1</sup>	R	0.104 ± 0.002
MMM + Porphyran 2 g L <sup>-1</sup>	P	0.108 ± 0.006
<i>Temperature</i>		
Zobell, 10 °C	10 °C	0.050 ± 0.000
Zobell, 20 °C	20 °C	0.178 ± 0.003
Zobell, 44 °C	44 °C	0.209 ± 0.008
<i>Oxidative Stress</i>		
Zobell + KI 100 mM	KI	0.148 ± 0.003
Zobell + H <sub>2</sub> O <sub>2</sub> 500 μM	H <sub>2</sub> O <sub>2</sub>	-
Zobell + KI 100 mM + H <sub>2</sub> O <sub>2</sub> 500 μM	KI + H <sub>2</sub> O <sub>2</sub>	-

### 2.2. Nucleic acid extraction and cDNA synthesis

Genomic DNA was extracted as described by Barbeyron et al. (2008). RNA extractions were conducted as follows: 2 mL aliquots of bacterial cultures in mid- to late-log phase were placed in RNA protect reagent (Qiagen) and total RNA was isolated by using the RNA mini kit (Qiagen) following the manufacturer's instructions. RNA samples were treated with 4 Units of Turbo DNase (Ambion). Total elimination of genomic DNA was checked by quantitative PCR. RNA concentration was quantified on a Nanodrop ND-1000 spectrophotometer. The ratios OD<sub>260</sub>/OD<sub>280</sub> and OD<sub>230</sub>/OD<sub>260</sub> were calculated to assess the purity of RNA extracts. The integrity of RNA preparations was examined by 0.8% agarose gel electrophoresis. Reverse transcription reactions were performed on 800 ng total RNA using the Phusion RT-PCR Kit (Finnzymes) with random hexamer primers according to the supplier's instructions. cDNA samples were diluted to 1 ng μL<sup>-1</sup> and stored at -20 °C.

### 2.3. Primer design

For each of the 14 candidate genes, primers were designed using the Perl Primer open-source software (Marshall, 2004) with lengths between 17 and 22 nucleotides, T<sub>m</sub> between 58 and 60 °C and G/C content between 47 and 53%. The PCR products ranged from 60 to 120 bp and had a G/C content over 60%. Gene-specificity of all primers was confirmed using BLAST searches on the genomic sequence of *Z. galactanivorans* (Barbeyron et al., unpublished).

### 2.4. Real-time PCR and data analysis

Quantitative real-time PCR assay was performed in 96-well plates (ThermoScientific) on a Chromo4 real-time PCR Instrument (Bio-Rad, Hercules, CA). The composition of each reaction was as follows in 20 μL final volume: 10 ng of cDNA, primers (250 nM each) and 10 μL of Absolute Sybr Green ROX Mix (AbGene, Epson, UK). The reactions for each gene were technically duplicated. The program was: 15 min at 95 °C, followed by 40 cycles of 95 °C (15 s), 60 °C (30 s) and 72 °C (30 s). A single lot of cDNAs was used to quantify the 14 genes, to minimize the potential variability due to reverse transcription. Genomic DNA serial dilutions ranging from 10 to 10<sup>5</sup> copies were amplified by qPCR, in the same run than cDNA samples. Efficiencies were determined using the resulting standard curves, by the formula  $E = 10^{(-1/\text{slope})} - 1$ . Dissociation curves were obtained by heating the PCR products from 65 to 95 °C. Opticon 3 Monitor software (AbGene, Epson, UK) was used to obtain C<sub>q</sub> values and copy numbers for each reaction. Data were further analyzed using geNorm and Normfinder softwares (Vandesompele et al., 2002; Andersen et al., 2004).

## 3. Results and discussion

### 3.1. Treatments of *Z. galactanivorans* and RNA extraction

Various culture conditions were tested for *Z. galactanivorans* to assess the stability of candidate reference genes (Table 1). Carbohydrates used as sole carbon sources in minimal medium included glucose and four marine polysaccharides from brown algae (alginate and laminarin) and red algae (agar and porphyran). Bacteria were also incubated in rich medium (ZoBell medium) in various conditions: (i) At three different temperatures 10, 20 and 44 °C; (ii) addition of either iodine (100 mM KI) or hydrogen peroxide (500 μM H<sub>2</sub>O<sub>2</sub>), and with both compounds. These two substances are indeed produced as defence mechanisms by macroalgae (Cosse et al., 2007) and are susceptible to modify the physiology of this algae-associated bacterium. As expected, the growth of *Z. galactanivorans* varied according the culture conditions (Table 1), suggesting that variations of the gene expression may occur. The temperature had the largest

**Table 2**  
Candidate reference genes from *Zobellia galactanivorans*.

Gene	EMBL code	Product	Pathway
<i>recF</i>	FQ073855	DNA replication and repair protein	Replication/Repair
<i>dnaG</i>	FQ073966	DNA primase	Replication
<i>rho</i>	FQ073851	Transcription termination factor	Transcription
<i>rpoA</i>	FQ073850	DNA-dependent RNA polymerase, $\alpha$ subunit	Transcription
<i>rpoB</i>	FQ073837	DNA-dependent RNA polymerase, $\beta$ subunit	Transcription
<i>rrs</i>	AF208293	16 S ribosomal rRNA subunit	Translation
<i>rplD</i>	FQ073849	Ribosomal protein L4	Translation
<i>glyA</i>	FQ073857	Serine/Glycine hydroxymethyltransferase A	Amino acid catabolism
<i>gmkA</i>	FQ073846	Guanylate kinase	Nucleotide metabolism
<i>dfrA</i>	FQ073861	Dihydrofolate reductase	Dihydrofolate metabolism
<i>murG</i>	FQ073965	Undecaprenyldiphospho-muramoylpentapeptide beta-N- acetylglucosaminyltransferase	Peptidoglycan synthesis
<i>icdA</i>	FQ073844	NADP-dependent isocitrate dehydrogenase	Citric acid cycle
<i>ribF</i>	FQ073964	Riboflavin biosynthesis protein	FMN and FAD synthesis
<i>aroE</i>	FQ073854	Shikimate dehydrogenase	Aromatic compounds synthesis

effect, with the growth rate (GR) four times higher at 44 °C ( $0.21 \text{ h}^{-1}$ ) than at 10 °C ( $0.05 \text{ h}^{-1}$ ). In minimal medium, the growth rate is somewhat lower with complex polysaccharides (Pol) than with glucose (Glc), with a ratio  $\text{GR}_{\text{Pol}}/\text{GR}_{\text{Glc}}$  ranging from 65% for alginate to 82% for porphyran. Finally, *Z. galactanivorans* tolerates quite well the presence of high concentration of iodine (100 mM KI), the GR being only decreased by ~20% compared to cultures at 20 °C in Zobell medium. Due to the experimental procedure (1 h incubation with  $\text{H}_2\text{O}_2$ ), the effect of hydrogen peroxide on GR cannot be measured (see [Materials and methods](#)). Total RNA was purified from mid- to late-exponential growth phase cultures in the aforementioned conditions, treated with DNase, and analysed for concentration, purity and integrity (see [Materials and methods](#) and Additional files 1 and 2). The extraction protocol led to non-degraded RNA extracts ranging from 235 to 1200  $\text{ng } \mu\text{l}^{-1}$ , with  $A_{260}/A_{280}$  ratios between 2.0 and 2.3 indicating no detectable protein contamination.

### 3.2. Choice of reference genes and primer validation

The candidate reference genes were chosen from various functional classes, in order to minimize the risk of co-regulation. Fourteen genes belonging to ten different pathways were selected ([Table 2](#)) and primers were designed for each gene ([Table 3](#)). All these gene sequences have been deposited to the EMBL database ([Table 2](#)). To evaluate the PCR efficiency of each primer pair, 10-fold serial dilutions of genomic DNA were amplified by qPCR, in the same run than cDNA samples. Efficiencies were determined using the resulting standard curves, by the formula  $E = 10^{(-1/\text{slope})} - 1$ , and were all above 80%. The amount of transcripts of each candidate reference gene in the cDNA samples were calculated using the appropriate standard curve. The specificity of amplification was verified for each primer pair by the detection of a single peak on melting curves from 65 °C to 95 °C (data not shown). The repeatability of the method (or intra-assay

**Table 3**  
List of primers used for qPCR validation.

Gene	Primer sequences (5'-3') Forward - Reverse	Product length (bp)	Calibration curve	E (%)	Intra-assay CV (%)
<i>recF</i>	CAGTATGTCGCTATTTTCAGG CCAACAAGCCTTGATCC	79	$y = -3.36x + 35.41$ $R^2 = 0.999$	98	3.24
<i>dnaG</i>	CGAGGAGAAATCACAAGAC CGCTTTATAGGATCGTTGG	77	$y = -3.43x + 36.22$ $R^2 = 0.999$	96	3.20
<i>rho</i>	CCTCTTCAAGTACAAGACG CAACAGGATTCATGTCCG	99	$y = -3.79x + 36.89$ $R^2 = 0.997$	84	3.29
<i>rpoA</i>	GGAACGCACTAAGAAGAG CATCTTCTACAACGCTCG	114	$y = -3.39x + 33.99$ $R^2 = 1.000$	97	5.42
<i>rpoB</i>	CTTCTCGTATGAACATCCG CCCAATTAAGACCTGCC	64	$y = -3.38x + 34.06$ $R^2 = 0.999$	98	10.39
<i>rrs</i>	CCCGTCAATTCCTTTGAG ATGGATACTAGCTGTCCG	109	$y = -3.93x + 39.57$ $R^2 = 0.978$	80	6.34
<i>rplD</i>	GGATTCCGAAGTCATAACC CCAATCCTTCTAAGGCAC	95	$y = -3.33x + 33.35$ $R^2 = 0.999$	100	7.49
<i>glyA</i>	GCTTTGATCTATCCCACG CATCGTAATCAAGGGTCC	120	$y = -3.35x + 35.55$ $R^2 = 1.000$	99	3.52
<i>gmkA</i>	TATGCGAATAGCGAAGG AAATCAGCCCAACTC	117	$y = -3.36x + 34.30$ $R^2 = 1.000$	99	10.31
<i>dfrA</i>	CATCATAGCCTTGATAGCG CGGTAATGCCATAGTAGG	76	$y = -3.45x + 35.47$ $R^2 = 1.000$	95	7.02
<i>murG</i>	GGCCTATGATGAAATGGAC GTCATTCAACCAAGTCTC	87	$y = -3.58x + 40.46$ $R^2 = 0.998$	90	5.35
<i>icdA</i>	GTAAGGCTTACCTTCCC CGGTAGCTTCTTGGTG	73	$y = -3.44x + 36.35$ $R^2 = 1.000$	95	5.45
<i>ribF</i>	CTACCGTAGAAGGAAAGGA GGTGTAAGATGTCGATCTG	97	$y = -3.33x + 33.34$ $R^2 = 0.997$	100	7.88
<i>aroE</i>	AGATCAGTTACGACCAGG GGCGTTCCTCAATATTGG	99	$y = -3.32x + 36.86$ $R^2 = 1.000$	100	5.65

variation) was evaluated by running technical duplicates of each sample on the same plate and measuring the coefficient of variation (CV) of the number of copies of each gene. The low overall means of intra-assay CVs, between 3.0 and 10.5% (Table 3), denote a good precision and robustness of the assay.

### 3.3. Transcript levels of candidate reference genes

Transcript levels were measured by RT-qPCR for each gene in the 11 conditions tested in biological triplicates. Fig. 1 shows that, except for *rrs* which encodes the 16S rRNA, mean  $C_q$  values were comprised between 16.4 and 25.2. Standard deviations of transcripts quantification across the 11 conditions were all under 2.2  $C_q$ , with *dfrA* and *murG* exhibiting the lowest and the largest variations, respectively.  $C_q$  values were converted to transcript numbers using the standard curve generated for each gene (Fig. 1B). The transcripts of *rplD* and *rpoA* were more numerous, whereas transcripts of *dnaG* and *recF* were the less abundant. Due to the abundance of the *rrs* transcripts, cDNA samples had to be diluted  $10^3$ -fold before qPCR in order to obtain a mean  $C_q$  value within the range of the standard curve ( $19.3 \pm 1.6$ ). Even though *rrs* expression was relatively stable in the various culture conditions, it is not representative of lower mRNA quantities. The use of *rrs* as a reference gene is therefore not recommended, as concluded in other bacterial models (Theis et al., 2007; Nieto et al., 2009; Ritz et al., 2009). The *rrs* gene was not further considered for the determination of the optimal reference genes.

### 3.4. Determination of the optimal reference genes

Two complementary programs were used to rank the candidate reference genes for their stability over the 11 tested culture conditions: geNorm (Vandesompele et al., 2002) and Normfinder (Andersen et al., 2004). Although the two softwares use different algorithms to calculate a stability value for each gene, a very good correlation was obtained between the results provided by geNorm and Normfinder (Fig. 2). *recF*, *dfrA* and *rho* were the least stable genes according to both methods. The four most stable genes were the same with geNorm and Normfinder, namely *dnaG*, *gmkA*, *glyA* and *icdA*, although the two approaches did not rank them in the same order.

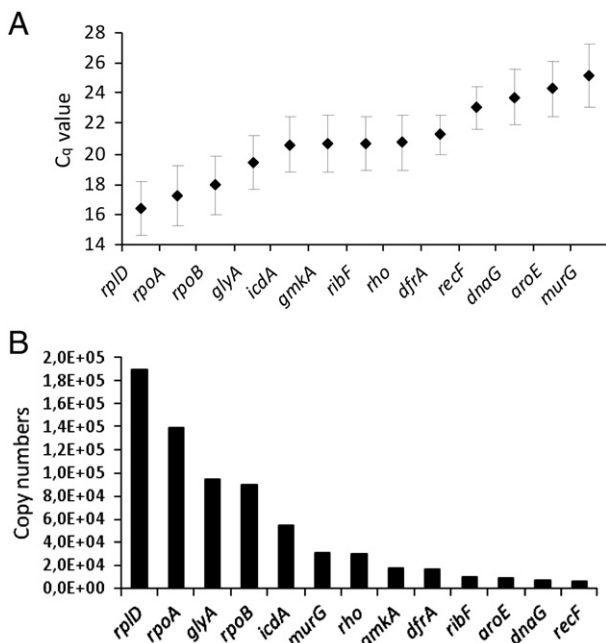


Fig. 1. A.  $C_q$  values for the reference genes. Expression levels of the reference genes over the 11 conditions (mean  $\pm$  SD). B. Mean copy numbers for the reference genes.

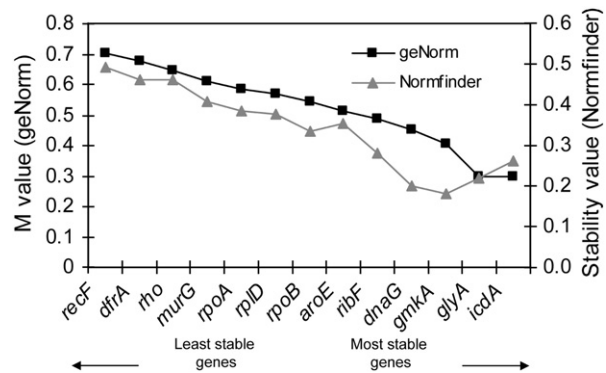


Fig. 2. Ranking of the stability of the reference genes. Analyses were conducted both with geNorm and Normfinder softwares. A low stability value denotes that the gene is stably expressed over the 11 experimental conditions.

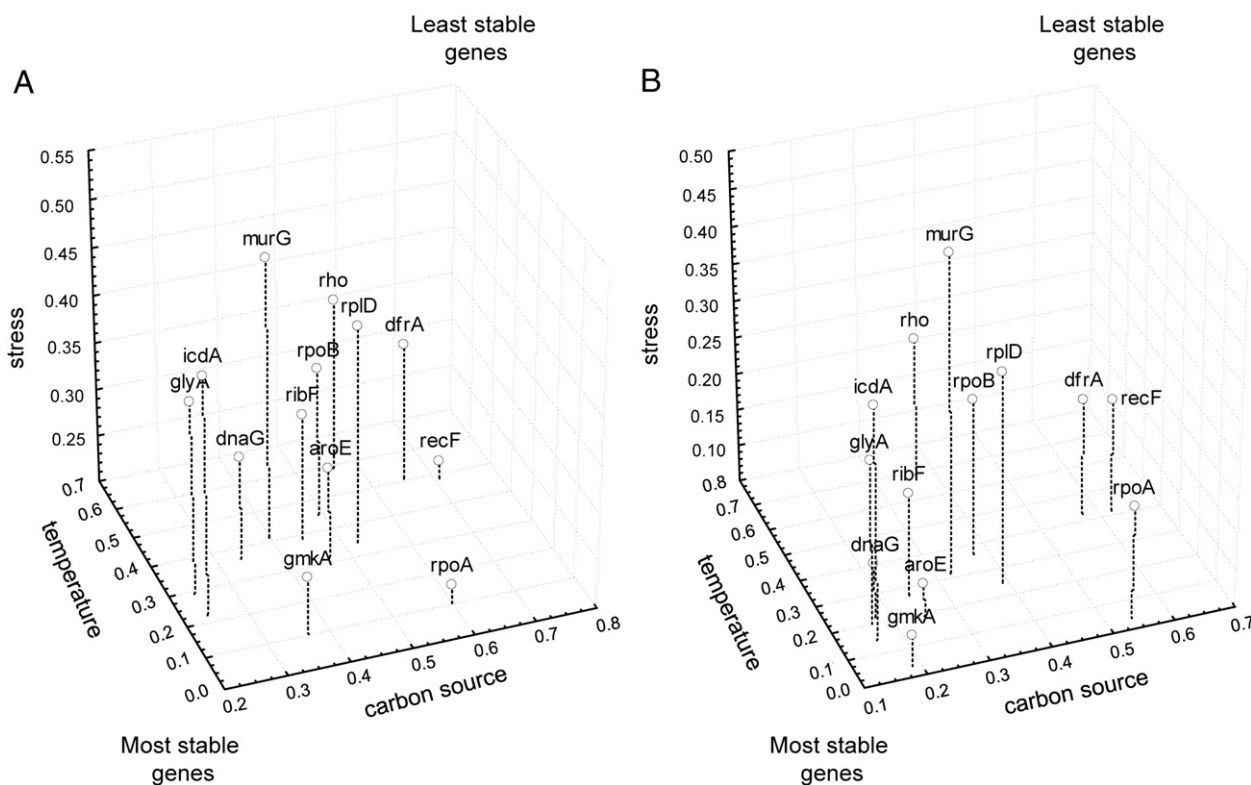
The aforementioned evaluation of the candidate genes was made by combining data from the eleven conditions tested in this study. However, if the experimental model changes, so does the best gene combination. GeNorm and Normfinder analyses were performed on the three distinct series of treatments (temperature, carbon source and oxidative stress) and the calculated stability values were plotted on 3-D graphs (Fig. 3). *gmkA* remained among the most stable genes whatever the type of treatment. However, this approach demonstrated that the nature of the tested parameters influenced the rank of the reference genes. For instance, the transcript abundance of *recF* and *rpoA* greatly varied in the “carbon source” treatments, but these genes were among the best references for the oxidative stress experiments.

### 3.5. Determination of the optimal number of reference genes

The geometric averaging of multiple reference genes is highly recommended to normalize the expression of interest genes (Vandesompele et al., 2002; Bustin et al., 2009). The geNorm software enables the determination of the optimal number of reference genes to be included in a RT-qPCR experiment. This program calculates the pairwise variation ( $V_{n/n+1}$ ) between two sequential normalization factors ( $NF_n$  and  $NF_{n+1}$ ). Vandesompele et al. (2002) recommended the minimal use of the three most stable reference genes for calculation of a normalization factor, and the inclusion of additional genes until  $V_{n/n+1}$  is lower than 0.15. According to our data (Table 4), the normalization factors are only slightly modified when including more than three genes ( $V_{3/4} = 0.109$ ). Thus, a factor including transcript levels of *glyA*, *icdA* and *gmkA* is sufficient to normalize the expression of genes of interest in these samples by RT-qPCR.

### 3.6. Use of the validated factor to normalize the expression of a target gene

To assess the ability of this method to detect changes in transcripts abundance, we used the validated factor to normalize the expression of a target gene. We selected *alyA1*, a gene encoding a putative alginate lyase, and predicted its expression to be higher in the presence of alginate. We quantified the expression of *alyA1* in *Z. galactanivorans* with either glucose or alginate as a sole carbon source, using specific primers (forward primer sequence: ATCACCGAATGTAGCCA; reverse primer sequence: CCATCACCATAACCTGAC). Transcripts numbers of *alyA1* were normalized using the average mean of *glyA*, *icdA* and *gmkA* values. In the presence of glucose, the relative expression of *alyA1* was  $0.12 \pm 0.02$ . When cells were grown with alginate, *alyA1* expression increased 8 fold, reaching  $1.01 \pm 0.10$ . Thus, the use of the normalization factor defined in this study is suitable to detect changes in the gene expression.



**Fig. 3.** Ranking of the stability of the reference genes over the three different sets of parameters. Genes were ordered both with geNorm (A) and Normfinder (B) and plotted on three axes corresponding to the three types of culture conditions tested. The position of each gene on the 3-D graph indicates its suitability as a reference gene, the most stable ones being at the front bottom position.

#### 4. Conclusion

Normalization is a key step in gene expression studies. As no universal reference gene exists, the normalization strategy should be validated before use in a particular experimental setup. The aim of this study was to develop a reliable procedure for gene expression analysis in the marine flavobacterium *Z. galactanivorans*. Using RT-qPCR, the expression of fourteen candidate reference genes was measured in diverse culture conditions. The two approaches, geNorm and Normfinder, gave consistent results and identified *dnaG*, *gmkA*, *glyA* and *icdA* as the genes with the most stable expression throughout our tested conditions. Furthermore, a pairwise analysis indicated that three genes were sufficient to calculate a valid normalization factor. Therefore, the geometric average of the expression of *glyA*, *icdA* and *gmkA* can be confidently used to normalize the transcript abundance of target-genes. We applied this procedure to quantify the expression of a selected target gene and showed that it is suitable to detect changes in transcripts abundance. We also show that the choice of the reference genes has to be adapted if only one distinct type of

treatment is considered. To our knowledge, this study provides the first validated set of reference genes for a species belonging to the Flavobacteria class.

#### Acknowledgements

We are grateful to Thierry Tonon and Mirjam Czjzek to helpful discussion on this study. We also thank Drs Carole Dossat and Valérie Barbe (Génoscope, France) for the sequence deposition at the EMBL database. The 'Marine Plants and Biomolecules' laboratory is funded by the French National Scientific Research Centre (CNRS) and the University Pierre and Marie Curie (UPMC). FT is supported by a PhD grant from the UPMC. This work was also supported by the European Community within the Seventh Framework Programme under grant agreement n° 222628 (Large collaborative project PolyMode, <http://www.polymode.eu/>).

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at [doi:10.1016/j.mimet.2010.10.016](https://doi.org/10.1016/j.mimet.2010.10.016).

#### References

- Ahn, K., Huh, J.W., Park, S.J., Kim, D.S., Ha, H.S., Kim, Y.J., et al., 2008. Selection of internal reference genes for SYBR green qRT-PCR studies of rhesus monkey (*Macaca mulatta*) tissues. *BMC Mol. Biol.* 9, 78.
- Andersen, C.L., Jensen, J.L., Orntoft, T.F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64, 5245–5250.
- Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., Potin, P., 2001. *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of. *Int. J. Syst. Evol. Microbiol.* 51, 985–997.
- Barbeyron, T., L'Haridon, S., Michel, G., Czjzek, M., 2008. *Mariniflexile fucanivorans* sp. nov., a marine member of the *Flavobacteriaceae* that degrades sulphated fucans from brown algae. *Int. J. Syst. Evol. Microbiol.* 58, 2107–2113.

**Table 4**

Pairwise variations (V) of the sequential normalization factors. V was calculated using geNorm and represents the pairwise variation between two sequential normalization factors containing an increasing number of genes.

Sequential normalization factors	Pairwise variation (V)
V3/4	0.109
V4/5	0.094
V5/6	0.080
V6/7	0.077
V7/8	0.067
V8/9	0.059
V9/10	0.060
V10/11	0.063
V11/12	0.059
V12/13	0.056



- Barsalobres-Cavallari, C.F., Severino, F.E., Maluf, M.P., Maia, I.G., 2009. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. *BMC Mol. Biol.* 10, 1.
- Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., et al., 2006. Whole genome analysis of the marine *Bacteroidetes* 'Gramella forsetii' reveals adaptations to degradation of polymeric organic matter. *Environ. Microbiol.* 8, 2201–2213.
- Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., et al., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622.
- Colin, S., Deniaud, E., Jam, M., Descamps, V., Chevolut, Y., Kervarec, N., et al., 2006. Cloning and biochemical characterization of the fucanase FcnA: definition of a novel glycoside hydrolase family specific for sulfated fucans. *Glycobiology* 16, 1021–1032.
- Cosse, A., Leblanc, C., Potin, P., 2007. Dynamic defense of marine macroalgae against pathogens: from early activated to gene-regulated responses. *Adv. Bot. Res.* 46, 221–266.
- de Boer, M.E., de Boer, T.E., Marien, J., Timmermans, M.J., Nota, B., van Straalen, N.M., et al., 2009. Reference genes for QRT-PCR tested under various stress conditions in *Folsomia candida* and *Orchesella cincta* (Insecta, Collembola). *BMC Mol. Biol.* 10, 54.
- Dheda, K., Huggett, J.F., Bustin, S.A., Johnson, M.A., Rook, G., Zumla, A., 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37, 112–119.
- Duchaud, E., Boussaha, M., Loux, V., Bernardet, J.F., Michel, C., Kerouault, B., et al., 2007. Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*. *Nat. Biotechnol.* 25, 763–769.
- Eilers, H., Pernthaler, J., Peplies, J., Glockner, F.O., Gerdt, G., Amann, R., 2001. Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* 67, 5134–5142.
- Flament, D., Barbeyron, T., Jam, M., Potin, P., Czjzek, M., Kloareg, B., et al., 2007. Alpha-agarases define a new family of glycoside hydrolases, distinct from beta-agarase families. *Appl. Environ. Microbiol.* 73, 4691–4694.
- Gomez-Pereira, P.R., Fuchs, B.M., Alonso, C., Oliver, M.J., van Beusekom, J.E., Amann, R., 2010. Distinct flavobacterial communities in contrasting water masses of the north Atlantic Ocean. *ISME J.* 4, 472–487.
- Guibet, M., Colin, S., Barbeyron, T., Genicot, S., Kloareg, B., Michel, G., et al., 2007. Degradation of lambda-carrageenan by *Pseudoalteromonas carrageenovora* lambda-carrageenase: a new family of glycoside hydrolases unrelated to kappa- and iota-carrageenases. *Biochem. J.* 404, 105–114.
- Hehemann, J.H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., Michel, G., 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* 464, 908–912.
- Higuchi, R., Dollinger, G., Walsh, P.S., Griffith, R., 1992. Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* 10, 413–417.
- Huggett, J., Dheda, K., Bustin, S., Zumla, A., 2005. Real-time RT-PCR normalisation: strategies and considerations. *Genes Immun.* 6, 279–284.
- Jooste, P.J., Hugo, C.J., 1999. The taxonomy, ecology and cultivation of bacterial genera belonging to the family Flavobacteriaceae. *Int. J. Food Microbiol.* 53, 81–94.
- Kirchman, D.L., 2002. The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol. Ecol.* 39, 91–100.
- Le Bail, A., Dittami, S.M., de Franco, P.O., Rousvoal, S., Cock, M.J., Tonon, T., et al., 2008. Normalisation genes for expression analyses in the brown alga model *Ectocarpus siliculosus*. *BMC Mol. Biol.* 9, 75.
- Lee, K.S., Alvarenga, T.A., Guindalini, C., Andersen, M.L., Castro, R.M., Tufik, S., 2009. Validation of commonly used reference genes for sleep-related gene expression studies. *BMC Mol. Biol.* 10, 45.
- Marshall, O.J., 2004. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* 20, 2471–2472.
- Michel, G., Nyval-Collen, P., Barbeyron, T., Czjzek, M., Helbert, W., 2006. Bioconversion of red seaweed galactans: a focus on bacterial agarases and carrageenases. *Appl. Microbiol. Biotechnol.* 71, 23–33.
- Michel, G., Tonon, T., Scornet, D., Cock, J.M., Kloareg, B., 2010a. Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in Eukaryotes. *New Phytol.* 188, 67–81.
- Michel, G., Tonon, T., Scornet, D., Cock, J.M., Kloareg, B., 2010b. The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytol.* 188, 82–97.
- Nieto, P.A., Covarrubias, P.C., Jedlicki, E., Holmes, D.S., Quatrini, R., 2009. Selection and evaluation of reference genes for improved interrogation of microbial transcriptomes: case study with the extremophile *Acidithiobacillus ferrooxidans*. *BMC Mol. Biol.* 10, 63.
- Nolan, T., Hands, R.E., Bustin, S.A., 2006. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* 1, 1559–1582.
- Paolacci, A.R., Tanzarella, O.A., Porceddu, E., Ciaffi, M., 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Mol. Biol.* 10, 11.
- Rebuffet, E., Barbeyron, T., Jeudy, A., Jam, M., Czjzek, M., Michel, G., 2010. Identification of catalytic residues and mechanistic analysis of family GH82 iota-carrageenases. *Biochemistry* 49, 7590–7599.
- Ritz, M., Garenaux, A., Berge, M., Federighi, M., 2009. Determination of rpoA as the most suitable internal control to study stress response in *C. jejuni* by RT-qPCR and application to oxidative stress. *J. Microbiol. Meth.* 76, 196–200.
- Staufenberger, T., Thiel, V., Wiese, J., Imhoff, J.F., 2008. Phylogenetic analysis of bacteria associated with *Laminaria saccharina*. *FEMS Microbiol. Ecol.* 64, 65–77.
- Theis, T., Skurray, R.A., Brown, M.H., 2007. Identification of suitable internal controls to study expression of a *Staphylococcus aureus* multidrug resistance system by quantitative real-time PCR. *J. Microbiol. Meth.* 70, 355–362.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3 RESEARCH0034.
- Wong, M.L., Medrano, J.F., 2005. Real-time PCR for mRNA quantitation. *Biotechniques* 39, 75–85.
- ZoBell, C., 1941. Studies on marine bacteria I The cultural requirements of heterotrophic aerobes. *J. Mar. Res.* 4, 42–75.

Chapitre 4  
Etude du système alginolytique  
de *Z. galactanivorans*

---



## **Présentation de l'article**

*Zobellia galactanivorans* est capable d'utiliser l'alginate comme seule source de carbone, suggérant qu'elle possède une voie complète de dégradation de ce polysaccharide. L'article reporté ci-après présente l'identification des gènes impliqués dans cette voie, constituant un véritable système alginolytique.

L'analyse du génome de *Z. galactanivorans* a permis le recensement de gènes potentiellement impliqués dans la dégradation de l'alginate. *Z. galactanivorans* possède 7 alginate lyases putatives (AlyA1-AlyA7) appartenant à 4 familles CAZy différentes. Par ailleurs, elle possède un gène codant une enzyme de la famille PL15, où la seule activité caractérisée à ce jour est "oligo-alginate lyase". En analysant le contexte génomique des alginate lyases prédites, j'ai pu identifier deux clusters de gènes dont la fonction pourrait être reliée au catabolisme de l'alginate. Le premier cluster de 6 kb regroupe les trois gènes d'alginate lyases *alyA4*, *alyA5* et *alyA6*. Le deuxième cluster de 15,9 kb regroupe 12 gènes dont la fonction est reliée au métabolisme des sucres, tels qu'une kinase, un facteur de transcription, un système SusC/SusD-like, etc. Des analyses en RT-PCR et en *Northern Blot* ont montré que ces deux clusters peuvent être transcrits sous forme d'ARN polycistroniques en présence d'alginate. Ils constituent ainsi les premiers opérons alginolytiques caractérisés. Cette organisation rappelle celle des PUL retrouvés dans les génomes de *Bacteroidetes* et qui pourraient expliquer leur spécialisation en tant que dégradeur de matière organique polymérique (cf. Chapitre 2).

En appliquant la méthode de RT-qPCR développée et présentée dans le Chapitre 3, les niveaux d'expression des gènes potentiellement impliqués dans la dégradation de l'alginate ont été quantifiés pour des cultures en présence de glucose ou d'alginate comme seule source de carbone. Les résultats montrent que l'expression de la plupart des gènes est fortement induite en présence d'alginate. Ce effet au niveau transcriptomique est corrélé avec l'augmentation de l'activité enzymatique alginate lyase chez des bactéries cultivées en présence d'alginate. Ces différents résultats ont permis de bâtir un modèle hypothétique de régulation du système alginolytique chez *Z. galactanivorans*.

Par ailleurs, j'ai cherché à caractériser la distribution environnementale des gènes du système alginolytique, en recherchant des orthologues dans les bases de données génomiques



publiques. Des clusters synténiques ont été détecté dans d'autres *Flavobacteriaceae* marines, ainsi que chez des *Alphaproteobacteria*, *Gammaproteobacteria* et 5 espèces de *Bacteroides*. Des approches de génomique comparative et de phylogénie ont révélé que le système alginolytique identifié provient d'une flavobactérie marine ancestrale. Il a été acquis par des membres d'autres phyla par plusieurs événements indépendants de transferts latéraux de gènes (LGT). Ces transferts ne sont pas restreints au milieu marin. De manière surprenante, des événements de LGT ont eu lieu entre une flavobactérie marine et des *Bacteroides* du microbiome intestinal de personnes asiatiques. Ce type de transfert a probablement été favorisé par la consommation traditionnelle d'algues brunes dans la cuisine asiatique, sous forme de kombu ou de wakamé. Le contact entre une bactérie marine ingérée en même temps que l'algue et un microbionte intestinal aurait permis l'échange de gènes. L'acquisition de cette nouvelle capacité métabolique représenterait un avantage sélectif pour ces bactéries du tube digestif, leur permettant d'utiliser l'alginate ingéré comme source de carbone.

## **Characterization of the first Alginate Utilization System:**

**From its emergence in marine *Flavobacteria* to its independent transfers to marine  
*Proteobacteria* and human gut *Bacteroides***

**François Thomas<sup>1,2</sup>, Tristan Barbeyron<sup>1,2</sup>, Thierry Tonon<sup>1,2</sup>, Mirjam Czjzek<sup>1,2</sup> and  
Gurvan Michel<sup>1,2\*</sup>**

<sup>1</sup>UPMC University Paris 6, UMR 7139 Marine Plants and Biomolecules, Station Biologique  
de Roscoff, F-29682 Roscoff, Brittany, France

<sup>2</sup>CNRS, UMR 7139 Marine Plants and Biomolecules, Station Biologique de Roscoff, F-29682  
Roscoff, Brittany, France

\*Corresponding author:

E-mail address: [gurvan@sb-roscoff.fr](mailto:gurvan@sb-roscoff.fr); Tel (33) 298 29 23 30; Fax (33) 298 29 23 24

Running title: Emergence and transfer of alginolytic operons

Keywords: alginate lyase, operon, marine bacteria, gut bacteria, lateral gene transfer,  
comparative genomic

## SUMMARY

Marine heterotrophic bacteria play a crucial role in the carbon cycle by recycling algal biomass which is mainly composed of recalcitrant polysaccharides. However, the catabolism of algal polysaccharides remains largely unknown, impairing the correct interpretation of marine (meta)genomic data. Using an integrative approach, we here decipher in *Zobellia galactanivorans* the mechanisms of perception, uptake and utilization of alginate. Notably, this marine flavobacterium possesses 7 alginate lyase genes, 5 of them localized within two clusters comprising additional carbohydrate-related genes. The transcription of these genes and the alginolytic activity were strongly induced when *Z. galactanivorans* used alginate as sole carbon source. Using RT-PCR and Northern blot, these clusters were shown to be transcribed as polycistronic mRNAs and thus to constitute operons. Similar alginolytic operons were found in ~40 bacteria belonging to the *Bacteroidetes* and *Proteobacteria* phyla, including a majority of marine bacteria but also Japanese gut *Bacteroides*. In-depth phylogenomic analyses reveal that these alginolytic operons originated from an ancestral marine flavobacterium and were independently transferred to marine proteobacteria and gut *Bacteroides*. These bacteria thus gained the capacity to assimilate the main polysaccharide of brown algae, an adaptive advantage in the coastal environment but also in the gut microbiota of specific human population.

## Introduction

Coastal areas are the most productive marine ecosystems, accounting for 20% of the global carbon production. These environments host dense communities of macrophytes (macroalgae and seagrasses) which dominate the primary production. The world-wide surface of these macrophyte beds is estimated as  $2.10^6$  km<sup>2</sup> (Whittaker and Likens, 1973). Due to their slow turnover time (ca. 1 year) and the large biomass they represent, macrophytes act as a carbon sink in marine environments (Smith, 1981). Polysaccharides account in average for half of the algal tissues, where they are found either as cell wall (Popper et al., 2011) or carbon storage compounds (Michel et al., 2010b). By degrading this organic matter that would otherwise be lost for the food web, marine heterotrophic bacteria play a key role in the global carbon cycle. Bacterial degradation processes increase the nutritional value of algal tissues and allow the dissolution of organic matter into the water column, where it can be ingested by filter feeders and other animals (Norderhaug et al., 2003; Azam and Malfatti, 2007).

To help understanding the functional role of marine microbes in oceanic ecosystems, genome and metagenome sequencing projects have been launched ten years ago. These efforts have led to the massive release of new genomic data, which now need to be analyzed to infer useful knowledge (Glöckner and Joint, 2010). However, compared to the degradation of terrestrial plant biomass, the metabolic pathways involved in the degradation of marine polysaccharides are largely unknown. This hampers the annotation of new genome sequences being released. Thus, extensive studies on algal biomass utilization by bacteria are required to (i) help assigning functions to the genes found in marine bacteria and (ii) assess their biological roles in the marine carbon cycle.

Among marine heterotrophic bacteria, members of the *Bacteroidetes* phylum are increasingly recognized as specialists for the degradation of polysaccharides (Kirchman, 2002; Thomas et al., 2011b). Within this phylum, the flavobacterium *Zobellia galactanivorans* is being developed in our group as a model organism for the degradation of algal biomass. *Z. galactanivorans* was isolated in Roscoff (Brittany, France) from the red alga *Delesseria sanguinea* and utilizes a wide range of algal polysaccharides as a sole carbon source (Barbeyron et al., 2001; Michel et al., 2006; Thomas et al., 2011b). This suggests the existence of complete utilization pathways for these complex polymers. The recent sequencing of *Z. galactanivorans* genome confirmed this huge potential for polysaccharide bioconversion, with the presence of more than one hundred Carbohydrate Active Enzymes (CAZymes) and 72 sulfatases (Barbeyron et al, in preparation). Interestingly, it revealed that most of these genes co-localized within clusters, together with genes encoding transcriptional factors and homologs of the SusCD systems, known for their role in starch utilization in *Bacteroides* species (Martens et al., 2009). These gene clusters are thus potential operons specific for the utilization of algal polysaccharides.

Alginate is one of the polysaccharides from seaweeds which can be used by *Z. galactanivorans* as sole carbon source. This linear polysaccharide consists of two uronic monosaccharides, the  $\beta$ -D-mannuronic acid and its C5 epimer the  $\alpha$ -L-guluronic acid, arranged in blocks along the polysaccharide chain. Alginates form gels in the presence of divalent cations and their self-assembling properties depend on the relative proportion of guluronate homopolymeric blocks (Haug et al., 1974). These gelling properties find various industrial applications for alginate as a texturizing agent, including in food, cosmetics, textile printing and drug delivery. Alginate is the main component of the cell wall in brown algae, where it can account for 10 to 45% of the dry weight (Kloareg and Mabeau, 1987). This polymer thus represents an abundant source of carbon for consumers in coastal regions. Some

terrestrial or pathogenic bacteria of the family *Pseudomonadaceae* and *Azotobacteriaceae* also produce an acetylated type of alginate as an exopolysaccharide (Ramsey and Wozniak, 2005). The genomic analysis of the brown alga *Ectocarpus siliculosus* (Cock et al., 2010) has recently revealed that brown algae acquired the capacity to produce alginate by lateral gene transfer from an ancestral *Actinobacteria* (Michel et al., 2010a). Alginate, as a polysaccharide of bacterial origin, has then further evolved in brown algae in order to interact with the other cell wall components, cellulose, sulfated fucans and polyphenols (Popper et al., 2011).

The enzymatic pathway for the complete degradation of bacterial alginate and its assimilation into the central metabolism has been studied in the early 1960's in *Pseudomonas* spp. (Preiss and Ashwell, 1962b, a). It has since been found in other species such as *Sphingomonas* spp. and *Agrobacterium* spp. (Ochiai et al., 2006; Hashimoto et al., 2010). The first step in alginate catabolism is the depolymerization of the polysaccharide by alginate lyases, releasing unsaturated oligosaccharides which are further degraded down to the monosaccharide 4-deoxy-L-erythro-5-hexoseulose uronic acid (DEH). Dehydrogenases then convert DEH into 2-keto-3-deoxy-D-gluconate, further phosphorylated by a carbohydrate kinase. The final product is 2-keto-3-deoxy-6-phosphogluconate and can directly be assimilated through the Entner-Doudoroff pathway. To date, this complete pathway for alginate degradation has only been studied in a few terrestrial or pathogenic species, which degrade bacterial alginates. Several alginate lyases have also been characterized in various marine bacteria (Lombard et al., 2010). However, systems for the perception and complete degradation of algal alginate by marine bacteria have never been reported in the literature.

In the present study we report the molecular and transcriptomic characterization of the complex alginolytic system of *Z. galactanivorans*, demonstrating its organization into two genuine alginolytic operons. In addition, we provide phylogenomic evidences that this

alginate utilization system (AUS) emerged in an ancestral marine flavobacterium and was transferred through independent lateral gene transfers (LGT) to marine *Proteobacteria*, but also to several Japanese gut *Bacteroides*, due to the traditional consumption of brown algae in Japan.

## RESULTS

### Census of genes potentially involved in alginate degradation.

Seven genes encoding alginate lyases were predicted in the genome of *Z. galactanivorans* (*alyA1* - *alyA7*) (Figure 1). These new alginate lyases belong to four different CAZy families (Lombard et al., 2010). AlyA1, AlyA2 and AlyA5 are classified in family 7 of the polysaccharide lyases (PL7), where known activities are alginate lyase (EC 4.2.2.3) and  $\alpha$ -L-gulonate lyase (EC 4.2.2.11). Their sequences contain a cleaved lipoprotein signal, predicting localization in the outer membrane. AlyA2 and AlyA5 have similar molecular masses (35 and 39 kDa, respectively), whereas AlyA1 is larger due to the presence of a carbohydrate binding module (CBM) belonging to family 32. A similar CBM32 is present in a characterized alginate lyase, AlyPI from *Pseudoalteromonas* sp. CY24 (Duan et al., 2009), suggesting that this CBM binds alginate and potentiates the enzymatic activity of the PL7 module. AlyA3 belongs to the PL17 family, in which the only known activity is alginate lyase. This enzyme possesses a cleavable signal peptide, and is thus predicted as a periplasmic or extracellular protein (Figure 1). AlyA4 and AlyA6 both belong to the PL6 family in which known activities are alginate lyase and chondroitinase B (EC 4.2.2.4). To date, the only characterized alginate lyase in this family is Q06365 from *Pseudomonas* sp. OS-ALG-9 which has a mass of 44.5 kDa (Maki et al., 1993). AlyA4 and AlyA6 are twice as large, both featuring an additional C-terminal module of unknown function. AlyA4 contains an uncleavable signal peptide, and thus is likely localized in the periplasm anchored to the

cytoplasmic membrane. AlyA6 displays a cleavable signal peptide and is thus likely exported in the periplasm or in the extracellular medium. AlyA7 is classified in the PL14 family which comprises alginate lyases, exo-oligoalginate lyases (EC 4.2.2.-) and  $\beta$ -1,4-glucuronan lyases (EC 4.2.2.14). This protein features a cleavable signal peptide, and is predicted to be either periplasmic or extracellular. Moreover, another protein, Zg3452, displays 24% sequence identity on 492 amino acids (total length 768 amino acids) with the PL15 exotype alginate lyase Atu3025 from *Agrobacterium tumefaciens* C58 (Ochiai et al., 2010). Currently, the only known activities in the PL15 family are oligo-alginate lyase and alginate lyase, suggesting that Zg3452 may also be an alginate lyase despite its distant relationship with Atu3025. Finally, *Z. galactanivorans* possesses two enzymes, Zg3445 and Zg3804 (named UglA1 and UglA2, respectively), belonging the GH88 family, which encompasses glycoside hydrolases specific for unsaturated oligosaccharides released by polysaccharide lyases (Itoh et al., 2004). Thus, UglA1 and UglA2 are good candidates to further process the degradation products of alginate lyases and were included in the present analysis.

The genes *alyA1*, *alyA7*, *uglA1*, *uglA2* and *zg3452* are isolated in the genome of *Z. galactanivorans*, without any neighboring gene obviously related to alginate utilization. In contrast, *alyA4*, *alyA5* and *alyA6* constitute a small gene cluster (6 kbp), while *alyA2* and *alyA3* are localized in a large cluster of 13 genes (15.9 kbp, Figure 2). Most of these genes have a predicted function related to carbohydrates: a sugar permease (*zg2616*), two sugar dehydrogenases (*zg2615* and *zg2622*), two sugar kinases (*zg2613* and *zg2614*) and a lipoprotein constituted by a PKD module and a putative CBM (*zg2619*). The large cluster also contains 4 genes homologous to the genes *gntR* (*zg2617*), *susC* (*zg2621*), *susD* (*zg2620*) and *kdgF* (*zg2623*). GntR is a transcriptional factor negatively regulating the *gntRKU* operon involved in the uptake and catabolism of gluconate in *Escherichia coli* (Izu et al., 1997). SusC is a TonB-dependent receptor (TBDR) and has been shown in *Bacteroides thetaiotaomicron*



to be specific for the uptake of starch degradation products, while SusD is an associated outer membrane lipoprotein recently shown to bind malto-oligosaccharides (D'Elia and Salyers, 1996; Koropatkin et al., 2008). Finally, KdgF is involved in the pectinolytic system of *Dickeya dadantii* (formerly known as *Erwinia chrysanthemi*), but its exact function is unknown (Condemine and Robertbaudouy, 1991; Hugouvieux-Cotte-Pattat et al., 1996). This small protein is predicted by the Pfam server to adopt a cupin beta-barrel fold (Dunwell et al., 2004). The last gene of the cluster, *zg2612*, encodes a divalent cation transporter of the NRAMP family (Natural Resistance-Associated Macrophage Protein). Altogether, this large cluster encompasses various genes potentially involved in the uptake and degradation of alginate.

### ***Z. galactanivorans* produces alginate lyases in the presence of alginate**

To assess whether the presence of alginate influences the physiology of *Z. galactanivorans*, alginate lyase activities were measured in the extracellular, intracellular soluble and membrane-bound fractions of cultures grown in the presence of either glucose or alginate as sole carbon source. Very low enzymatic activities were recovered in the extracellular and intracellular soluble fractions from cultures with glucose, and no activity was detected in the membrane-bound fraction (Table 1). In contrast, the activities for cultures with alginate were higher by two orders of magnitude in the extracellular and soluble intracellular fractions, and activity was also determined in the membrane-bound fraction. The alginate lyase activity detected in the culture supernatant was purified by anion exchange chromatography ( $0.32 \text{ U}\cdot\text{mL}^{-1}$ ). SDS-PAGE analysis revealed that this enzymatic activity corresponded to a single protein with an estimated molecular mass was 40 kDa (Figure 3A). The protein was digested with trypsin and 35 peptides were identified by mass spectrometry. These peptides perfectly matched AlyA1, covering 60% of the whole sequence (Figure 3B). In addition, the calculated molecular mass of AlyA1 after cleavage of its peptide signal was in

agreement with the experimental mass determined on SDS-PAGE. Consequently, the protein purified from the culture supernatant was identified to be the alginate lyase AlyA1.

To experimentally validate the annotation of the alginate lyase genes, found in the genome of *Z. galactanivorans*, the seven corresponding coding regions were cloned into the pFO4 expression vector. This strategy allowed the heterologous expression of three soluble recombinant enzymes, namely AlyA5, AlyA7 and the catalytic module of AlyA1 (Table S6), that were found to be active on alginate, with specific activities of 189.7 U.mg<sup>-1</sup> for AlyA1, 49.7 U.mg<sup>-1</sup> for AlyA7 and 17.4 U.mg<sup>-1</sup> for AlyA5. These results confirm that AlyA1, AlyA5 and AlyA7 are true alginate lyases.

### **Prediction of transcription promoters and terminators**

DNA sequence analyses of the two clusters were performed to identify patterns that mediate transcription, such as promoters and Rho-independent terminators. Three and two sequences were found matching the consensus promoter, for the 15.9-kb and 6-kb clusters respectively (Figure 4A and Table 2A). These five putative promoters respected the TTG/TAnnTTTG consensus, except for P2 with one mismatch. The length of the spacer separating the -33 and -7 motifs was always 19 bp. Moreover, the identified sequences were relatively poor in GC (between 12 and 34 % of GC) compared to the mean value for the whole genome (43 % of GC). These parameters are consistent with studies in *Flavobacterium* species, showing that the promoter strength is influenced by the spacer length (19 bp being the optimal value), and is higher with AT-rich sequences (Chen et al., 2007b). Sequence analyses also revealed 2 putative terminators (T1 and T2, Figure 4A and Table 2B) in the intergenic regions of the 15.9-kb operon: the first one between *zg2622* and *zg2621* and the second one between *zg2618* and *zg2617*. Two other terminators (T3 and T4) were predicted in the 6-kb operon, between *zg4131* and *zg4132* and after *zg4132*.

### **The two clusters are transcribed as polycistronic messenger RNAs**

The genetic organization of the two identified clusters (Figure 2) suggested they could be transcribed as polycistronic mRNAs. To test this hypothesis, total RNA was extracted from cultures with either glucose or alginate as sole carbon source and used in RT-PCR reactions to amplify regions covering several genes. Reactions using three different primer pairs gave positive results with samples from the culture condition with alginate (Figure 4B, lanes 1A, 2A and 3A). This is consistent with the presence of messenger RNAs encoding at least the genes *zg2624* to *zg2618*, *zg2618* to *zg2613* and *zg4130* to *zg4132*, respectively. No PCR products were detected using samples from the culture with glucose, indicating the absence of the polycistronic transcripts in this condition. Northern hybridization was used to further analyze the transcription of the 15.9 kb cluster (Figure 4C). When *Z. galactanivorans* was grown with glucose, the probe against *zg2617* revealed the presence of two short transcripts of approximately 770 bases and 950 bases. Other transcripts were detected when cells were grown in the presence of alginate at around 9.5 kb, 5 kb and 2.6 kb. Compared to the glucose condition, the 950 base product was more abundant and the 770 base containing transcript was almost undetectable after cultivation with alginate.

### **Alginate induces the transcription of the alginolytic genes**

The transcriptomic regulations involved in *Z. galactanivorans* during alginate degradation were further studied using RT-qPCR. The relative expression of 20 genes predicted to be involved in alginate degradation was measured in triplicate (Table S3). Namely, the 12 genes from the 15.9-kb cluster, the three alginate lyase genes from the 6-kb cluster and five other alginate-related (AR) genes isolated in the genome, were included in this study. Furthermore, the gene *zg2612*, which encodes a divalent cation transporter, was added to the set of genes tested. Although its predicted function was not related to carbohydrate metabolism, its position directly upstream of the 15.9-kb operon suggested it

could be co-regulated. The expression of most of the target-genes was indeed higher in the presence of alginate compared to glucose, with 2<sup>n</sup>-fold variations between 1.3 for *uglA1* and 5.7 for *alyA3* (Figure 5 and Table S3). The transcription levels were relatively similar between glucose and alginate (2n-fold variation between 0 and 1) for *uglA2* and the gene *zg3452* encoding the PL15 enzyme. Interestingly, the expression level of genes encoding alginate lyases was relatively low in absence of alginate, except for *alyA7* which was significantly higher than the 6 others (Tukey-Kramer test, p-value < 0.001). To assess the possibility that the observed effects were due to catabolic repression by glucose, the expression of the target-genes was also measured when *Z. galactanivorans* was grown in the complex Zobell medium. Expression values were similar in this medium compared to glucose, except for the genes *alyA1* and *alyA7* (Table S3).

### **Phylogenetic distribution of the alginate-related clusters**

Homologues of the alginate lyases from *Z. galactanivorans* were searched for in all the genomes of Archae (93 genomes) and bacteria (1743 genomes) available on the NCBI genomic blast server (10<sup>th</sup> June 2011). The genomic vicinity of each identified alginate lyase gene was then screened for homologues of the other AR-genes from *Z. galactanivorans*, in order to determine the presence of AR-clusters in these microbes. Such AR-clusters are absent from Archae and are only present in two bacterial phyla, the *Bacteroidetes* (20 / 136 genomes) and the *Proteobacteria* (19 / 668 genomes). But the AR-cluster distribution is not even in the different bacterial classes (Table 3): they are very frequent in the *Flavobacteria* (36.1% of the sequenced genomes), while they are uncommon in the *Alphaproteobacteria* and *Gammaproteobacteria* (1.6% and 5.7% respectively). The marine habitat largely dominates the environmental distribution of the bacteria containing AR-clusters (31 / 39 genomes) (Figure S1). The exceptions are limited to three terrestrial microorganisms (the sphingobacterium *Pedobacter saltans* DSM 12145 and two strains of the *Xanthomonadales*

*Stenotrophomonas maltophilia*) and five species of human gut *Bacteroides*. Among these gut bacteria two species were isolated from individuals of Japanese origin (*B. clarus*, from Japanese volunteers, and *B. eggerthii* from Japanese Hawaiian volunteers) (Holdeman and Moore, 1974; Watanabe et al., 2010). The three other *Bacteroides* strains were isolated in USA from female patients during medical analyses, but there is no information concerning the ethnicity of these individuals ([http://www.broadinstitute.org/annotation/genome/bacteroides\\_group/GenomeDescriptions.html](http://www.broadinstitute.org/annotation/genome/bacteroides_group/GenomeDescriptions.html)).

With the exception of the putative phosphotransferase *zg2613* which is only conserved in *Cellulophaga algicola*, the large AR-cluster of *Z. galactanivorans* is almost entirely conserved throughout 12 flavobacterial genera (Figure 6 and Table S4). The synteny of this gene cluster is also well preserved in these marine bacteria. In most cases, there is only one large AR-cluster which also encompasses homologues of the alginate lyase genes localized in the small AR-cluster of *Z. galactanivorans* (*alyA4*, *alyA5* and *alyA6*). In contrast, the AR-clusters are much shorter in the other bacterial lineages and their gene composition is more variable. For instance, the *susD/susC*-like genes (*zg2620* and *zg2621*) are only conserved in other *Bacteroidetes* classes (*Bacteroidia*, *Cytophagia* and *Sphingobacteria*). Although there is no *susD*-like gene in *Proteobacteria*, a TBDR gene is nonetheless present in the AR-cluster of the *Alphaproteobacteria* and most of the *Gammaproteobacteria*. However, these genes are too distantly related to *zg2621* to be its orthologous genes. The closest gene from *Z. galactanivorans* is *zg1355*, an isolated TBDR gene without an associated *susD*-like gene, which displays less than 25% sequence identity with its proteobacterial homologues. The *Vibrio* species constitute an exception among the *Proteobacteria*, since none of them possess a significant homologue of alginate-related TBDR (*Bacteroidetes* or *Proteobacteria*-type) and of the sugar permease *Zg2616*. Furthermore, the PKD-CBM protein *Zg2619* is unique to

*Flavobacteria*, while the alginate lyase AlyA2 (PL7) is only conserved in two *Vibrio* species (Figure 6 and Table S4). The synteny is also less preserved among *Proteobacteria*, even between different species of a same genus (e.g. *Pseudoalteromonas haloplanktis* and *P. atlantica*), with frequent interruption of the AR-cluster by unrelated genes.

Since some AR-genes were found missing in several gene clusters, a systematic search for isolated genes, homologous to the AR-genes from *Z. galactanivorans* (>40% sequence identity), was performed in each genome. Few isolated AR-genes were indeed found in some species (highlighted in pink in Supplemental Table S4), but even with these additional genes the AR-clusters from *Sphingobacteria*, *Bacteroidia* and *Proteobacteria* remain incomplete in comparison to the flavobacterial operons. Finally, phylogenetic trees of each protein family of the alginolytic system (with the exception of Zg2613 and Zg2619 which are unique to *Flavobacteria*) were reconstructed using a maximum likelihood method (Figure S2). While the flavobacterial proteins cluster as a single group for three protein families (NRAMP, GntR-like and AlyA3), they emerge in two distinct clades in the other cases. These intriguing results are explained by gene duplication events followed by a divergent evolution of the duplicated genes. Indeed, several protein families comprise at least two divergent paralogous genes in the genomes of extant *Flavobacteria*. The presence of duplicated *kdgK*-, *susD*- and *susC*-like genes is limited to two species, *Flavobacteriales* bacterium ALC-1 and *Polaribacter dokdonensis* MED152. In contrast, the short chain dehydrogenases *zg2615* and *zg2622*, and the alginate lyases *alyA4* and *alyA6* (PL6 family) are conserved as paralogous genes in most flavobacterial species. The phylogenetic analysis also revealed several cryptic gene duplications. Such phenomena are explained by the occurrence of a gene duplication in the most recent common ancestor of a group of organisms (e.g. *genes a* and *b*), followed by the loss of *gene a* in the first subgroup of descendants, while the second subgroup lost *gene b*. This is clearly the case for the sugar permease (*zg2616*) and the *kdgF*-like gene (*zg2623*). A

cryptic duplication of *susD*- and *susC*-like genes also occurred in most flavobacterial species, but only *Flavobacteriales* bacterium ALC-1 and *P. dokdonensis* MED152 conserved both copies. In the subsequent paragraphs, the duplicated genes will be distinguished by the letters a and b, the a-group including by definition the gene(s) from *Z. galactanivorans*.

Interestingly, the AR-proteins from the gut *Bacteroides* constitute single clades which are sister groups of the proteins from marine *Flavobacteria* or are even rooted by them. The gut *Bacteroides* and the marine *Flavobacteria* have in common only one copy of each duplicated gene, namely KdgK\_a, SDR\_b, MFP\_a, *susD*-like\_a *susC*\_like\_a, Kdgf\_b and *alyA6* (Figure 6). In most phylogenetic trees the AR-proteins from *Vibrio* species emerge into monophyletic clades distinct from the other *Gammaproteobacteria*, which rather cluster with the *Alphaproteobacteria* (e.g. SDR, MFP, GntR-like, *AlyA3*, *AlyA6*). In several cases, the proteobacterial proteins constitute sister groups of the AR-proteins from marine *Flavobacteria* or even emerge with a flavobacterial clade (e.g. SDR, MFP, *AlyA3*). As observed for the *Bacteroidia*, *Cytophagia* and *Sphingobacteria*, the *Proteobacteria* have conserved only one copy of the flavobacterial duplicated genes. It is noteworthy that the *Vibrio* species have not conserved the same gene copies than the other *Gammaproteobacteria* (Figure 6).

## DISCUSSION

### ***Z. galactanivorans* possesses a complex alginolytic system organized into two operons**

*Z. galactanivorans* is able to grow using alginate as sole carbon source and indeed possesses a great number and variety of enzymes predicted to act in the degradation of alginate. Particularly, this marine bacterium displays at least 7 putative alginate lyases belonging to four different PL families, with various modular architectures. This complexity is reminiscent of the enzymatic systems involved in the deconstruction of plant cell walls

(Gilbert, 2010). It is likely an indication of the difficulty to hydrolyze alginates in the cell walls of brown algae. Indeed, alginates form fibers in tight interaction with the other components of brown algal cell walls, whereas they constitute a simple exopolysaccharide in bacteria. Most of the alginate lyase genes from *Z. galactanivorans* are not randomly distributed in its genome, but are co-localized within two gene clusters. The 6-kb cluster consists of the three genes *alyA4*, *alyA5* and *alyA6*, whereas the 15.9-kb cluster contains 12 genes with putative carbohydrate-related functions, including *alyA2* and *alyA3*, and a NRAMP family transporter gene (Figure 2). Consensus promoters and Rho-independent terminators were detected in these two genomic regions (Figure 4A and Table 2). RT-PCR analyses showed that the two clusters are transcribed as polycistronic mRNAs in the presence of alginate (Figure 4B). Since the primer sets 1 and 2 were designed to give two overlapping amplification products, this experiment showed that the 12 genes from *zg2624* to *zg2613* can be transcribed on a 15.9 kb polycistronic mRNA. Differential regulation of transcription seems to occur, as polycistronic mRNAs were not detected by RT-PCR when cells were grown with glucose. In the Northern Blot analysis focusing on the 15.9-kb operon, the probe against *zg2617* revealed only two small RNA species (770 and 950 bases) in the presence of glucose (Figure 4C). Since *zg2617* is a 729 bp ORF, the two products correspond to the gene transcribed alone with different initiation or termination sites. Interestingly, this gene encodes a transcription regulator belonging to the GntR family (Haydon and Guest, 1991), constituted by a N-terminal DNA-binding module (Helix-Turn-Helix, HTH) and a C-terminal FCD domain involved in effector binding and oligomerization. Therefore, *Zg2617* is a good candidate to act as a transcription regulator for the alginolytic pathway in *Z. galactanivorans*. The presence of alginate in the medium drastically changed the transcription profile (Figure 4C). The transcription of the 950 bases transcript for *zg2617* seemed to be favored in this culture condition. Moreover, new RNA species were detected. The larger transcript (9.5 kb)



could correspond to a transcription initiated at the P2 site and covering the region from *zg2621* to *zg2616*. The products of 5 kb and 2.6 kb could both correspond to initiations at the site P3 but with different termination sites. The 15.9 kb transcript evidenced by the RT-PCR results was not detected by Northern blot hybridization. This could reflect the low abundance of such a large transcript and/or a higher sensibility to degradation. Altogether, these results indicate that *Z. galactanivorans* possesses two alginolytic operons comprising 3 and 12 genes, respectively, and which are transcribed as different polycistronic messengers depending on the presence or absence of alginate in the extracellular medium.

### **Alginate acts as an inducer of the alginolytic pathway**

The regulation of the predicted AR-genes was analyzed by RT-qPCR. In glucose-supplemented medium, the expression values for these genes were relatively low (Table S3). Nevertheless, two genes had noteworthy transcription levels, namely *zg2617* and *alyA7*. *Zg2617* was the most transcribed among the set of genes tested in RT-qPCR, consistently with the high abundance of the *zg2617* transcript seen in the Northern Blot analysis (Figure 4C). These results strengthen the hypothesis that this GntR-like regulator likely act as a sensor waiting to detect the presence of alginate to mediate the transcription of AR-genes. Furthermore, *alyA7* was significantly more expressed in the presence of glucose than the 6 others alginate lyase genes, indicating that this putative alginate lyase is constitutively transcribed. The expression of all the genes within the two operons and of three isolated AR-genes (*alyA1*, *zg2612*, *uglA1*) was up-regulated in the presence of alginate compared to glucose (Figure 5). For most of the genes, this cannot be explained by a catabolic repression mediated by glucose, as similar expression patterns were found in Zobell medium and glucose-supplemented minimum medium (Table S3). However, a slight catabolic repression cannot be excluded for *alyA1*, which expression was 5-times higher in Zobell medium than in the presence of glucose. The transcription of *zg2617*, which was already high in glucose, was

still induced 4-times in response to alginate, suggesting that this GntR-like transcription factor probably acts as a positive regulator. Other members of the GntR family have already been characterized as transcription activator in other bacterial species, affecting various metabolic pathways, such as the biosynthesis of the antibiotic prodigiosin (Fineran et al., 2005), the utilization of taurin (Wiethaus et al., 2008), and of  $\beta$ -glucosides (Neelakanta et al., 2009). Since the protein is predicted as cytoplasmic, the effector cannot be a polysaccharide, but rather an alginate degradation product. Such a regulatory mechanism would require a basal expression of alginate lyase activity to produce the first oligo-alginates and initiate the signalization pathway. The alginate lyases AlyA1 and AlyA7, which are significantly expressed in Zobell medium and in the presence of glucose, respectively, are good candidates to play this potential role of "sentry" enzymes. Interestingly, the transcription of *alyA7* was down-regulated in response to alginate (Figure 5), suggesting that the physiological role of this alginate lyase is accomplished once the presence of alginate is detected. In contrast, all the other putative alginate lyases are highly up-regulated in the presence of alginate, suggesting that *Z. galactanivorans* adapts its transcriptome towards more efficient alginolytic systems including *alyA1* and the two AR-operons. In our experimental conditions, the transcription levels of *zg3452* (PL15) and *uglA2* were not affected by alginate ( $2^n$ -fold variation between 0 and 1), suggesting that these two genes are not involved in alginate catabolism.

The influence of alginate on the physiology of *Z. galactanivorans* was further investigated at the protein level. While very low alginate lyase activities were detected when cells were grown in glucose, these activities were drastically induced in the presence of alginate (Table 1). The detection of alginate lyase activities in the extracellular, intracellular and membrane-bound fractions reflects the various localizations predicted by the sequence analyses of the seven putative alginate lyases (Figure 1). Moreover, the extracellular

enzymatic activity was shown to essentially correspond to AlyA1, demonstrating that this protein is indeed an alginate lyase (Figure 3). Since this enzyme is predicted as a lipoprotein of the outer membrane, AlyA1 is eventually released by proteolysis in the extracellular medium. During the purification process, AlyA1 was co-precipitated with alginate, indicating that this enzyme is strongly bound to alginate microgels. Interestingly, AlyA1 is the only alginate lyase in *Z. galactanivorans* to contain a CBM (Figure 1). This additional non-catalytic module likely renders AlyA1 specific for the initial degradation of semi-crystalline alginate fibers. Altogether, our results demonstrate that *Z. galactanivorans* is able to perceive the presence of alginate in its environment and to regulate its metabolism to efficiently degrade this algal polysaccharide.

### **Alginate perception plays a role in algae / bacteria interactions**

The set of genes tested by RT-qPCR also contained *zg2612*, which encodes a divalent cation transporter of the NRAMP family (Natural Resistance-Associated Macrophage Protein). Although this function does not obviously relate to carbohydrate metabolism, the expression of this gene was surprisingly up-regulated in the presence of alginate, suggesting that this gene eventually belongs to the large AR-operon. The emergence of bacterial NRAMP transporters was previously proposed as an adaptation to oxidative environments, including those arising during infection of animals and plants (Cellier et al., 2001). In *Salmonella typhimurium* and *E. coli*, NRAMP transporters have been shown to participate in the resistance towards hydrogen peroxide (Kehres et al., 2000). In marine macroalgae, the perception of a pathogen attack (such as a bacterial aggression) is followed by an oxidative burst consisting of a rapid, strong, and transient emission of reactive oxygen species (ROS) (Cosse et al., 2007). Indeed, the production of H<sub>2</sub>O<sub>2</sub> by brown algae is a key step in their resistance towards infection (Küpper et al., 2002), and can control the growth of epiphytic bacteria (Küpper et al., 2001). Brown algae perceive oligo-alginates resulting from the

degradation of their cell walls as signal molecules (referred to as elicitors), triggering their defense mechanisms (Küpper et al., 2001). During host/pathogen interaction, alginate-consuming bacteria thus need to protect themselves against the elicitor-induced oxidative burst. The up-regulation of *zg2612* in *Z. galactanivorans* could therefore be involved in such a protection mechanism, quenching the release of ROS by consumed brown algae. Our results suggest that the perception of alginate could "warn" *Z. galactanivorans* about the imminence of an oxidative burst, triggering the synthesis and/or activation of resistance factors such as the NRAMP transporter.

### **Model of the Alginate Utilization System in *Z. galactanivorans***

Based on our results, we propose a model of the alginolytic pathway in *Z. galactanivorans* (Figure 7). The perception of alginate in the environment leads to an alginolytic response with two main aspects: (i) the degradation of the polysaccharide and (ii) potentially the resistance towards algal defense reactions such as ROS production. The cleavage of alginate is accomplished by several predicted alginate lyases, and the activity of three of them (AlyA1, AlyA5 and AlyA7) was confirmed biochemically. The signaling steps involved in this regulation are still unclear, but we propose that oligo-alginates play a key role as effectors. Given the available data on alginate catabolism in other bacteria (Preiss and Ashwell, 1962b, a), it appears that the identified alginate-responsive genes in *Z. galactanivorans* can account for the complete breakdown of alginate into 2-keto-3-deoxy-6-phosphogluconate (KDPG). This compound can be further processed into pyruvate by a KDPG aldolase (EC 4.1.2.14) encoded by *zg4702* in the genome of *Z. galactanivorans*. Although supported by our results and the literature, this model will need further validation.

Apart from proteins with enzymatic activities, the 15.9-kb operon also contains two genes, *zg2620* and *zg2621*, which encode homologues of the TonB-dependent receptor SucC

and its associated lipoprotein SusD, involved in the uptake of starch degradation products in *B. thetaiotaomicron* (Anderson and Salyers, 1989a, b). As for other members of the large operon, their expression was highly upregulated in response to alginate. Therefore, Zg2620 and Zg2621 likely play roles similar to those of SusC and SusD and are involved in the uptake of alginate-oligosaccharides released by the extracellular and outer membrane-bound alginate lyases. Clusters containing *susC/susD* homologues and other carbohydrate-related genes are referred to as Polysaccharide Utilization Loci (PUL) and are abundant in environmental and gut *Bacteroidetes* (Martens et al., 2009; Thomas et al., 2011b). However, very few PULs have been experimentally shown to be genuine operons. The transcription of the xylan utilization clusters from *Prevotella bryantii* B14 was recently analyzed, but these gene clusters were not experimentally proven to be transcribed as polycistronic mRNAs (Dodd et al., 2010; Dodd et al., 2011). To our knowledge, the *zg2613-zg2624* cluster is thus the first alginolytic operon in a marine bacterium and the second characterized polysaccharide utilization operon in *Bacteroidetes*, after the starch utilization system (D'Elia and Salyers, 1996). This Alginate Utilization System (AUS) contains additional genes completing the alginate catabolisingolytic machinery from *Z. galactanivorans*. Zg2616 is a sugar permease belonging to the Major Facilitator superfamily (MFP), which probably transports oligo-alginates from the periplasm to the cytoplasm. Finally, the function of Zg2619 and Zg2623 (KdgF-like protein) are unclear. Zg2619 is an outer membrane lipoprotein containing a PKD module and a CBM, and may participate to alginate binding together with the SusD/SusC homologues (Zg2620 and Zg2621). In *D. dadantii*, KdgF is necessary for the pectinolytic activity of this phytopathogen bacterium, but its exact role is unknown. Since Zg2623 is well conserved in AUS-containing bacteria, this protein is also probably essential to the alginolytic system.

## Origin and evolution of the Alginate Utilization System

The large alginolytic operon of *Z. galactanivorans* is conserved at various degrees in 39 bacteria. Most of these microorganisms (31 species) were isolated in coastal regions all around the world, including North Atlantic Ocean, North Pacific Ocean, Southern Ocean, North Sea and English Channel, Mediterranean Sea and Sea of Japan (Figure S1). All these marine ecosystems host various species of brown algae (see algaeBASE, <http://www.algaebase.org/>), and thus comprise significant alginate resources. These alginolytic operons are only found in two distant bacterial phyla, the *Bacteroidetes* and the *Proteobacteria*. It is noteworthy that the alginate-producing bacteria (e.g. the *Gammaproteobacteria* *Pseudomonas aeruginosa* and *Azotobacter vinelandii*) and the terrestrial microorganisms degrading bacterial alginate (e.g. the *Alphaproteobacteria* *Sphingomonas* spp.) do not feature AUS similar to the large alginolytic operon identified in *Z. galactanivorans*. Due to this patchy taxonomic distribution, it is highly unlikely that this AUS existed in the common ancestor of the *Bacteroidetes* and the *Proteobacteria*, since this hypothesis would require multiple gene losses to explain the current taxonomic distribution. Therefore, the most plausible scenario is the occurrence of one or several independent lateral gene transfers (LGT) between these various bacterial species. The completeness of the alginolytic operon from the marine *Flavobacteria*, its high frequency in this taxonomic group, and its degree of conservation, both in term of gene composition and of synteny, strongly support that this AUS arised from an ancestral marine flavobacterium (Figure 6). In comparison, the AUS from the other bacterial groups are much shorter, frequently fragmented and less conserved. Our phylogenetic analyses have unraveled gene duplications in the common ancestor of the AUS-containing flavobacteria. Several gene duplications are still visible in the genome of extant flavobacterial species (Figure 6 and S1, Table S5), but some of these ancestral gene duplications are now cryptic (i.e. MFP, *kdgF*-like, *susD*-like and *susC*-

like). The conservation of specific gene copies in the various bacterial groups supports the occurrence of independent LGT events with different ancestral marine flavobacteria differing by the exact composition of their AUS. For instance, extant *Bacteroidia* display a sugar permease belonging to the MFP\_a subgroup, whereas all the other bacteria possess the MFP\_b subgroup. Similarly, the *Vibrio* species contain the a-type SDR (Zg2615), which is strictly conserved in the marine flavobacteria, whereas the other microorganisms have the b-type SDR (Zg2622) which has been lost by several extant marine flavobacteria.

The microbial communities colonizing the surface of macroalgae are usually dominated by marine *Flavobacteria*, *Alphaproteobacteria* and *Gammaproteobacteria* (Staufenberger et al., 2008; Salaun et al., 2010; Tujula et al., 2010). Therefore, this ecological promiscuity has certainly favored the AUS transfers from the marine *Flavobacteria* to the marine *Proteobacteria*. The acquisition of an AUS of marine origin by several human gut *Bacteroides* is more surprising at first sight. However, this alginolytic operon is present in two species isolated from Japanese and Japanese-Hawaiian individuals (*B. clarus* and *B. eggerthii*). Our group has recently demonstrated that agarolytic and porphyranolytic genes from marine *Flavobacteria* were acquired by the gut bacterium *B. plebeius*, which was also isolated from Japanese volunteers. These LGT events were likely promoted by the traditional consumption of edible red algae, such as *Porphyra* (Nori in Japanese) (Hehemann et al., 2010; Rebuffet et al., 2011). Several brown algae are also part of the daily diet of Japanese, mainly used in soups and salads: *Eisenia bicyclis* (Arame), *Sargassum fusiforme* (Hijiki), *Saccharina japonica* (Kombu) and *Undaria pinnatifida* (Wakame). Therefore, the most plausible scenario explaining this LGT is the uptake of marine AUS-containing *Flavobacteria* at the surface of edible brown algae. Such transfers of CAZYme genes (Hehemann et al., 2010; Rebuffet et al., 2011) or even of complete polysaccharide utilization operons (this study) were thus promoted by the consumption of very different sea vegetables (edible red and brown algae), suggesting

that these LGTs are more frequent than initially expected (Hehemann et al., 2010) and may be associated to the consumption of seafoods in general.

Finally, the phylogenetic proximity between the *Flavobacteria* and the *Bacteroidia* has likely facilitated the conservation of the core alginolytic operon observed in the *Bacteroidetes* (*kdgK*-like, SDR, sugar permease, *gntR*-like, *susD*-/*susC*-like, *kdgF*-like, *alyA3*, *alyA6*). In contrast, the core operon within the *Proteobacteria* (without the *Vibrio* species) is more limited (SDR, sugar permease, *gntR*-like, *alyA3*). In particular, the TonB-dependent receptor of the *Bacteroidetes* has been replaced by another TBDR, specific for the *Proteobacteria*. This surprising event is likely explained by the fact that the SusD-like proteins are unique to the *Bacteroidetes* (Thomas et al., 2011b) and that these proteins physically and functionally interact with their associated TBDR (Martens et al., 2009). For yet unknown reasons, the SusD-like proteins seem incompatible with the physiology of the *Proteobacteria* and these transferred genes were lost by the proteobacterial species together with their associated *susC*-like genes. We propose that most of the AUS-containing *Proteobacteria* have subsequently recruited another TBDR of proteobacterial origin which evolved to fulfill the role of the lost alginate SusC/SusD homologues. The *Vibrio* species constitute an even more radical case, since these microorganisms do not possess any obvious alginate-related TBDR or sugar permease. This suggests that alginolytic *Vibrio* species have likely developed a different system for the uptake of alginate oligosaccharides.

## **Conclusion**

We have characterized the first Alginate Utilization System in a marine bacterium, which is mainly organized in two operons inducible by the presence of alginate. We also propose a model for the catabolism of this polysaccharide in this bacterium, including its detection, the uptake of its degradation products and their complete assimilation by *Zobellia*



*galactanivorans*. Such AUS emerged in an ancestral marine flavobacterium and were acquired by diverse bacteria, which thus gained the capacity to assimilate the main polysaccharide of brown algae, an adaptive advantage in the coastal environment but also in the gut microbiota of specific human populations.

## **EXPERIMENTAL PROCEDURES**

### **Reagents**

Unless otherwise stated, all reagents were purchased from Sigma-Aldrich.

### **Bacterial strain and culture conditions**

The type strain Dsij<sup>T</sup> of *Z. galactanivorans* was routinely grown at 20°C in Zobell medium 2216E (5 g/L tryptone, 1 g/L yeast extract, sea water, ZoBell, 1941) until mid-exponential phase ( $A_{600\text{nm}}$  around 0.3). Cells were transferred either in fresh Zobell medium or in the Marine Mineral Medium supplemented with 2 g.L<sup>-1</sup> glucose or alginate as sole carbon sources. Briefly, one liter of Marine Mineral Medium is composed of 24.7 g NaCl, 6.3 g MgSO<sub>4</sub>.7H<sub>2</sub>O, 4.6 g MgCl<sub>2</sub>.H<sub>2</sub>O, 2 g NH<sub>4</sub>Cl, 0.7 g KCl, 0.6 g CaCl<sub>2</sub>, 200 mg NaHCO<sub>3</sub>, 100 mg K<sub>2</sub>HPO<sub>4</sub>, 50 mg yeast extract, 20 mg FeSO<sub>4</sub>.7H<sub>2</sub>O, in Tris-HCl 50 mM pH 8.0 (Thomas et al., 2011a). Cultures were grown at 20°C and 150 rpm. For RNA isolation, 5 mL cultures were grown in 50 mL Falcon tubes. For protein extraction and purification, cultures were grown in flasks with baffles, allowing a 1:5 liquid to air ratio.

### **RT-PCR and quantitative PCR**

Total RNA isolation and cDNA synthesis were performed as described previously (Thomas et al., 2011a). Briefly, RNA was isolated from 2 mL of cultures in mid-log phase ( $A_{600\text{nm}}$  0.3-0.5) using RNA Mini Kit (Qiagen) and treated with Turbo DNase (Ambion). Total elimination of genomic DNA was checked by PCR on RNA samples. Eight hundreds ng of RNA were retro-transcribed using the Phusion RT-PCR Kit (Finnzymes) with random

hexamer primers. cDNA samples were stored at -20°C and used in PCR reactions within a month. PCR reactions were performed using the Advantage 2 PCR Kit (Clontech) with 0.2 µM of each primer (Table S1) for 30 cycles (95°C for 30 s, 68°C for 6 min). Samples were analyzed by electrophoresis on 0.8% agarose gel in TAE buffer and visualized by ethidium bromide. Quantitative PCR reactions were performed on cDNA samples using gene-specific primers (Table S1). Primers were validated (Table S2) as previously described (Thomas et al., 2011a). The geometric average of the transcript abundance of *glyA*, *icdA* and *gmkA* was used to normalize the expression of the genes of interest, as described previously (Thomas et al., 2011a); 2<sup>n</sup>-fold variations of expression in alginate relative to growth in glucose were calculated.

### **Northern Blot analysis**

Total RNA was extracted from 1 mL of culture using the TRI reagent (Sigma) following the manufacturer's instructions. Samples containing 1 µg of RNA were separated by electrophoresis for 4 h at 40 V on 1% agarose gel containing 2% formaldehyde. As a control of sample quantity and integrity, duplicate samples were run on the same gel and stained with ethidium bromide. Nucleic acids were transferred by capillarity to a nylon membrane (Hybond N+, Amersham) and cross-linked with UV (1 min, 125 mJ). The DIG-labelled *zg2617* specific DNA probe was constructed by PCR (PCR DIG Probe synthesis kit, Roche Applied Science) following the manufacturer's instructions, using 0.2 µM of each primer (Table S1) and 10 ng of *Z. galactanivorans* genomic DNA prepared as described previously (Barbeyron et al., 1984). RNA was hybridized with 2 µL of probe for 16 h at 50°C and detected using anti-DIG/alkaline phosphatase antibody and CDP-Star (Roche Applied Science) following the manufacturer's instructions. The membrane was exposed for 30 min to a detection film (Kodak). Approximate sizes of the hybridizing mRNA were estimated by using a 0.2 to 10 kb RNA ladder (RNA marker R7020, Sigma).

### **Protein extraction**

After microscopic observation to ensure that no significant cellular lysis had occurred, 8 mL of late-log phase ( $A_{600\text{nm}}$  0.9-1.0) cultures in Marine Mineral Medium supplemented with either glucose or alginate were centrifuged 10 min at 3,000 *g*. Supernatants (hereafter referred to as “extracellular fraction”) were conserved at 4°C. Cells were resuspended in 5 mL of ice-cold Tris-HCl buffer (0.1 M, pH 7.5) and disrupted in a French pressure cell under 1.6 kbar. The suspensions were centrifuged for 30 min at 20,000 *g* and at 4°C to remove cell debris. Supernatants were conserved at 4°C (hereafter referred to as “intracellular soluble fraction”). Cell debris were washed in 3 mL of Tris-HCl buffer (0.1 M, pH 7.5) and centrifuged for 30 min at 20,000 *g* and 4°C. Pellets were resuspended in 1 mL of Tris-HCl buffer containing 0.1% Tween 20 and incubated overnight to solubilize membranes. After centrifugation (22,000 *g*, 25 min, 4°C), supernatants were conserved at 4°C (hereafter referred to as “membrane-bound fraction”). Protein concentrations in each fraction were determined by the Bradford method (Bradford, 1976) using BSA as a standard.

### **Alginate lyase activity**

Alginate lyase activity was assayed by measuring the increase in absorbance at 235 nm, due to the release of unsaturated uronic acids, for 5 min at 25°C in a 1 cm quartz cuvette (Preiss and Ashwell, 1962b). Reaction mixtures (1 mL) were composed of 0.1% (w/v) alginate in Tris-HCl buffer (0.1 M, pH 7.5). One unit of enzyme activity was defined as an increase of 1 in absorbance at 235 nm, at 25°C, and per min.

### **Purification and identification of a secreted enzyme**

Bacteria were grown until early stationary phase ( $A_{600\text{nm}}$  1.2) in mineral medium supplemented with alginate. After microscopic observation to ensure that no significant lysis had occurred, cells were removed by centrifugation for 20 min at 12,000 *g* and 4°C. Unless

otherwise stated, the following treatments were conducted at 4°C. The alginate present in the supernatant was precipitated by slow addition of 0.25 volume of 100% ice-cold ethanol under agitation, and recovered by centrifugation for 30 min at 30,000 g. Alginate lyase activity was measured on the supernatant and the pellet. The pellet, which contained most of the alginolytic activity was resuspended in 10 mL of Buffer A (Tris-HCl 0.1 M, pH 7.5), incubated for 4 h at 30°C and passed through 0.2 µm filter. The sample was brought to 80% ammonium sulfate saturation and centrifuged for 20 min at 15,000 g. The supernatant (9.7 mL) was dialyzed against Buffer A in membrane with 6-8 kDa MWCO (Spectrum Laboratories, USA) and injected on an anion exchange column (MonoQ, 1 mL, GE Healthcare) equilibrated with the same buffer. Proteins were eluted at 1 mL.min<sup>-1</sup> with a 20 mL linear gradient of 0 to 1M NaCl in Buffer A and 0.5 mL fractions were collected. Elution was followed by the absorbance at 280 nm. Ten µL of each fraction were assessed for alginate lyase activity. An active fraction was obtained at around 300 mM NaCl and precipitated with 3 volumes of acetone prior to analysis by SDS-PAGE on 15% polyacrylamide gel. Coomassie blue-stained protein slices were cut into small pieces, washed with distilled water and destained using acetonitrile (ACN). The cysteine residues were reduced by 100 µL of 10 mM DTT at 56°C and alkylated by 150 µL of 55 mM iodoacetamide at room temperature. Iodoacetamide solution was replaced by 100 µL of 100 mM NH<sub>4</sub>HCO<sub>3</sub> and gel dehydration was achieved with ACN. After evaporation in a Speed Vac (Thermo), proteins were digested overnight at 37°C by a solution containing 0.9 µg of modified bovine trypsin (Promega, Madison, WI, USA) prepared in 25 mM NH<sub>4</sub>HCO<sub>3</sub>. Finally, a double extraction was performed, first with 1% (v/v) trifluoroacetic acid (TFA) solution, and subsequently with 100% (v/v) ACN. The resulting peptide mixture was extracted and dried under vacuum and resuspended in 50% ACN, 0.1% TFA solution. Peptide Mass Fingerprinting by MALDI-TOF MS was performed using a Voyager DE-STR MALDI-TOF mass spectrometer (Applied

Biosystems). Each pooled tryptic digest (1  $\mu$ L) was mixed with an equal volume of  $\alpha$ -Cyano-4-hydroxycinnamic acid matrix (10 mg.mL<sup>-1</sup> in 50% ACN, 0.1% TFA) and spotted onto the MALDI target. Spectra were acquired in positive ion reflector mode under 20 kV accelerating voltage and a mass range of 450–5000 Da. Internal calibration was performed using trypsin autolysis fragments at m/z 1433.70, 2163.05, and 2289.1.

### **Cloning, expression and purification of recombinant alginate lyases**

The seven genes encoding putative alginate lyases (alyA1-alyA7) were selected for a medium-throughput cloning strategy performed as previously described (Groisillier et al., 2010). Briefly, primers were designed to amplify the coding region corresponding to the mature proteins, except for AlyA1 for which only the catalytic module was cloned. The genes were amplified by PCR from *Z. galactanivorans* genomic DNA using the oligonucleotide primers reported in Table S6. Using the restriction sites indicated in Table S6, the PCR products were then ligated into the expression vector pFO4, resulting in a gene encoding the protein flanked by an N-terminal hexa-histidine tag. The recombinant plasmids were transformed into *E. coli* BL21(DE3) expression strains. Recombinant strains were cultured at 20°C during 3 days in 200 mL of ZYP medium (Studier, 2005) supplemented with 100 mg.mL<sup>-1</sup> ampicillin. The cells were resuspended in 25 mM Tris-HCl pH 7.5, 200 mM NaCl and 5 mM imidazole buffer, supplemented with one dose of an anti-protease cocktail (Complete EDTA free, Roche) and 0.1 mg.mL<sup>-1</sup> of DNase. Cells were disrupted in a French pressure cell under 1.6 kbar. The samples were centrifuged for 2 h at 20,000 g and at 4°C. The supernatant was loaded onto a HyperCell PAL column charged with NiSO<sub>4</sub> (0.1 M) and pre-equilibrated with resuspension buffer. The His-tagged proteins were eluted with a linear imidazole gradient (5-500 mM imidazole) at a flow rate of 1 ml.min<sup>-1</sup>. Alginate lyase activity tests were performed as described above on the obtained fractions.

## Sequence analyses

The alginate-related (AR) genes of *Z. galactanivorans* were identified in its complete genome (sequenced by Genoscope: <http://www.genoscope.cns.fr/spip/Marine-cytophagales-specialized.html> ; Barbeyron et al., in preparation) using the program GenDB (Meyer et al., 2003). For each corresponding protein, evidence of conserved protein modules was queried on the Pfam database (Bateman et al., 2004) and by BlastP searches against the GenBank nr database. Signal peptide and transmembrane helices were predicted using SignalP v2.0 (Nielsen et al., 2007) and TMHMM (Sonnhammer et al., 1998), respectively. The presence of homologous AR genes in all available prokaryotic genomes was screened using the genomic BlastP (Altschul et al., 1997) at NCBI ([http://www.ncbi.nlm.nih.gov/sutils/genom\\_table.cgi](http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi)). The corresponding proteins were subjected to multiple sequence alignment using MAFFT (Kato et al., 2002), with the iterative refinement method and the scoring matrix Blosum62. These alignments were manually edited using Bioedit (© Tom Hall). The different phylogenetic trees were derived from these refined alignments using the Maximum Likelihood method with the program PhyML (Guindon and Gascuel, 2003). The reliability of the trees was tested by bootstrap analysis using 100 resamplings of the dataset. The trees were displayed with MEGA 5 (Tamura et al., 2011). To identify putative promoters, intergenic regions were searched using DNAMAN 4.15 for the motif -33/-7 (TTG/TAnnTTTG) with a spacer length ranging between 18 and 21 bp. This motif has been identified as the consensus promoter in *Flavobacterium* species, and is conserved in *Bacteroides fragilis* (Chen et al., 2007b; Chen et al., 2007a). The TransTermHP software was used to predict putative Rho-independent transcriptional terminators (Kingsford et al., 2007).

## ACCESSION NUMBERS

The nucleotide sequence for the genome of *Zobellia galactanivorans* has been deposited at EMBL with the accession number FP476056.

## ACKNOWLEDGEMENTS

We are grateful to Drs Carole Dossat and Valérie Barbe (Génoscope, France) for the sequence deposition at the EMBL database. The ‘Marine Plants and Biomolecules’ laboratory is funded by the French National Scientific Research Centre (CNRS) and the University Pierre and Marie Curie (UPMC). FT is supported by a PhD grant from the UPMC. This work was also supported by the European Community within the Seventh Framework Programme under grant agreement n° 222628 (Large collaborative project PolyModE, <http://www.polymode.eu/>).

## FIGURE LEGENDS

**Figure 1: Modular analysis of the putative alginate lyases predicted in *Z. galactanivorans* genome.**

PL, polysaccharide lyase (CAZy family). CBM, Carbohydrate Binding Module. CM, cytoplasmic membrane.

**Figure 2: The two alginolytic clusters of *Z. galactanivorans*.**

**Figure 3: Purification and identification of AlyA1 in the extracellular fraction.**

A. SDS-PAGE analysis of the active, purified protein. B. Superposition of AlyA1 protein sequence and peptides identified in mass spectrometry (red).

**Figure 4: Transcription analysis of the two alginolytic clusters of *Z. galactanivorans*.**

A. Genetic organization of the two alginolytic clusters found in *Z. galactanivorans* genome. Predicted promoters (P1 to P5) and Rho-independent terminators (T1 to T4) are depicted, as well as primers used for the RT-PCR analysis (black arrows; fw: forward; rv: reverse).

B. RT-PCR analysis of the transcription pattern for the two clusters. Agarose gel of RT-PCR products resulting from primer pairs shown in panel A. RT-PCR were performed on RNA extracted from culture with glucose (G) or alginate (A). Lanes: MW, molecular mass standard; 1, primer set 1; 2, primer set 2; 3, primer set 3.

C. Northern Blot analysis of 1  $\mu$ g total RNA extracted from cultures with glucose (G) or alginate (A). Membrane was hybridized with an internal probe targeting *zg2617* as shown in panel A. Molecular masses of transcripts were extrapolated from a molecular ladder in an adjacent lane.

**Figure 5: Differential gene expression in the presence of alginate compared to glucose.**

Transcript levels were quantified by RT-qPCR. Values are mean  $\pm$  s.e.m. (n=3). Dotted lines denote the significance threshold ( $2^n$  fold-variation  $>1$ , upregulated;  $2^n$  fold-variation  $<1$ , down-regulated).

**Figure 6: Phylogenetic distribution of the Alginate Utilization System.**

The number of strains for which each gene was found is indicated in the arrow. The number of strains considered in each bacterial group is given between brackets. The gene organization was arbitrarily represented as found in *Z. galactanivorans*. Duplicated genes share the same name, followed by "\_a" or "\_b". This figure was built from the information presented in Table S4.

**Figure 7: Model of the alginolytic system in *Z. galactanivorans*.**

The genes encoding enzymes putatively involved at each step are given. The color code represents the variation of the gene expression in alginate compared to glucose (blue, down-regulated; yellow, up-regulated).



## TABLES

**Table 1: Total alginate lyase activity in the extracellular, intracellular soluble and membrane-bound fractions.**

Carbon source	Total activity in fraction (mU)*		
	extracellular	intracellular soluble	membrane-bound
Glucose	4.9	5.7	not detected
Alginate	266.1	535.9	27.3

\* one unit of activity was defined as an increase of 1.0 in absorbance at 235 nm per min

**Table 2: Predicted promoters and terminators in the alginolytic clusters.**

A. Promoters identified by searching for the consensus sequence -33/-7 (TTG/TAnnTTTG) characterized in *Flavobacterium* sp. and *Bacteroides fragilis*.

B. Rho-independent transcriptional terminators predicted by the TransTermHP software.

### A

promoter	sequence*	spacer length	GC%
15.9-kb cluster			
P1	TTGctagtccacaaatactTAttTTTG	19 bp	24.4
P2	TTttttaacaaactattaagtTAAtTTTG	19 bp	12.2
P3	TTGttttgagtagattggcagTAatTTTG	19 bp	34.2
6-kb cluster			
P4	TTGttttaaataataaaacgacTAAtTTTG	19 bp	17.1
P5	TTGtcggttctgtaataatagTAttTTTG	19 bp	22.0

\*uppercase letters refer to nucleotides matching the consensus TTG/TAnnTTTG

### B

terminator	sequence <sup>#</sup>	confidence
15.9-kb cluster		
T1	uaaauuuuuuuuaaaaGGAUGGGCgugcACCCAUCCuuuuuuuuuugaaaa	93%
T2	aagcgatttttagaaTTCTACCGACCCGtggtgagCGGGTTCGGCAAAAAtt actgccaatctaa	89%
6-kb cluster		
T3	taataatatgtattTTGGTAAACCAaacTGGTTAACCAAttagaactatgaga	82%
T4	gatcatactaaaaaGGCCTA- AGTTctataAGCTCTAGGCCttttctatattgg	77%

<sup>#</sup>uppercase letters refer to nucleotides involved in the complementary interaction and lowercase letters to the 5'- and 3'- tail or internal loops

**Table 3: Distribution of the Alginate Utilization System in the different bacterial classes**

<b>Taxon</b>	<b>Number of genomes</b>	<b>Species with alginolytic operon</b>	<b>%</b>
<b><i>Bacteroidetes</i></b>	<b>135</b>	<b>20</b>	<b>14.8</b>
<i>Flavobacteria</i>	36	13	36.1
<i>Cytophagia</i>	8	1	12.5
<i>Sphingobacteria</i>	9	1	11.1
<i>Bacteroidia</i>	79	5	6.3
<b><i>Proteobacteria</i></b>	<b>668</b>	<b>19</b>	<b>2.8</b>
<i>Gammaproteobacteria</i>	279	16	5.7
<i>Alphaproteobacteria</i>	193	3	1.6

## REFERENCES

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.
- Anderson, K.L., and Salyers, A.A. (1989a) Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. *J Bacteriol* **171**: 3199-3204.
- Anderson, K.L., and Salyers, A.A. (1989b) Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *J Bacteriol* **171**: 3192-3198.
- Azam, F., and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nat Rev Microbiol* **5**: 782-791.
- Barbeyron, T., Kean, K., and Forterre, P. (1984) DNA adenine methylation of GATC sequences appeared recently in the *Escherichia coli* lineage. *Journal of bacteriology* **160**: 586-590.
- Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., and Potin, P. (2001) *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of [*Cytophaga*] *uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **51**: 985-997.
- Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S. et al. (2004) The Pfam protein families database. *Nucleic Acids Res* **32**: D138-141.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.
- Cellier, M.F., Bergevin, I., Boyer, E., and Richer, E. (2001) Polyphyletic origins of bacterial Nrapm transporters. *Trends Genet* **17**: 365-370.
- Chen, S., Bagdasarian, M., Kaufman, M.G., and Walker, E.D. (2007a) Characterization of strong promoters from an environmental *Flavobacterium hibernum* strain by using a green fluorescent protein-based reporter system. *Appl Environ Microbiol* **73**: 1089-1100.
- Chen, S., Bagdasarian, M., Kaufman, M.G., Bates, A.K., and Walker, E.D. (2007b) Mutational analysis of the *ompA* promoter from *Flavobacterium johnsoniae*. *J Bacteriol* **189**: 5108-5118.
- Cock, J.M., Sterck, L., Rouze, P., Scornet, D., Allen, A.E., Amoutzias, G. et al. (2010) The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* **465**: 617-621.
- Condemine, G., and Robertbaudouy, J. (1991) Analysis of an *Erwinia chrysantemi* gene cluster involved in pectin degradation. *Mol Microbiol* **5**: 2191-2202.
- Cosse, A., Leblanc, C., and Potin, P. (2007) Dynamic defense of marine macroalgae against pathogens: from early activated to gene-regulated responses. *Adv Bot Res* **46**: 221-266.
- D'Elia, J.N., and Salyers, A.A. (1996) Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. *J Bacteriol* **178**: 7180-7176.

- Dodd, D., Mackie, R.I., and Cann, I.K. (2011) Xylan degradation, a metabolic property shared by rumen and human colonic *Bacteroidetes*. *Mol Microbiol* **79**: 292-304.
- Dodd, D., Moon, Y.H., Swaminathan, K., Mackie, R.I., and Cann, I.K. (2010) Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic *Bacteroidetes*. *J Biol Chem* **285**: 30261-30273.
- Duan, G., Han, F., and Yu, W. (2009) Cloning, sequence analysis, and expression of gene alyPI encoding an alginate lyase from marine bacterium *Pseudoalteromonas* sp. CY24. *Can J Microbiol* **55**: 1113-1118.
- Dunwell, J.M., Purvis, A., and Khuri, S. (2004) Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**: 7-17.
- Fineran, P.C., Everson, L., Slater, H., and Salmond, G.P. (2005) A GntR family transcriptional regulator (PigT) controls gluconate-mediated repression and defines a new, independent pathway for regulation of the tripyrrole antibiotic, prodigiosin, in *Serratia*. *Microbiology* **151**: 3833-3845.
- Gilbert, H.J. (2010) The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol* **153**: 444-455.
- Glöckner, F.O., and Joint, I. (2010) Marine microbial genomics in Europe: current status and perspectives. *Microb Biotechnol* **3**: 523-530.
- Groisillier, A., Herve, C., Jeudy, A., Rebuffet, E., Pluchon, P.F., Chevotot, Y. et al. (2010) MARINE-EXPRESS: taking advantage of high throughput cloning and expression strategies for the post-genomic analysis of marine organisms. *Microb Cell Fact* **9**: 45.
- Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696-704.
- Hashimoto, W., Kawai, S., and Murata, K. (2010) Bacterial supersystem for alginate import/metabolism and its environmental and bioenergy applications. *Bioengineered Bugs* **1**: 97-109.
- Haug, A., Larsen, B., and Smidsrød, O. (1974) Uronic acid sequence in alginate from different sources. *Carbohydr Res* **32**: 217-225.
- Haydon, D.J., and Guest, J.R. (1991) A new family of bacterial regulatory proteins. *FEMS Microbiol Lett* **79**: 291-296.
- Hehemann, J.H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**: 908-912.
- Holdeman, L.V., and Moore, W.E.C. (1974) New Genus, *Coprococcus*, Twelve New Species, and Emended Descriptions of Four Previously Described Species of Bacteria from Human Feces. *Int J Syst Bacteriol* **24**: 260-277.
- Hugouvieux-Cotte-Pattat, N., Condemine, G., Nasser, W., and Reverchon, S. (1996) Regulation of pectinolysis in *Erwinia chrysanthemi*. *Annu Rev Microbiol* **50**: 213-257.
- Itoh, T., Akao, S., Hashimoto, W., Mikami, B., and Murata, K. (2004) Crystal structure of unsaturated glucuronyl hydrolase, responsible for the degradation of glycosaminoglycan, from *Bacillus* sp. GL1 at 1.8 Å resolution. *J Biol Chem* **279**: 31804-31812.

- Izu, H., Adachi, O., and Yamada, M. (1997) Gene organization and transcriptional regulation of the gntRKU operon involved in gluconate uptake and catabolism of *Escherichia coli*. *J Mol Biol* **267**: 778-793.
- Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059-3066.
- Kehres, D.G., Zaharik, M.L., Finlay, B.B., and Maguire, M.E. (2000) The NRAMP proteins of *Salmonella typhimurium* and *Escherichia coli* are selective manganese transporters involved in the response to reactive oxygen. *Mol Microbiol* **36**: 1085–1100.
- Kingsford, C.L., Ayanbule, K., and Salzberg, S.L. (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* **8**: R22.
- Kirchman, D.L. (2002) The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol Ecol* **39**: 91-100.
- Kloareg, B., and Mabeau, S. (1987) Isolation and Analysis of the Cell Walls of Brown Algae: *Fucus spiralis*, *F. ceranoides*, *F. vesiculosus*, *F. serratus*, *Bifurcaria bifurcata* and *Laminaria digitata*. *J Exp Bot* **38**: 1573-1580.
- Koropatkin, N.M., Martens, E.C., Gordon, J.I., and Smith, T.J. (2008) Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* **16**: 1105-1115.
- Küpper, F.C., Kloareg, B., Guern, J., and Potin, P. (2001) Oligoguluronates elicit an oxidative burst in the brown algal kelp *Laminaria digitata*. *Plant Physiology* **125**: 278-291.
- Küpper, F.C., Müller, D.G., Peters, A.F., Kloareg, B., and Potin, P. (2002) Oligoalginate recognition and oxidative burst play a key role in natural and induced resistance of sporophytes of Laminariales. *Journal of Chemical Ecology* **28**: 2057-2081.
- Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P.M., and Henrissat, B. (2010) A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem J* **432**: 437-444.
- Maki, H., Mori, A., Fujiyama, K., Kinoshita, S., and Yoshida, T. (1993) Cloning, sequence analysis and expression in *Escherichia coli* of a gene encoding an alginate lyase from *Pseudomonas* sp. OS-ALG-9. *J Gen Microbiol* **139**: 987-993.
- Martens, E.C., Koropatkin, N.M., Smith, T.J., and Gordon, J.I. (2009) Complex glycan catabolism by the human gut microbiota: The *Bacteroidetes* Sus-like paradigm. *Journal of Biological Chemistry* **284**: 24673-24677.
- Meyer, F., Goesmann, A., McHardy, A.C., Bartels, D., Bekel, T., Clausen, J. et al. (2003) GenDB--an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* **31**: 2187-2195.
- Michel, G., Nyval-Collen, P., Barbeyron, T., Czjzek, M., and Helbert, W. (2006) Bioconversion of red seaweed galactans: a focus on bacterial agarases and carrageenases. *Applied Microbiology and Biotechnology* **71**: 23-33.
- Michel, G., Tonon, T., Scornet, D., Cock, J.M., and Kloareg, B. (2010a) The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytol* **188**: 82-97.

- Michel, G., Tonon, T., Scornet, D., Cock, J.M., and Kloareg, B. (2010b) Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in Eukaryotes. *New Phytol* **188**: 67-81.
- Neelakanta, G., Sankar, T.S., and Schnetz, K. (2009) Characterization of a beta-glucoside operon (bgc) prevalent in septicemic and uropathogenic *Escherichia coli* strains. *Appl Environ Microbiol* **75**: 2284-2293.
- Nielsen, H., Emanuelsson, O., Brunak, S., and von Heijne, G. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* **2**: 953-971.
- Norderhaug, K.M., Fredriksen, S., and Nygaard, K. (2003) Trophic importance of *Laminaria hyperborea* to kelp forest consumers and the importance of bacterial degradation to food quality. *Mar Ecol Prog Ser* **255**: 135-144.
- Ochiai, A., Hashimoto, W., and Murata, K. (2006) A biosystem for alginate metabolism in *Agrobacterium tumefaciens* strain C58: molecular identification of Atu3025 as an exotype family PL-15 alginate lyase. *Research in Microbiology* **157**: 642-649.
- Ochiai, A., Yamasaki, M., Mikami, B., Hashimoto, W., and Murata, K. (2010) Crystal structure of exotype alginate lyase Atu3025 from *Agrobacterium tumefaciens*. *J Biol Chem* **285**: 24519-24528.
- Popper, Z.A., Michel, G., Herve, C., Domozych, D.S., Willats, W.G., Tuohy, M.G. et al. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu Rev Plant Biol* **62**: 567-590.
- Preiss, J., and Ashwell, G. (1962a) Alginic acid metabolism in bacteria. II. The enzymatic reduction of 4-deoxy-L-erythro-5-hexoseulose uronic acid to 2-keto-3-deoxy-D-gluconic acid. *J Biol Chem* **237**: 317-321.
- Preiss, J., and Ashwell, G. (1962b) Alginic acid metabolism in bacteria. I. Enzymatic formation of unsaturated oligosaccharides and 4-deoxy-L-erythro-5-hexoseulose uronic acid. *J Biol Chem* **237**: 309-316.
- Ramsey, D.M., and Wozniak, D.J. (2005) Understanding the control of *Pseudomonas aeruginosa* alginate synthesis and the prospects for management of chronic infections in cystic fibrosis. *Mol Microbiol* **56**: 309-322.
- Rebuffet, E., Groisillier, A., Thompson, A., Jeudy, A., Barbeyron, T., Czjzek, M., and Michel, G. (2011) Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ Microbiol* **13**: 1253-1270.
- Salaun, S., Kervarec, N., Potin, P., Haras, D., Piotto, M., and La Barre, S. (2010) Whole-cell spectroscopy is a convenient tool to assist molecular identification of cultivatable marine bacteria and to investigate their adaptive metabolism. *Talanta* **80**: 1758-1770.
- Smith, S.V. (1981) Marine Macrophytes as a Global Carbon Sink. *Science* **211**: 838-840.
- Sonnhammer, E.L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**: 175-182.
- Staufenberger, T., Thiel, V., Wiese, J., and Imhoff, J.F. (2008) Phylogenetic analysis of bacteria associated with *Laminaria saccharina*. *FEMS Microbiol Ecol* **64**: 65-77.

- Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**: 207-234.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol.*
- Thomas, F., Barbeyron, T., and Michel, G. (2011a) Evaluation of reference genes for real-time quantitative PCR in the marine flavobacterium *Zobellia galactanivorans*. *J Microbiol Methods* **84**: 61-66.
- Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M., and Michel, G. (2011b) Environmental and gut *Bacteroidetes*: the food connection. *Frontiers in Microbiology* **2**: 93.
- Tujula, N.A., Crocetti, G.R., Burke, C., Thomas, T., Holmstrom, C., and Kjelleberg, S. (2010) Variability and abundance of the epiphytic bacterial community associated with a green marine Ulvacean alga. *Isme J* **4**: 301-311.
- Watanabe, Y., Nagai, F., Morotomi, M., Sakon, H., and Tanaka, R. (2010) *Bacteroides clarus* sp. nov., *Bacteroides fluxus* sp. nov. and *Bacteroides oleiciplenus* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* **60**: 1864-1869.
- Whittaker, R.H., and Likens, G.E. (1973) Carbon in Biota. *Brookhaven Symposia in Biology*: 281-302.
- Wiethaus, J., Schubert, B., Pfander, Y., Narberhaus, F., and Masepohl, B. (2008) The GntR-like regulator TauR activates expression of taurine utilization genes in *Rhodobacter capsulatus*. *J Bacteriol* **190**: 487-493.
- ZoBell, C. (1941) Studies on marine bacteria I The cultural requirements of heterotrophic aerobes. *J Mar Res* **4**: 42-75.

## Supporting information

**Table S1: Primers used in this study.**

**Table S2: Validation data for qPCR primer pairs.**

**Table S3: Quantification of the transcript levels for the 21 genes analyzed.**

**Table S4: Census of genes of the Alginate Utilization System in different species.**

### **Figure S1: Geographical distribution of the species containing an Alginate Utilization System**

Each color-coded spot locates the site of isolation of the sequenced strain. Zg, *Zobellia galactanivorans*; Cal, *Cellulophaga algicola*; Cly, *Cellulophaga lytica*; Gf, *Gramella forsetii*; Ka, *Kordia algicida*; Ca, *Croceibacter atlanticus*; Dd, *Dokdonia donghaensis*; ALC-1, *Flavobacteriales* bacterium ALC-1; 2A, *Flavobacteria* MS024-2A; 3C, *Flavobacteria* MS024-3C; SCB49, *Bacteroidetes* SCB49; Pd, *Polaribacter dokdonensis* MED152; BBFL7, *Flavobacteria* bacterium BBFL7; Ps, *Pedobacter saltans*; Rm, *Rhodothermus marinus*; Bc, *Bacteroides clarus*; Be, *Bacteroides eggerthii*; Mm, *Maricaulis maris*; Hb, *Hirschia baltica*; Bs, *Brevundimonas subvibrioides*; Ph, *Pseudoalteromonas haloplanktis*; Pa, *Pseudoalteromonas atlantica*; Am, *Alteromonas macleodii*; Sd, *Saccharophagus degradans* 2-40; TW7, *Alteromonadales* bacterium TW-7; Sf, *Shewanella frigidimarina*; K279a, *Stenotrophomonas maltophilia* K279a; Vs, *Vibrio sinaloensis*; SWAT3, *Vibrionales* bacterium SWAT-3; Va, *Vibrio alginolyticus*; Vh, *Vibrio harveyi* 1DA3; MED222, *Vibrio* sp. MED222; 12B01, *Vibrio splendidus* 12B01; LGP32, *Vibrio splendidus* LGP32.

**Figure S2: Phylogenetic trees of the protein families related to the alginolytic system.**











Name	Gene ID	Modularity	Signal peptide	Size (kDa)
AlyA1	zg1182		Lipoprotein	48.1
AlyA2	zg2618		Lipoprotein	35.1
AlyA3	zg2624		Cleaved	85.1
AlyA4	zg4130		Uncleavable (CM)	86.9
AlyA5	zg4131		Lipoprotein	39.3
AlyA6	zg4132		Cleaved	86.6
AlyA7	zg4327		Cleaved	44.0
-	zg3452		Uncleavable (CM)	87.5

Figure 1

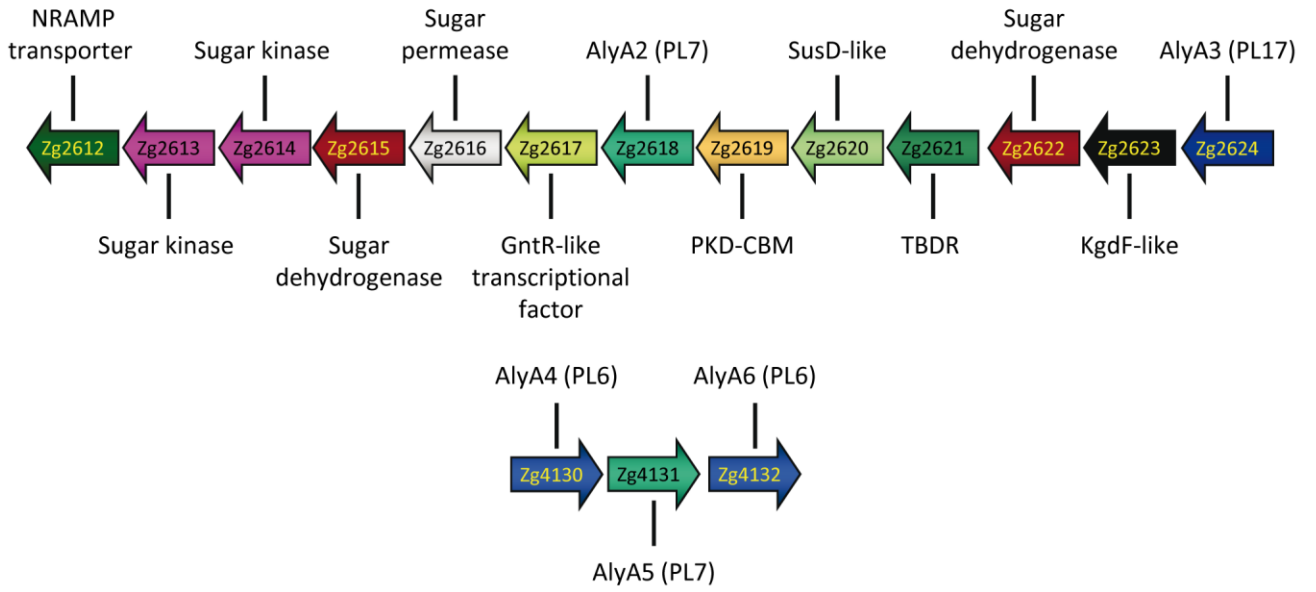


Figure 2

**A**

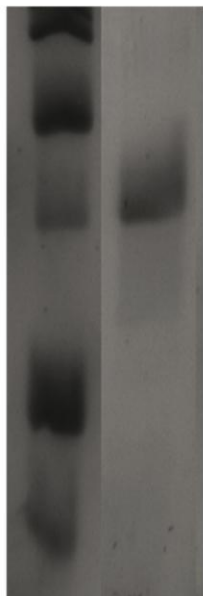
75 kDa -

50 kDa -

37 kDa -

25 kDa -

20 kDa -

**B**

Database: &gt;zobellia\_1182 alyA1

Protein MW: 48105

MKKNVFTTLRTRVVNGDIMWKLIPVFFLALCLGSCSEEPVD  
PEEEAVLTK**LSANSTAIGISSVSASTSQSPNVASNTLDGS**  
**TSTRWSGYGDGASITYDLGSSANIDYVKIAFYKGSRKTK**  
**YEVWVGNSTSSLTKIKSKTSSGSTSDYETIDLPNSTARYI**  
**RIVGKGYVLNSGGSTVLWNSITKFQAWGSGGSSTLPISGN**  
**SPASVLGITANTWK**INSFIGSPGSSATYYDDITDASGISY  
NTYSDDNYFYTDGEWVYFKCYR**GLGGSANSQNPRVELREM**  
DNGNLASWTGDSGTHMEWTVQVNQLPQD TDGDGGVLCFG  
QIHGPSK**NSDGVEVDDVVRVQFIGEENQSSGSVKLK**ISGY  
VTEEQGGSQTFSGYSLD TTYNCK**LVYSGGYVELFMNGSSV**  
**FRKKMEVDDLSENYFKVGNYLQSVKGASYTGSYGLVRIKN**  
**LSVTHN**

Figure 3

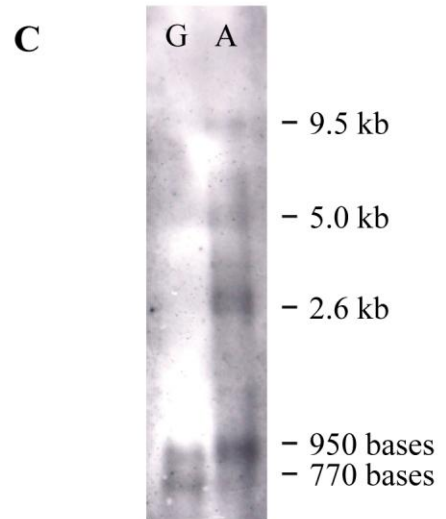
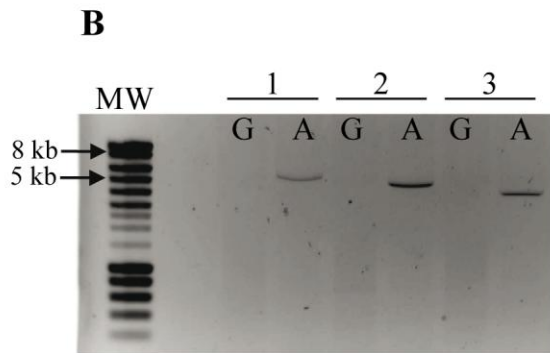
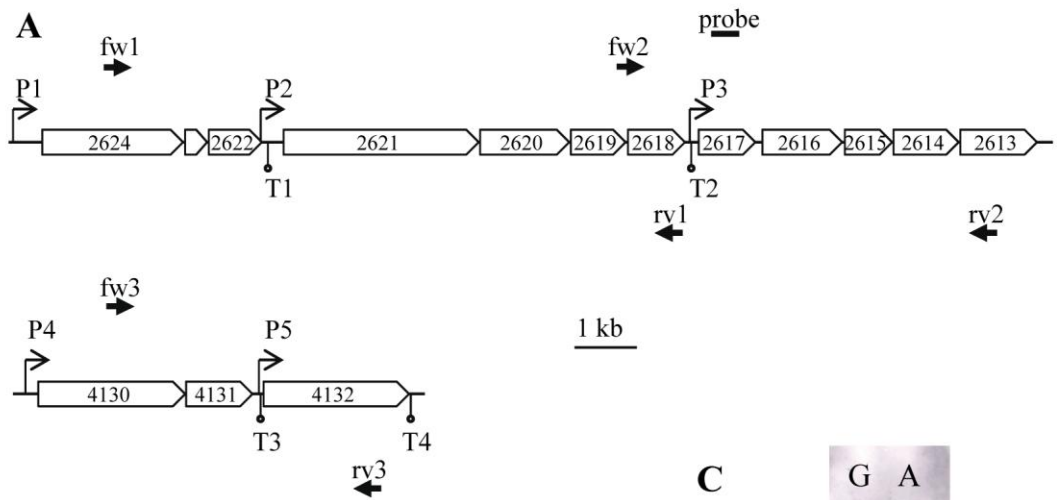


Figure 4

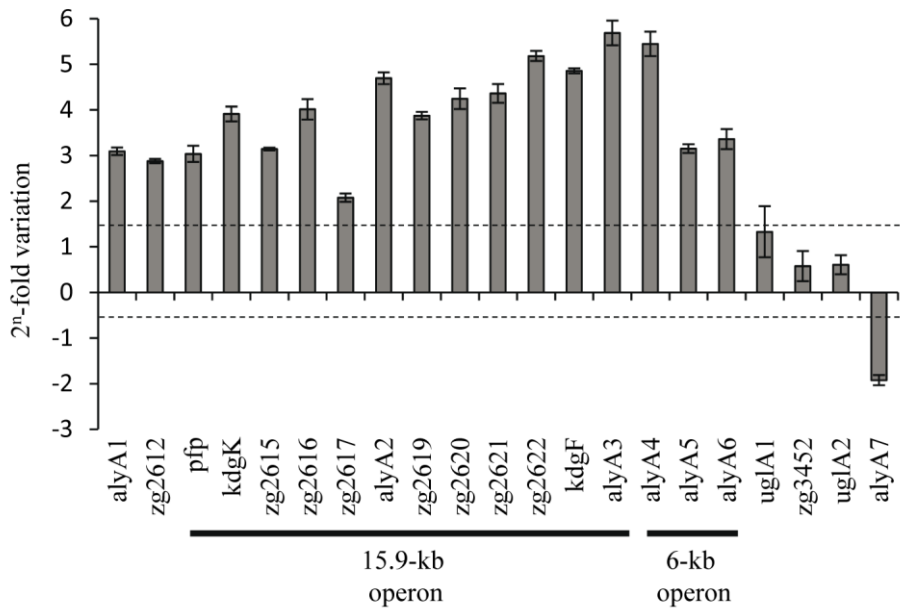


Figure 5

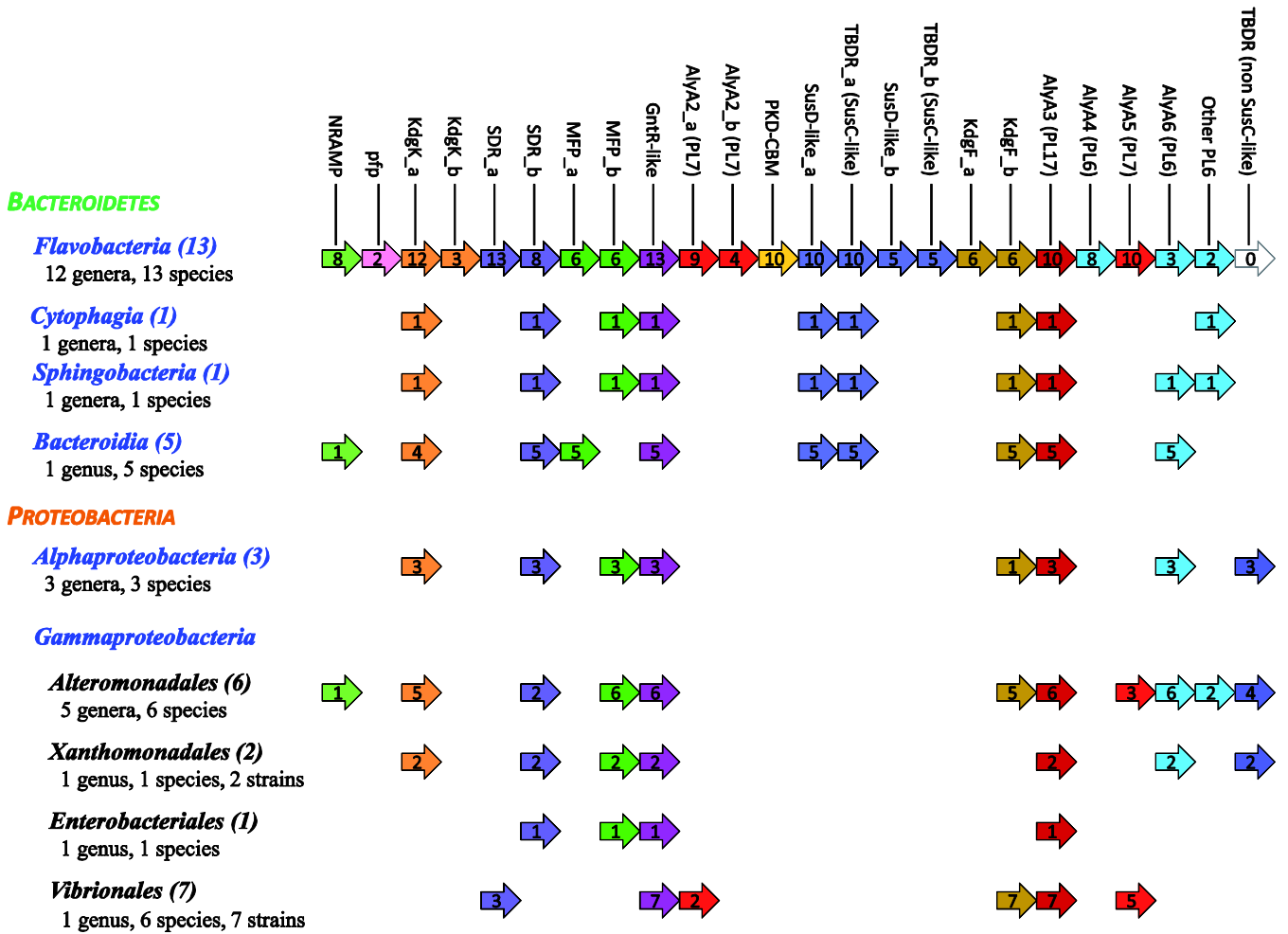


Figure 6

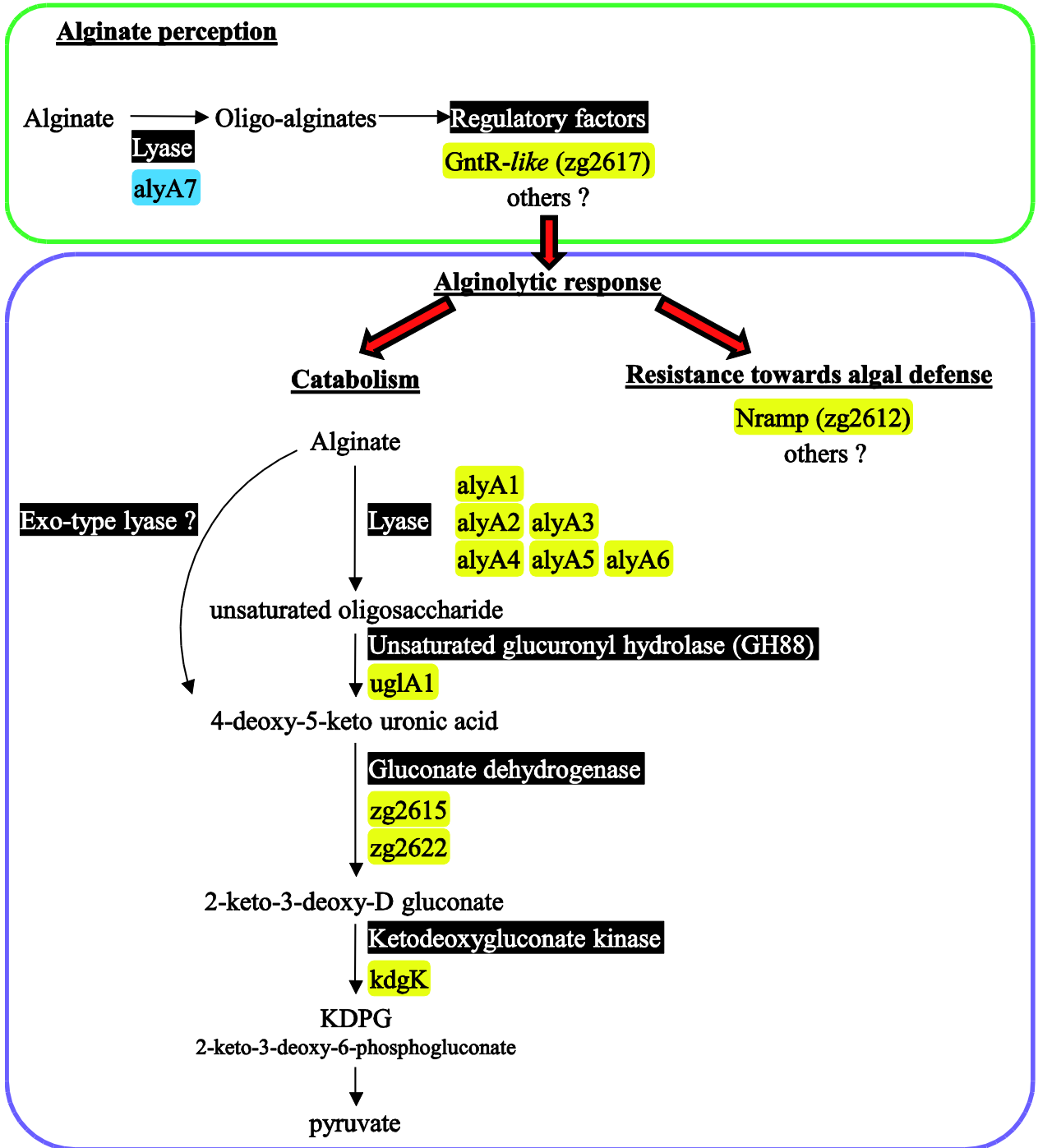


Figure 7

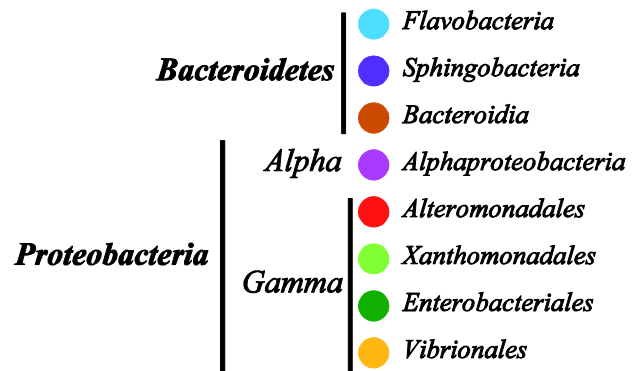
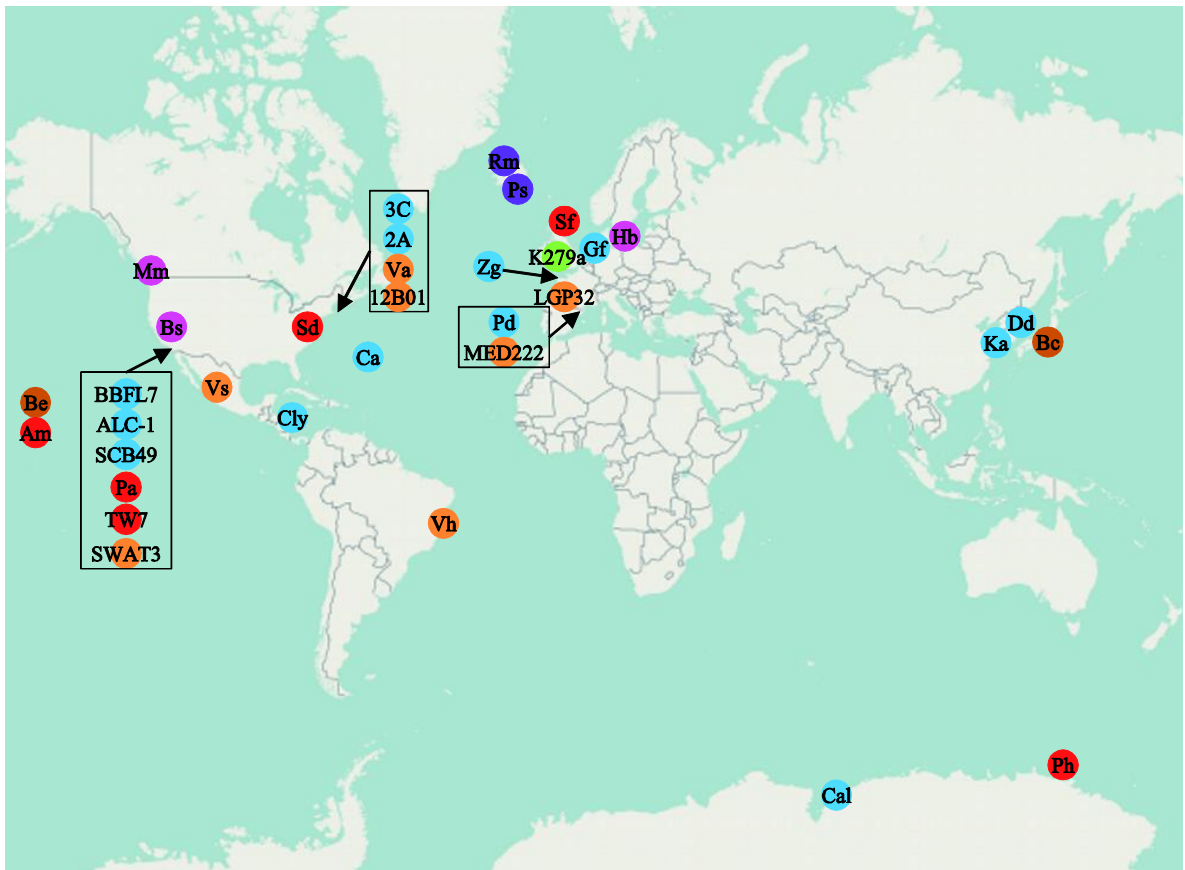
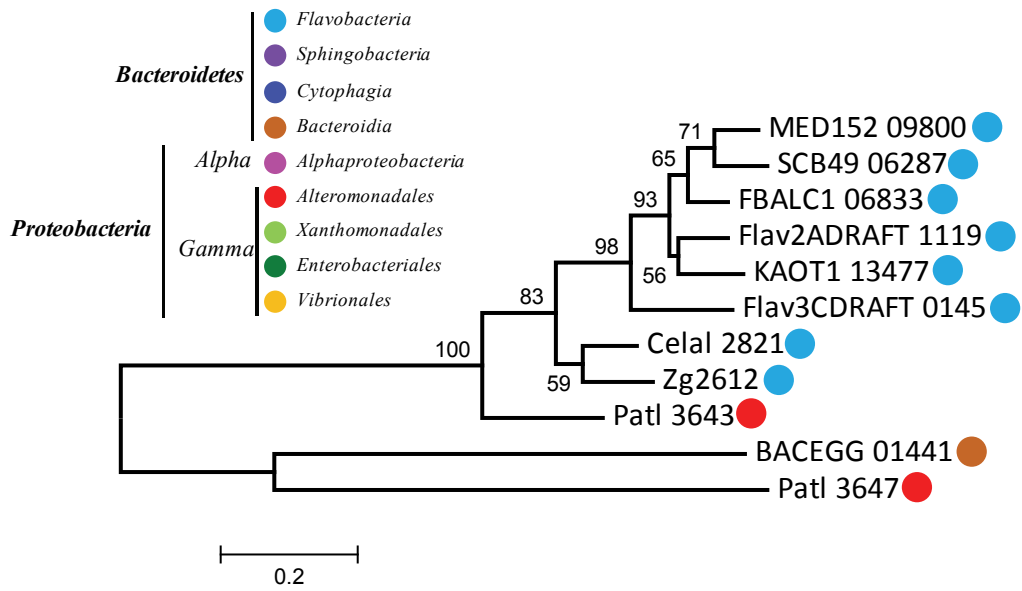
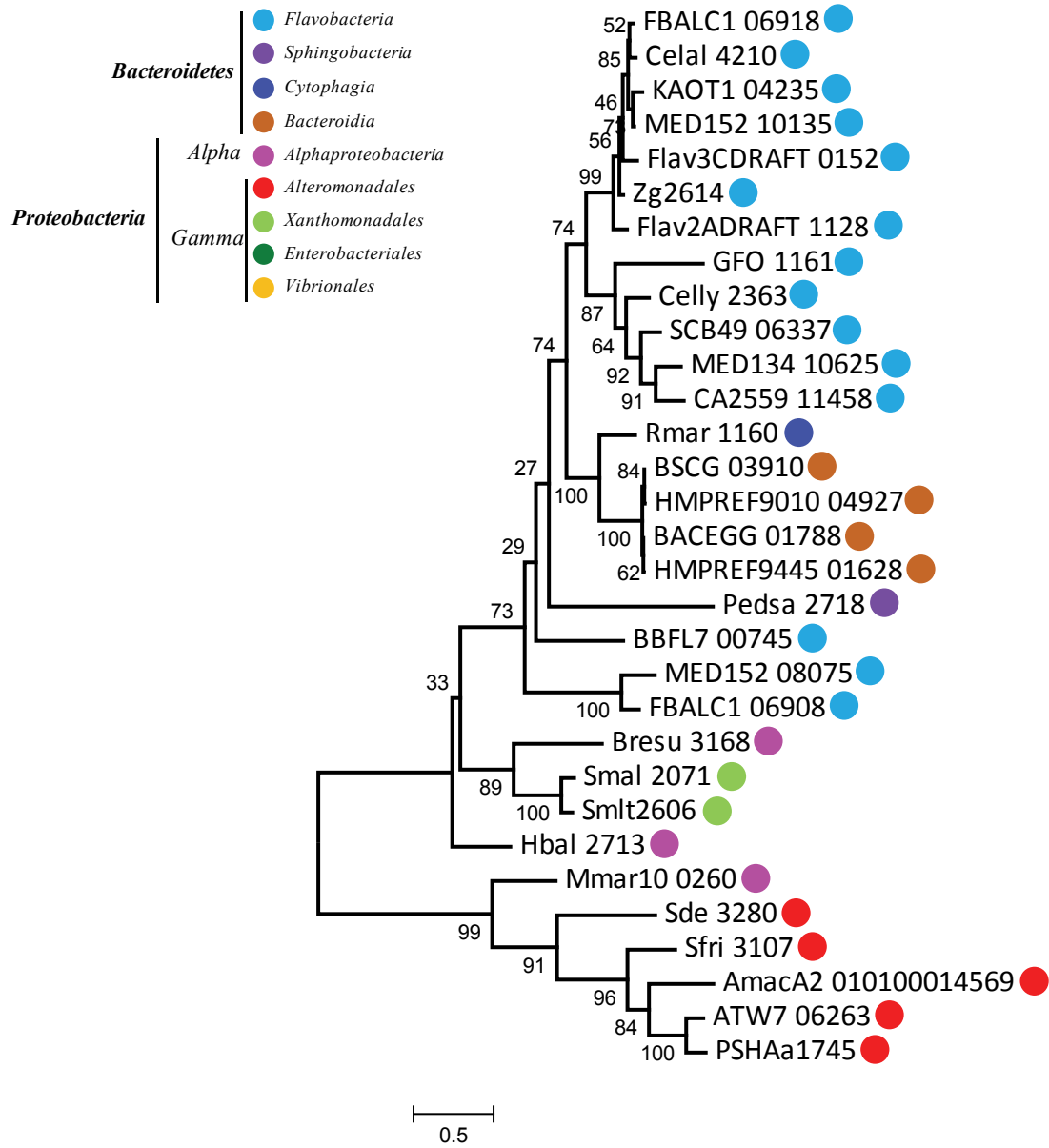


Figure S1

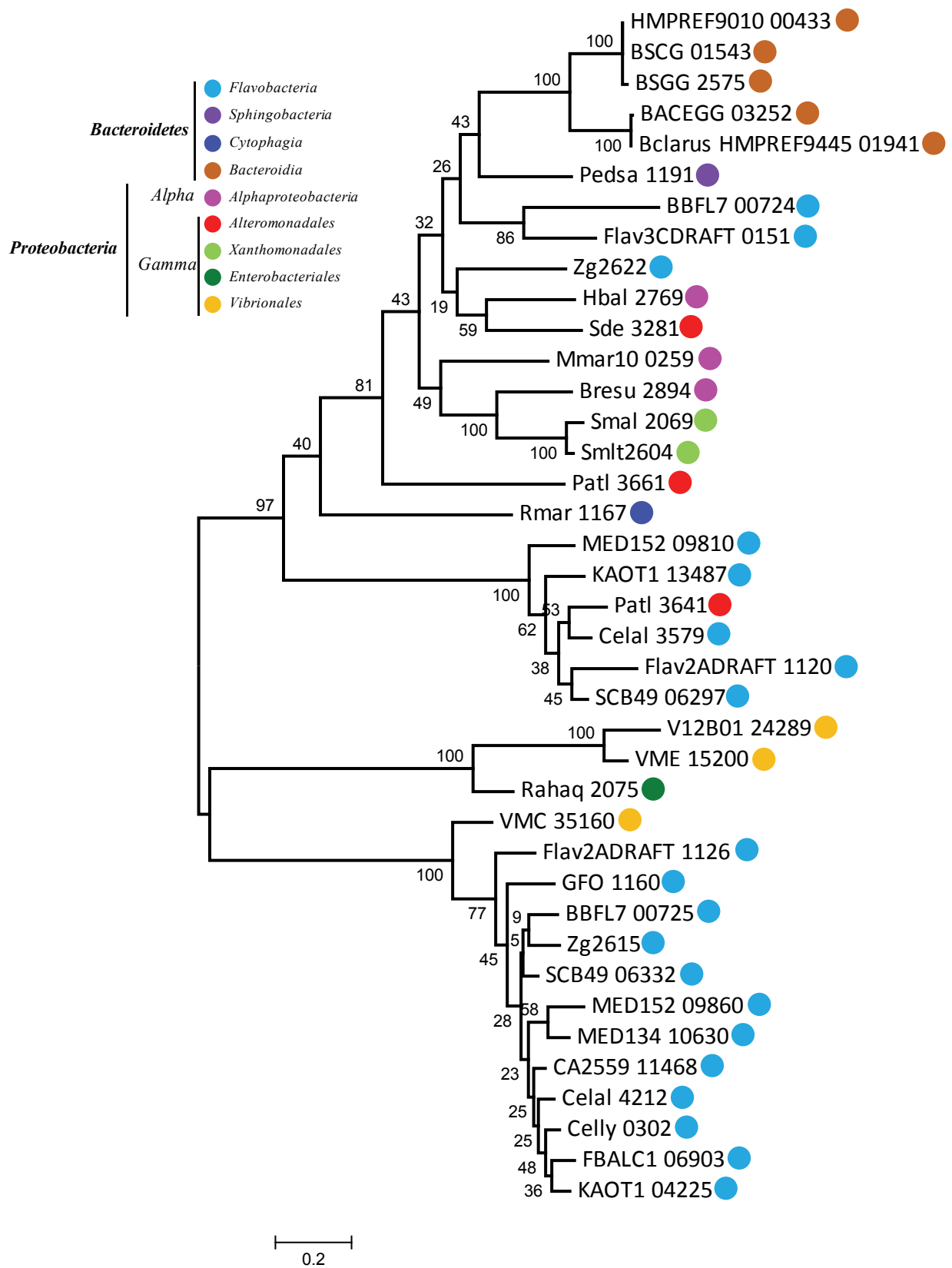




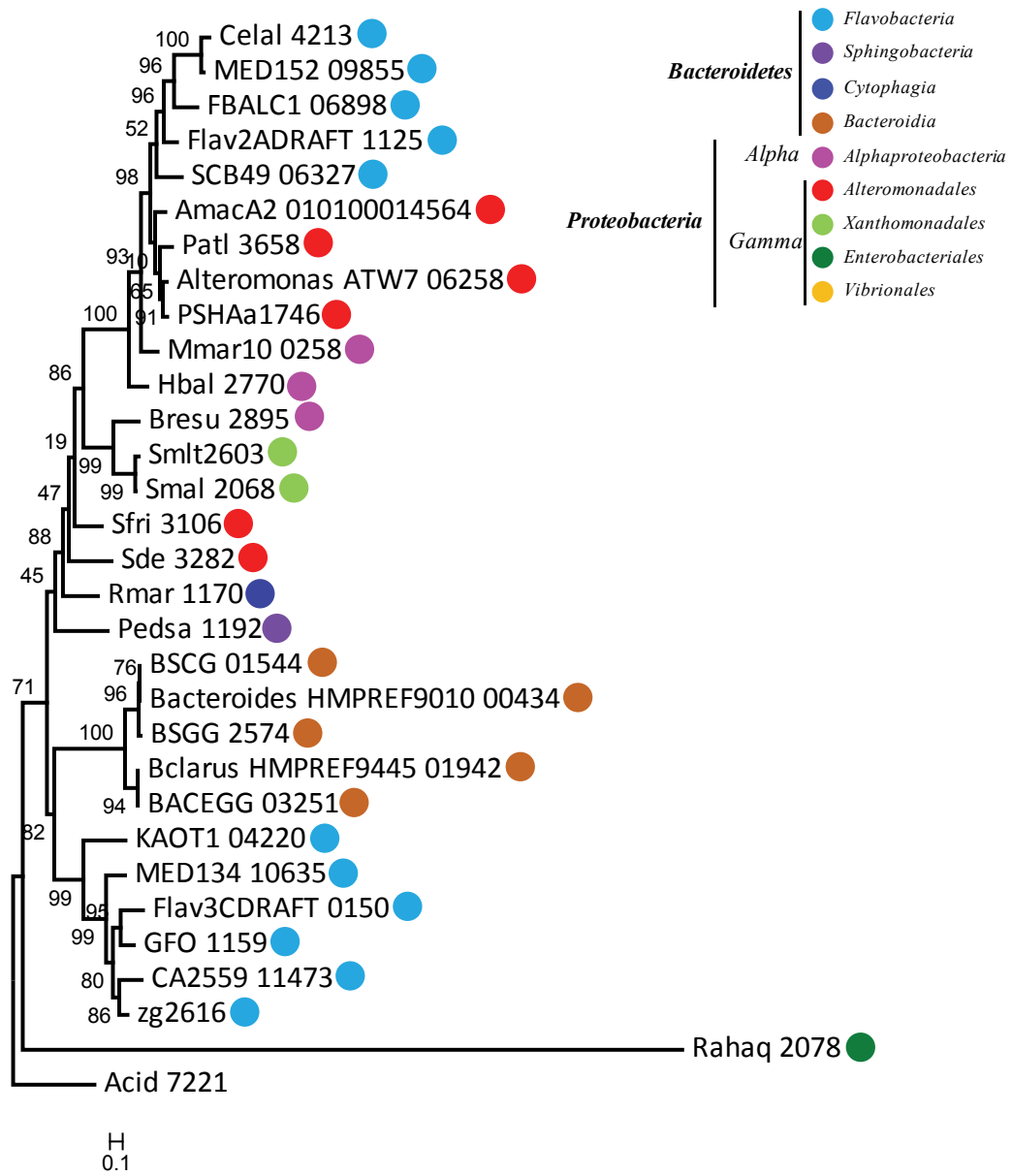
NRAMP (Zg2612)



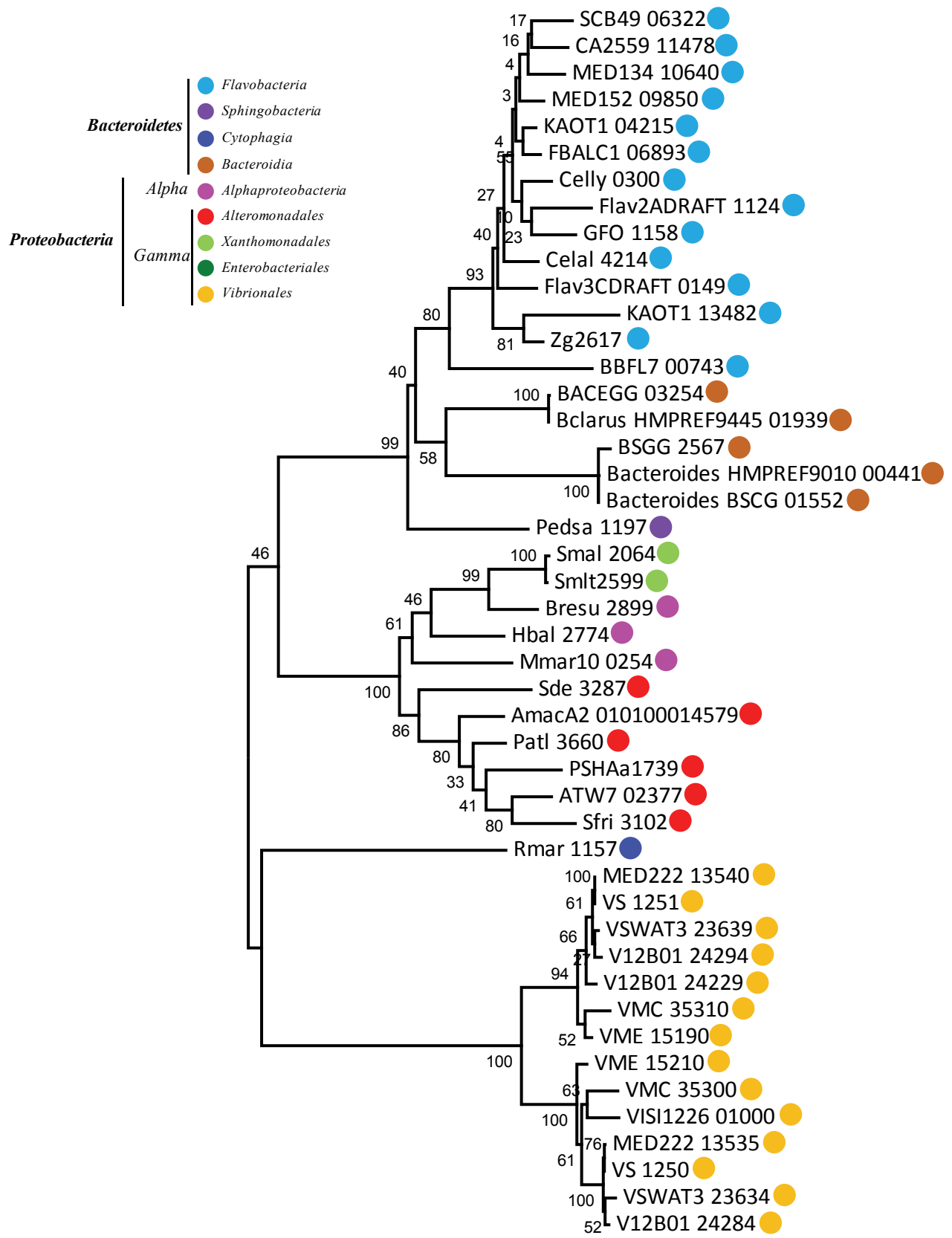
KdgK (Zg2614)



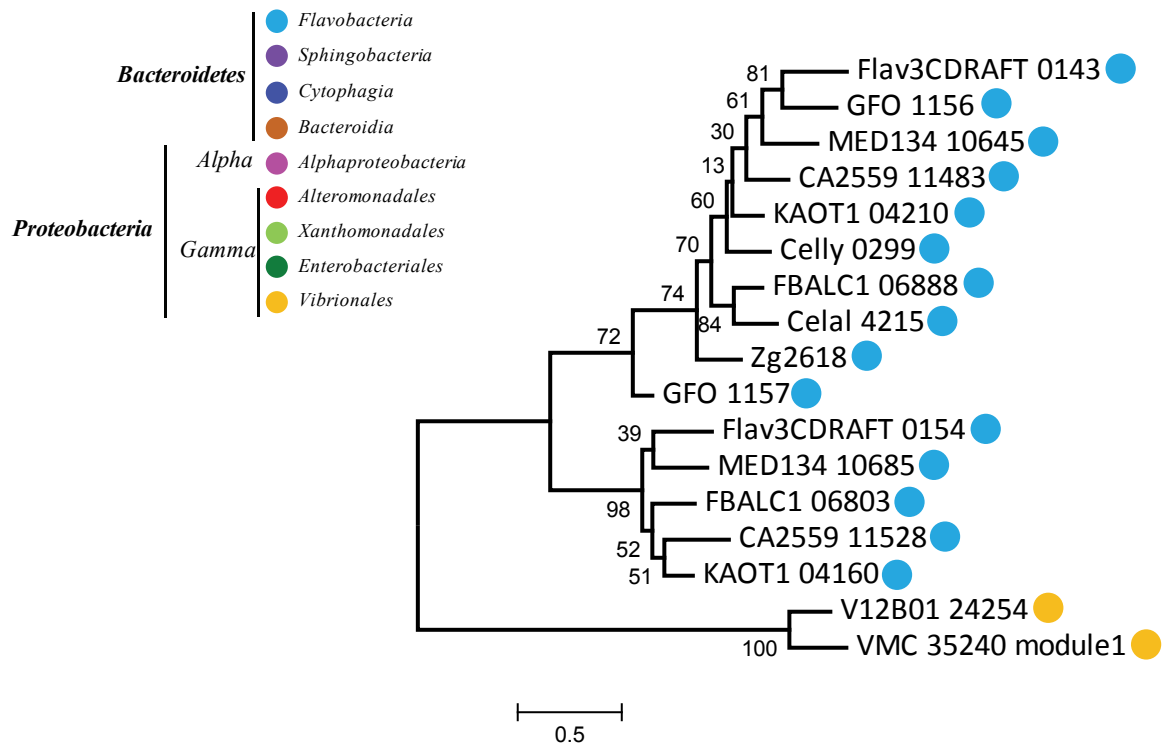
SDR (Zg2615, Zg2622)



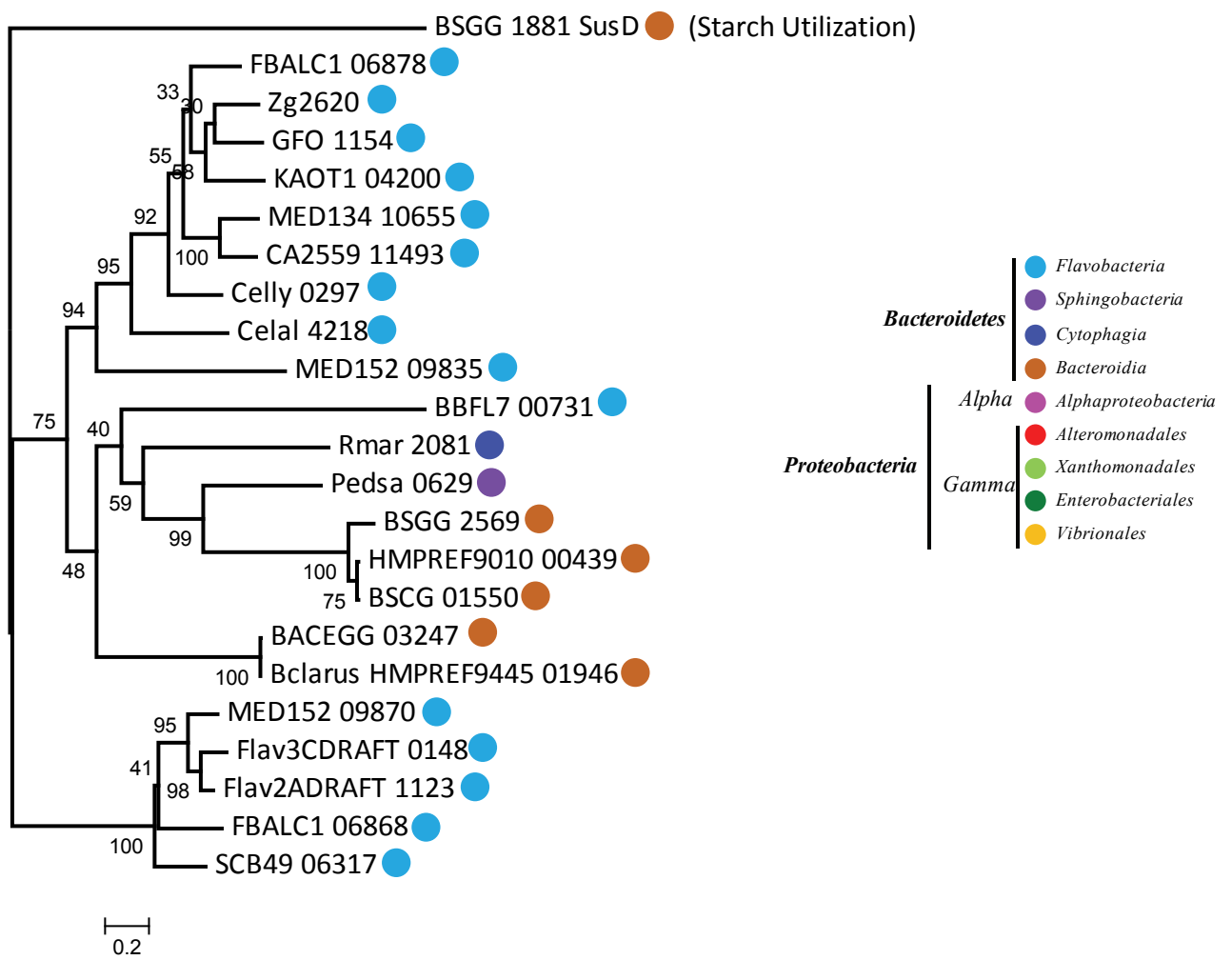
MFP(Zg2616)



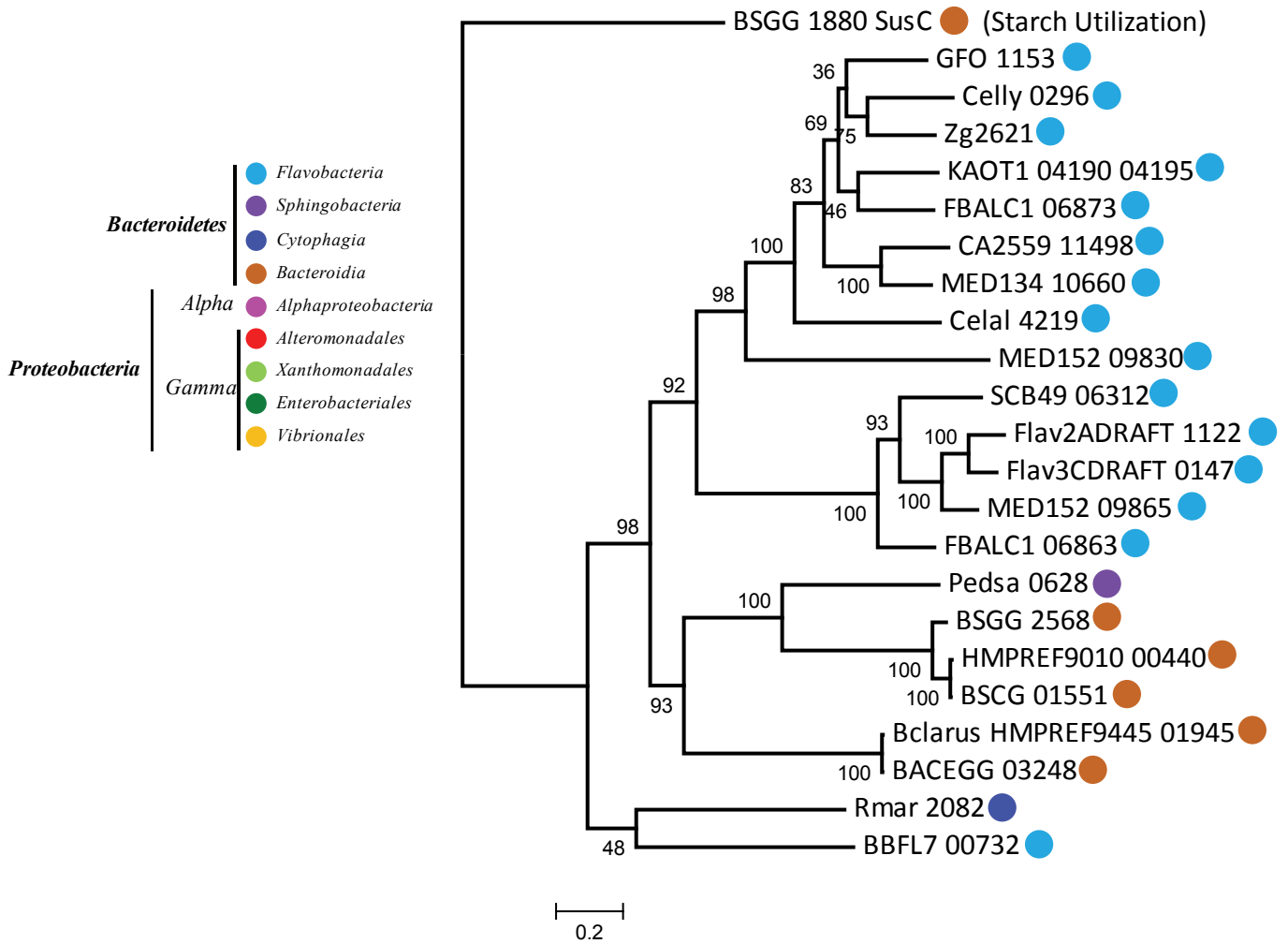
0.2  
GntR-like(Zg2617)



AlyA2 (Zg2618)

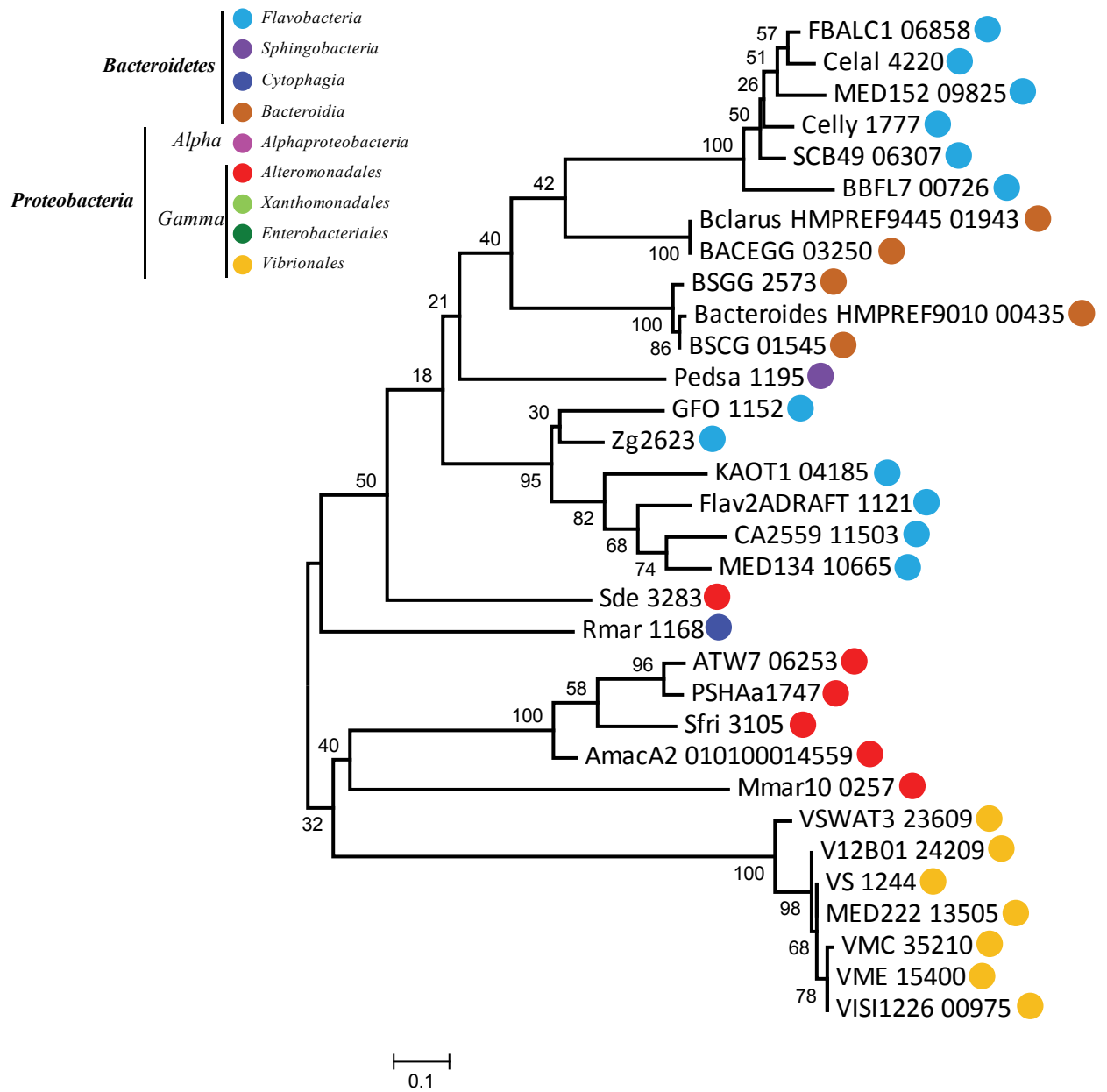


SusD-like (Zg2620)

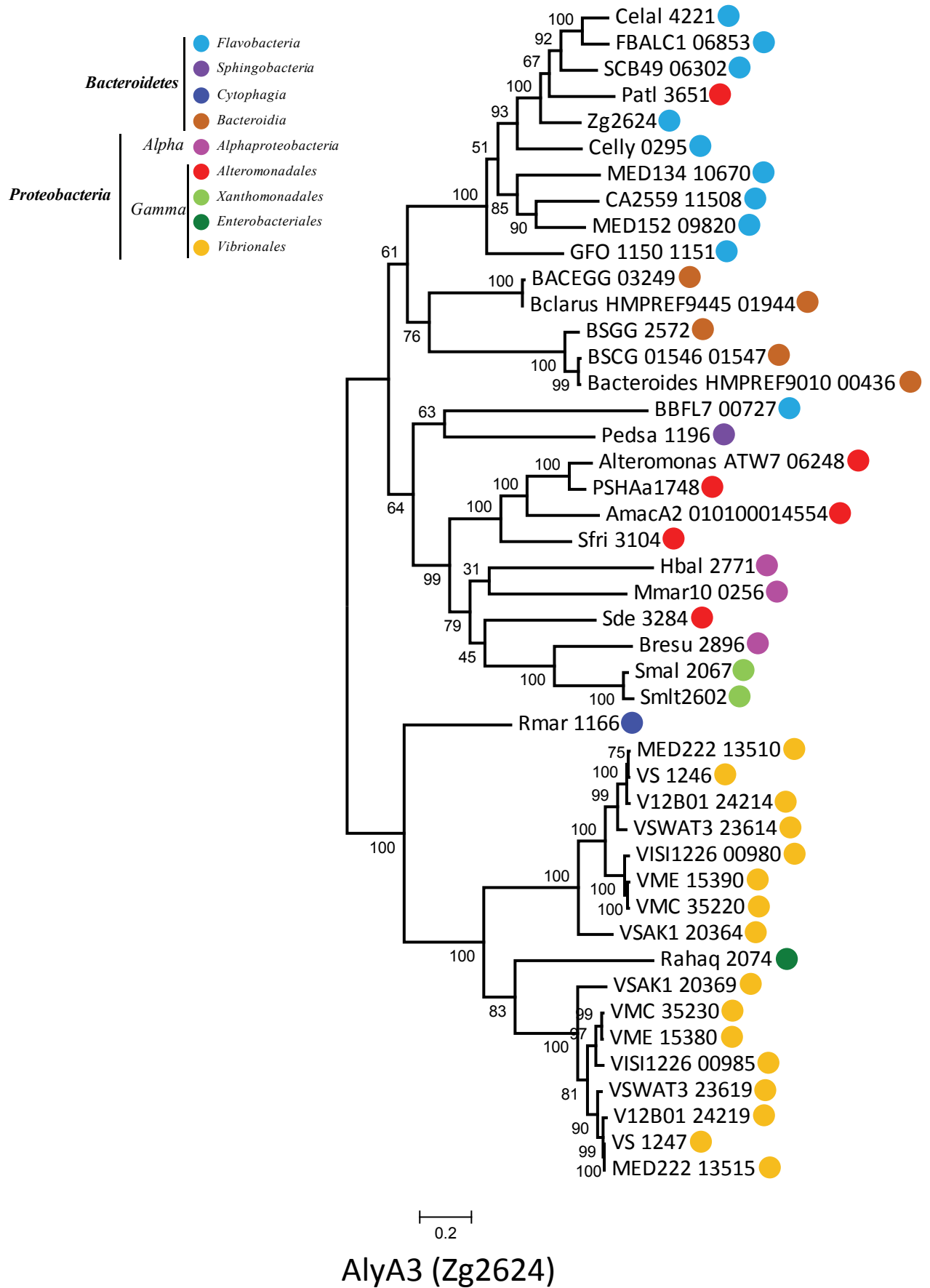


TBDR SusC-like (Zg2621)



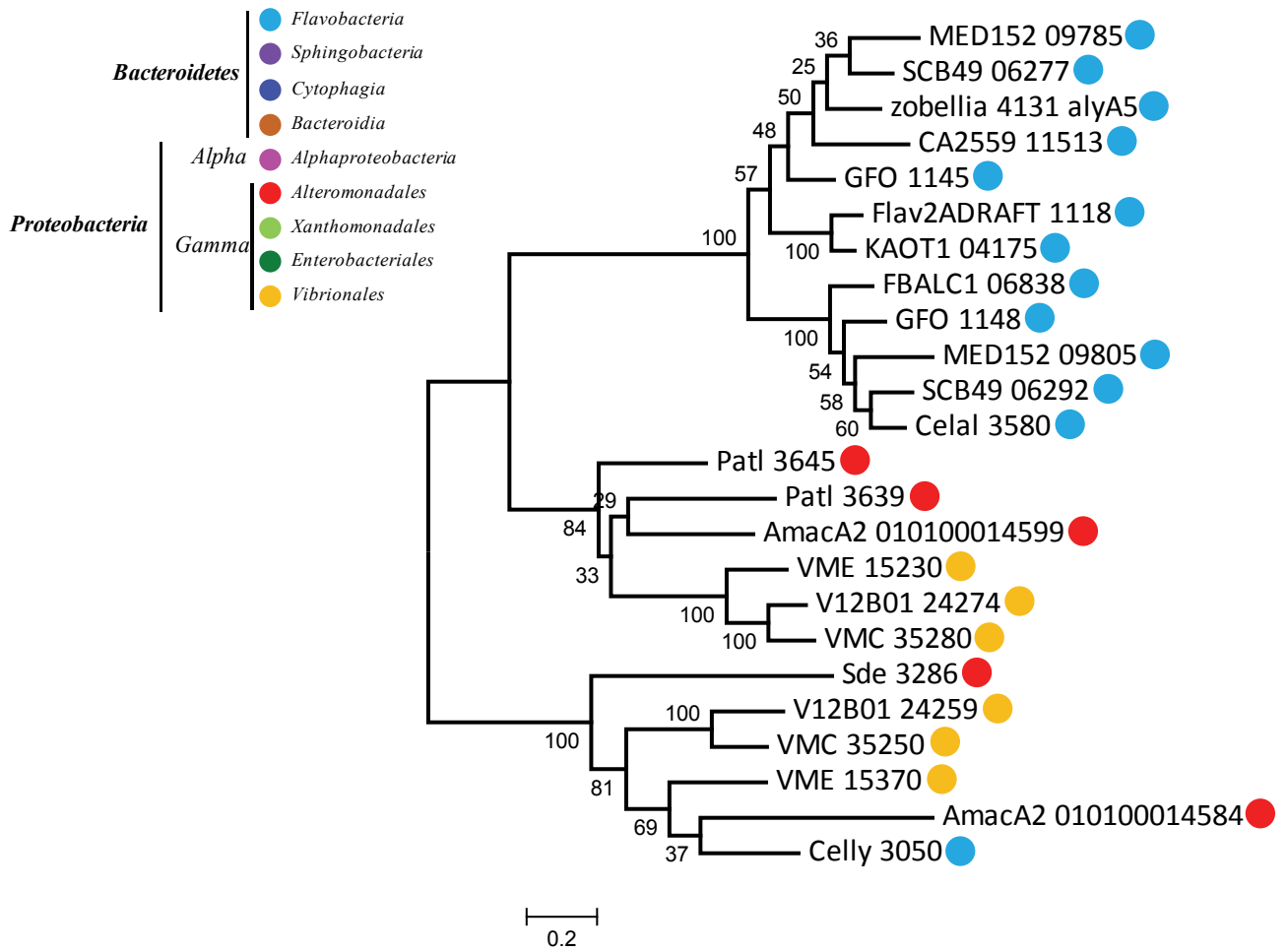


KdgF(Zg2623)





AlyA4 (Zg4130), AlyA6 (Zg4132) and other PL6



AlyA5 (Zg4131)

**Table S1: Primers used for the transcriptomic studies**

Target gene(s)	Primer sequence (5'-3') forward (top)-reverse (bottom)	Product length (bp)	Use
<i>glyA</i>	GCTTTGATCTATCCCACG CATCGTAATCAAGGGTGC	120	RT-qPCR normalisation
<i>gmkA</i>	TATGCGAATAGCGAAGG AAATCAGCCACCAACTC	117	RT-qPCR normalisation
<i>icdA</i>	GTAAAGGCTTACCTTCCC CGGTAGCTTCTTTAGGTG	73	RT-qPCR normalisation
<i>zg2624</i>	ACGCTTCTATACCCGAC CCCTTCTTAGGTTGAAGTC	95	RT-qPCR
<i>zg2623</i>	GTCAGATTATGGGCTACG CTTGTGAGTGGTAGTGC	96	RT-qPCR
<i>zg2622</i>	CTCAACTCCACCTTCTG GTTTATGATCGAGGCACC	69	RT-qPCR
<i>zg2621</i>	GGCTCCGTTGATTATTACC CTGTAGTCACTGAAATCCG	70	RT-qPCR
<i>zg2620</i>	CTATGTAACCAAGTGAAGCC TACCTGATACTGACCTACC	103	RT-qPCR
<i>zg2619</i>	CTACACGGTTACCTTGAC CGCTAGTGGAATTTTCAGG	94	RT-qPCR
<i>zg2618</i>	GTACGGGTAAAGACTAAGG GGTATATCCCTCATCATCG	93	RT-qPCR
<i>zg2617</i>	CTCAGAAGAACCGAGGA TCAAGGCAATCATCACC	89	RT-qPCR
<i>zg2616</i>	CTCTTTGTAGGATGGATGC CAATACTTCCTGCCATACC	112	RT-qPCR
<i>zg2615</i>	CCGGTGTAGTATCTTTCCCT GATTCCTAAAGCTTGCCC	89	RT-qPCR
<i>zg2614</i>	CCGACAGTATCAGATTACC GACCGTATATCAATCCACC	77	RT-qPCR
<i>zg2613</i>	GAACCGCAAATCAAGGA TCAAAGTAGATCCGCCA	70	RT-qPCR
<i>zg2612</i>	GTAAACACCCTTGAGCC CCCGCACTTAATGTACC	71	RT-qPCR
<i>zg4130</i>	CGAAACCATACTAAGGG GAAATGACTTCCACCTCTC	105	RT-qPCR
<i>zg4131</i>	CCAGAACGGAATTAGGG TGCATTCAACTTCCCTC	62	RT-qPCR
<i>zg4132</i>	GACACAGAAGGTAGCAG TCGCCACCATTATTTCC	86	RT-qPCR
<i>zg1182</i>	ATCACCGAATGTAGCCA CCATCACATAACCTGAC	69	RT-qPCR
<i>zg4327</i>	CAAGTTACGAAGGGCA CAGCCTCGCATAATACC	86	RT-qPCR
<i>zg3452</i>	CTTCTACATACCGATGCC CCCAACTCTTACGAACC	76	RT-qPCR
<i>zg3445</i>	CCGCTAAAGACCTAATCG GTGGAATGCAGTAGAAGTG	82	RT-qPCR
<i>zg3804</i>	ACAAGACCACCAATCAC AAGGTATCCGTTACCGT	92	RT-qPCR
<i>zg2619 to zg2624</i>	CTAGGTTCCGATAACGGC TGTCTGAATTCGCATAGGG	7734	RT-PCR
<i>zg2613 to zg2619</i>	CACAAGAATCAGGAACCGT GATCAAAGTAGATCCGCCA	5706	RT-PCR
<i>zg4130 to zg4132</i>	CCAATCACTCCCGAATTACC AGGTAGGAATCAAGAAGCCC	4444	RT-PCR
<i>zg2617</i>	GGTGTAGCAGAAGTAACG ATCTTGGTCATAGGCCAC	413	DNA probe synthesis

**Table S2: Validation of qPCR primer pairs**

accession number	calibration curve	efficiency (%)
zg2612	$y = -3.61x + 38.49$ $R^2 = 0.979$	90%
zg2613	$y = -3.55x + 35.41$ $R^2 = 0.999$	91%
zg2614	$y = -3.55x + 36.04$ $R^2 = 0.996$	91%
zg2615	$y = -3.44x + 34.35$ $R^2 = 1.000$	95%
zg2616	$y = -3.44x + 34.35$ $R^2 = 1.000$	95%
zg2617	$y = -3.32x + 34.50$ $R^2 = 0.997$	100%
zg2618	$y = -3.46x + 35.60$ $R^2 = 0.998$	95%
zg2619	$y = -3.45x + 34.58$ $R^2 = 0.996$	95%
zg2620	$y = -3.44x + 36.44$ $R^2 = 1.000$	95%
zg2621	$y = -3.38x + 34.22$ $R^2 = 0.999$	98%
zg2622	$y = -3.54x + 35.75$ $R^2 = 0.999$	92%
zg2623	$y = -3.46x + 34.57$ $R^2 = 0.998$	95%
zg2624	$y = -3.41x + 34.43$ $R^2 = 0.988$	97%
zg4130	$y = -3.42x + 34.81$ $R^2 = 0.994$	96%
zg4131	$y = -3.36x + 35.63$ $R^2 = 0.996$	99%
zg4132	$y = -3.61x + 40.82$ $R^2 = 0.982$	90%
zg1182	$y = -3.44x + 35.78$ $R^2 = 1.000$	95%
zg4327	$y = -3.36x + 34.40$ $R^2 = 0.999$	98%
zg3452	$y = -3.59x + 34.99$ $R^2 = 0.985$	90%
zg3445	$y = -3.60x + 36.68$ $R^2 = 0.992$	90%
zg3804	$y = -3.45x + 37.31$ $R^2 = 0.991$	95%

**Table S3: Quantification of the transcript levels for the 21 genes analyzed.**

Before RNA extraction, *Z. galactanivorans* was grown in Zobell medium or in mineral medium supplemented either with glucose or alginate.

accession number	putative fonction	expression level (mean $\pm$ s.e.m, n=3)*			2 <sup>n</sup> -fold variation** alginate vs. glucose
		Zobell	Glucose	Alginate	
zg1182	alyA1	0.57 $\pm$ 0.05	0.12 $\pm$ 0.02	1.01 $\pm$ 0.10	3.1
zg2612	NRAMP	n.d.	0.08 $\pm$ 0.03	0.60 $\pm$ 0.03	2.9
zg2613	pfp	n.d.	0.20 $\pm$ 0.00	1.69 $\pm$ 0.33	3.0
zg2614	kdgK	0.10 $\pm$ 0.00	0.25 $\pm$ 0.00	3.93 $\pm$ 0.81	3.9
zg2615	sdr	0.14 $\pm$ 0.01	0.31 $\pm$ 0.03	2.71 $\pm$ 0.10	3.1
zg2616	sugar permease	0.03 $\pm$ 0.00	0.10 $\pm$ 0.02	1.77 $\pm$ 0.52	4.0
zg2617	GntR-like	1.22 $\pm$ 0.11	0.95 $\pm$ 0.09	4.04 $\pm$ 0.45	2.1
zg2618	alyA2	0.02 $\pm$ 0.00	0.03 $\pm$ 0.00	0.78 $\pm$ 0.11	4.7
zg2619	PKD/CBM	0.07 $\pm$ 0.00	0.09 $\pm$ 0.01	1.40 $\pm$ 0.14	3.9
zg2620	susD	0.08 $\pm$ 0.00	0.07 $\pm$ 0.01	1.46 $\pm$ 0.37	4.2
zg2621	tldr	0.03 $\pm$ 0.00	0.06 $\pm$ 0.01	1.28 $\pm$ 0.32	4.4
zg2622	sdr	0.14 $\pm$ 0.01	0.18 $\pm$ 0.01	6.79 $\pm$ 0.94	5.2
zg2623	kdgF	0.12 $\pm$ 0.01	0.17 $\pm$ 0.01	5.00 $\pm$ 0.34	4.9
zg2624	alyA3	0.02 $\pm$ 0.00	0.02 $\pm$ 0.00	1.19 $\pm$ 0.37	5.7
zg4130	alyA4	0.08 $\pm$ 0.01	0.04 $\pm$ 0.00	1.86 $\pm$ 0.64	5.4
zg4131	alyA5	0.31 $\pm$ 0.02	0.11 $\pm$ 0.01	1.02 $\pm$ 0.12	3.2
zg4132	alyA6	0.10 $\pm$ 0.01	0.09 $\pm$ 0.02	1.01 $\pm$ 0.24	3.4
zg3445	uglA1	n.d.	0.01 $\pm$ 0.00	0.03 $\pm$ 0.02	1.3
zg3452	PL15	n.d.	0.02 $\pm$ 0.00	0.03 $\pm$ 0.01	0.6
zg3804	uglA2	n.d.	0.01 $\pm$ 0.00	0.02 $\pm$ 0.00	0.6
zg4327	alyA7	0.17 $\pm$ 0.00	0.47 $\pm$ 0.07	0.13 $\pm$ 0.02	-1.9

\* transcripts levels were normalized against the geometric average of *glyA*, *icdA* and *gmkA* (Thomas et al., 2011)

\*\* expression level in alginate compared to glucose as sole carbon source

**Table S4: Census of genes of the Alginate Utilization System in the different species.**

<b>Marine <i>Flavobacteria</i></b>			
<b><i>Cellulophaga algicola</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Celal_4222	zg2624 (AlyA3)	poly(beta-d-mannuronate) lyase	69%
Celal_4221	zg2624 (AlyA3)	heparinase ii/iii family protein	69%
Celal_4220	zg2623 (KdgF)	hypothetical protein	50%
Celal_4219	zg2621 (TBDR)	tonb-dependent receptor plug	53%
Celal_4218	zg2620 (SusD-like)	ragb/susd domain-containing protein	46%
Celal_4217	zg2619 (PKD-CBM)	pkd domain containing protein	34%
Celal_4216	zg2619 (PKD-CBM)	hypothetical protein	25%
Celal_4215	zg2618 (AlyA2)	alginate lyase 2	72%
Celal_4214	zg2617 (GntR-like)	transcriptional regulator, gntR family	70%
Celal_4213	zg2616 (Sugar permease)	major facilitator superfamily mfs_1	44%
Celal_4212	zg2615 (Acetoin(diacetyl) reductase)	3-oxoacyl-(acyl-carrier-protein) reductase	86%
<i>Celal_4211</i>	<i>zg3348 (LacI-like)</i>	<i>transcriptional regulator, lacI family</i>	<i>38%</i>
Celal_4210	zg2614 (KdgK1)	pfkb domain protein	89%
Celal_4209	zg2613 (pfp)	6-phosphofructokinase	87%
Celal_3579	zg2622 (SDR)	3-oxoacyl-(acyl-carrier-protein) reductase	42%
Celal_3580	zg4131 (AlyA5)	alginate lyase 2	62%
Celal_2821	Zg2612	NRAMP	85%
<b><i>Cellulophaga lytica</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Celly_0294	zg4132 (AlyA6)	poly(beta-D-mannuronate) lyase	61%
Celly_0295	zg2624 (AlyA3)	Heparinase II/III family protein	65%
Celly_0296	zg2621 (TBDR)	TonB-dependent receptor plug	64%
Celly_0297	zg2620 (SusD-like)	RagB/SusD domain-containing protein	59%
Celly_0298	zg2619 (PKD-CBM)	PKD domain-containing protein	42%
Celly_0299	zg2618 (AlyA2)	Alginate lyase 2	62%
Celly_0300	zg2617 (GntR-like)	GntR domain-containing protein	66%
Celly_0301	zg2616 (Sugar permease)	major facilitator superfamily protein	47%
Celly_0302	zg2615 (Acetoin(diacetyl) reductase)	3-oxoacyl-(acyl-carrier-protein) reductase	86%
Celly_3050	Zg4131	poly(beta-D-mannuronate) lyase	29%
Celly_1777	Zg2623	Cupin 2 barrel domain-containing protein	52%
Celly_2363	Zg2614	Pfkb domain-containing protein	60%
<b><i>Gramella forsetii</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
GFO_1145	zg4131 (AlyA5)	secreted alginate lyase-like	71%
<i>GFO_1146</i>	<i>zg2626 (TIM barrel)</i>	<i>TIM barrel domain-containing protein</i>	<i>46%</i>
GFO_1147	zg4130 (AlyA4)	secreted alginate lyase-like	52%
GFO_1148	zg4131 (AlyA5)	secreted alginate lyase-like	61%
<i>GFO_1149</i>	<i>no homolog</i>	<i>hypothetical protein</i>	<i>-</i>
GFO_1150_1151	zg2624 (AlyA3)	chondroitin AC/alginate lyase domain-containing protein	55%
GFO_1152	zg2623 (KdgF)	pectin degradation protein	75%
GFO_1153	zg2621 (TBDR)	TonB-dependent outer membrane receptor	62%
GFO_1154	zg2620 (SusD-like)	SusD/RagB family protein	66%
GFO_1155	zg2619 (PKD-CBM)	PKD domain-containing protein	60%
GFO_1156	zg2618 (AlyA2)	secreted alginate lyase-like	63%
GFO_1157	zg2618 (AlyA2)	secreted alginate lyase-like	58%
GFO_1158	zg2617 (GntR-like)	GntR family transcriptional regulator	62%
GFO_1159	zg2616 (Sugar permease)	major facilitator superfamily permease hexuronate/hexarate transporter	78%
GFO_1160	zg2615 (Acetoin(diacetyl) reductase)	short-chain dehydrogenase/reductase family protein	82%
GFO_1161	zg2614 (KdgK1)	2-dehydro-3-deoxygluconokinase	43%
<b><i>Kordia algicida</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
KAOT1_04160	zg2618 (AlyA2)	Poly(beta-D-mannuronate) lyase	35%
<i>KAOT1_04165</i>	<i>no homolog</i>	<i>hypothetical protein</i>	<i>-</i>
KAOT1_04170	zg4131 (AlyA5)	putative alginate lyase	49%
KAOT1_04175	zg4131 (AlyA5)	alginate lyase	68%
KAOT1_04185	zg2623 (KdgF)	diguanylate cyclase/phosphodiesterase	69%



KAOT1_04190_04195	zg2621 (TBDR)	putative outer membrane protein	57%
KAOT1_04200	zg2620 (SusD-like)	putative outer membrane protein	63%
KAOT1_04205	zg2619 (PKD-CBM)	hypothetical protein	41%
KAOT1_04210	zg2618 (AlyA2)	hypothetical protein	63%
KAOT1_04215	zg2617 (GntR-like)	transcriptional regulator (GntR family)	69%
KAOT1_04220	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	59%
KAOT1_04225	zg2615 (Acetoin(diacetyl) reductase)	Short-chain dehydrogenase/reductase SDR	87%
KAOT1_04230	zg3343	<i>LacI</i> family transcriptional regulator	38%
KAOT1_04235	zg2614 (KdgK1)	D-xylose transporter	89%
KAOT1_13477	zg2612 (NRAMP)	natural resistance-associated macrophage protein	72%
KAOT1_13482	zg2617 (GntR-like)	regulatory protein GntR, HTH:GntR, C-terminal	59%
KAOT1_13487	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	42%
<b><i>Croceibacter atlanticus</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
CA2559_11458	zg2614 (KdgK1)	2-dehydro-3-deoxygluconokinase	52%
CA2559_11468	zg2615 (Acetoin(diacetyl) reductase)	acetoin(diacetyl)reductase	86%
CA2559_11473	zg2616 (Sugar permease)	putative hexuronate transport protein	79%
CA2559_11478	zg2617 (GntR-like)	transcriptional regulator, GntR family	66%
CA2559_11483	zg2618 (AlyA2)	hypothetical protein	69%
CA2559_11488	zg2619 (PKD-CBM)	serine protease, subtilase family protein	45%
CA2559_11493	zg2620 (SusD-like)	putative outer membrane protein	58%
CA2559_11498	zg2621 (TBDR)	putative outer membrane protein	55%
CA2559_11503	zg2623 (KdgF)	putative pectin degradation protein	73%
CA2559_11508	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	61%
CA2559_11513	zg4131 (AlyA5)	putative alginate lyase	65%
CA2559_11518	zg4130 (AlyA4)	putative alginate lyase precursor	48%
CA2559_11523	<i>no homolog</i>	<i>hypothetical protein</i>	-
CA2559_11528	zg2618 (AlyA2)	putative alginate lyase	37%
<b><i>Dokdonia donghaensis</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
MED134_10625	zg2614 (KdgK1)	2-dehydro-3-deoxygluconokinase	54%
MED134_10630	zg2615 (Acetoin(diacetyl) reductase)	short chain dehydrogenase	84%
MED134_10635	zg2616 (Sugar permease)	hexuronate transporter	76%
MED134_10640	zg2617 (GntR-like)	transcription regulator, GntR family	64%
MED134_10645	zg2618 (AlyA2)	alginate lyase	65%
MED134_10650	zg2619 (PKD-CBM)	conserved hypothetical protein	39%
MED134_10655	zg2620 (SusD-like)	outer membrane protein	59%
MED134_10660	zg2621 (TBDR)	TonB dependent/ligand-gated channel	58%
MED134_10665	zg2623 (KdgF)	conserved hypothetical protein	78%
MED134_10670	zg2624 (AlyA3)	heparinase II/III-like protein	57%
MED134_10675	<i>no homolog</i>	<i>heparinase II/III-like protein</i>	-
MED134_10680	zg4130 (AlyA4)	alginate lyase precursor	51%
MED134_10685	zg2618 (AlyA2)	alginate lyase	36%
<b><i>Flavobacteriales bacterium ALC-1</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
FBALC1_06803	zg2618 (AlyA2)	Poly(beta-D-mannuronate) lyase	38%
FBALC1_06808	<i>no homolog</i>	<i>CHU large protein; uncharacterized</i>	-
FBALC1_06813	zg4130 (AlyA4)	Poly(beta-D-mannuronate) lyase	57%
FBALC1_06818	zg4132 (AlyA6)	alginate lyase	67%
FBALC1_06823	<i>no homolog</i>	hypothetical protein	-
FBALC1_06828	zg67 ( <i>GndA</i> )	<i>gluconokinase and phosphogluconate dehydrogenase</i>	39%
FBALC1_06833	zg2612 (NRAMP)	natural resistance-associated macrophage protein	73%
FBALC1_06838	zg4131 (AlyA5)	alginate lyase	59%
FBALC1_06843	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	45%
FBALC1_06848	zg4130 (AlyA4)	Leucyl-tRNA synthetase class Ia	35%

FBALC1_06853	zg2624 (AlyA3)	Heparinase II/III-like protein	69%
FBALC1_06858	zg2623 (KdgF)	putative pectin degradation protein	48%
FBALC1_06863	zg2621 (TBDR)	putative outer membrane protein	35%
FBALC1_06868	zg2620 (SusD-like)	putative outer membrane protein	31%
FBALC1_06873	zg2621 (TBDR)	putative outer membrane protein	63%
FBALC1_06878	zg2620 (SusD-like)	putative outer membrane protein	62%
FBALC1_06883	zg2619 (PKD-CBM)	surface antigen gene	42%
FBALC1_06888	zg2618 (AlyA2)	lyase, putative	68%
FBALC1_06893	zg2617 (GntR-like)	regulatory protein GntR, HTH	70%
FBALC1_06898	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	39%
FBALC1_06903	zg2615 (Acetoin(diacetyl) reductase)	Short-chain dehydrogenase/reductase SDR	86%
FBALC1_06908	zg2614 (KdgK1)	PfkB domain protein	39%
FBALC1_06913	zg3343	<i>LacI</i> family transcriptional regulator	40%
FBALC1_06918	zg2614 (KdgK1)	histidyl-tRNA synthetase	89%
<b>Flavobacteria MS024-2A</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Flav2ADRAFT_1118	zg4131 (AlyA5)	Alginate lyase 2	65%
Flav2ADRAFT_1119	zg2612 (NRAMP)	natural resistance-associated macrophage protein	72%
Flav2ADRAFT_1120	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	43%
Flav2ADRAFT_1121	zg2623 (KdgF)	Cupin 2 conserved barrel domain protein	74%
Flav2ADRAFT_1122	zg2621 (TBDR)	TonB-dependent receptor plug	34%
Flav2ADRAFT_1123	zg2620 (SusD-like)	RagB/SusD domain protein	30%
Flav2ADRAFT_1124	zg2617 (GntR-like)	GntR domain protein	57%
Flav2ADRAFT_1125	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	45%
Flav2ADRAFT_1126	zg2615 (Acetoin(diacetyl) reductase)	short-chain dehydrogenase/reductase SDR	81%
Flav2ADRAFT_1128	zg2614 (KdgK1)	PfkB domain protein	84%
<b>Flavobacteria MS024-3C</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Flav3CDRAFT_0142	zg2619 (PKD-CBM)	protein containing PKD domain	34%
Flav3CDRAFT_0143	zg2618 (AlyA2)	Alginate lyase 2	59%
Flav3CDRAFT_0144	zg1280( <i>GntT2</i> )	<i>gluconate transporter</i>	51%
Flav3CDRAFT_0145	zg2612 (NRAMP)	natural resistance-associated macrophage protein	71%
Flav3CDRAFT_0146	zg4757 ( <i>GntK2</i> )	<i>6-phosphogluconate dehydrogenase</i>	49%
Flav3CDRAFT_0147	zg2621 (TBDR)	TonB-dependent receptor plug	35%
Flav3CDRAFT_0148	zg2620 (SusD-like)	RagB/SusD domain protein	32%
Flav3CDRAFT_0149	zg2617 (GntR-like)	GntR domain protein	65%
Flav3CDRAFT_0150	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	74%
Flav3CDRAFT_0151	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	61%
Flav3CDRAFT_0152	zg2614 (KdgK1)	PfkB domain protein	86%
Flav3CDRAFT_0153	zg4702 ( <i>Eda</i> )	<i>KDPG and KHG aldolase</i>	71%
Flav3CDRAFT_0154	zg2618 (AlyA2)	Alginate lyase 2	36%
<b>Bacteroidetes SCB49</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
SCB49_06272	zg4130 (AlyA4)	Poly(beta-D-mannuronate) lyase	57%
SCB49_06277	zg4131 (AlyA5)	Poly(beta-D-mannuronate) lyase	68%
SCB49_06282	zg67 ( <i>GndA</i> )	<i>gluconokinase and phosphogluconate dehydrogenase</i>	42%
SCB49_06287	zg2612 (NRAMP)	natural resistance-associated macrophage protein	72%
SCB49_06292	zg4131 (AlyA5)	alginate lyase	57%
SCB49_06297	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	42%
SCB49_06302	zg2624 (AlyA3)	Heparinase II/III-like protein	72%
SCB49_06307	zg2623 (KdgF)	putative pectin degradation protein	50%
SCB49_06312	zg2621 (TBDR)	putative outer membrane protein	34%
SCB49_06317	zg2620 (SusD-like)	putative outer membrane protein	32%
SCB49_06322	zg2617 (GntR-like)	GntR family transcriptional regulator	63%
SCB49_06327	zg2616 (Sugar permease)	putative hexuronate transport protein	45%

SCB49_06332	zg2615 (Acetoin(diacetyl) reductase)	Short-chain dehydrogenase/reductase SDR	87%
SCB49_06337	zg2614 (KdgK1)	PfkB domain protein	56%
<b><i>Polaribacter dokdonensis</i> MED152</b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
MED152_09780	zg4130 (AlyA4)	alginate lyase precursor	55%
MED152_09785	zg4131 (AlyA5)	alginate lyase precursor	63%
MED152_09790	zg4757 ( <i>GntK2</i> )	gluconokinase	58%
MED152_09795	zg67 ( <i>GndA</i> )	6-phosphogluconate dehydrogenase	40%
MED152_09800	zg2612 (NRAMP)	Mn2+ and Fe2+ transporter of the NRAMP family	73%
MED152_09805	zg4131 (AlyA5)	alginate lyase precursor	55%
MED152_09810	zg2622 (SDR)	short chain dehydrogenase	45%
MED152_09815	no homolog	hypothetical protein	-
MED152_09820	zg2624 (AlyA3)	heparinase II/III-like protein	61%
MED152_09825	zg2623 (KdgF)	conserved hypothetical protein	49%
MED152_09830	zg2621 (TBDR)	TonB dependent/ligand-gated channel	39%
MED152_09835	zg2620 ( <i>SusD-like</i> )	outer membrane protein	32%
MED152_09840	zg2619 (PKD-CBM)	conserved hypothetical protein	31%
MED152_09845	zg2619 (PKD-CBM)	conserved hypothetical protein	26%
MED152_09850	zg2617 ( <i>GntR-like</i> )	transcriptional regulator, GntR family	67%
MED152_09855	zg2616 (Sugar permease)	hexuronate transporter	44%
MED152_09860	zg2615 (Acetoin(diacetyl) reductase)	short chain dehydrogenase	83%
MED152_09865	zg2621 (TBDR)	TonB dependent/ligand-gated channel	35%
MED152_09870	zg2620 ( <i>SusD-like</i> )	outer membrane protein	32%
MED152_09875	zg4132 (AlyA6)	conserved hypothetical protein	35%
MED152_10135	Zg2614 (KgdK)	2-dehydro-3-deoxygluconokinase	90%
MED152_08075	Zg2614 (KgdK)	2-dehydro-3-deoxygluconokinase	38%
<b><i>Flavobacteria bacterium</i> BBFL7</b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
BBFL7_00723	zg4130 (AlyA4)	alginate (poly beta-D-mannuronate) lyase	50%
BBFL7_00724	zg2622 (SDR)	putative dehydrogenase/reductase	57%
BBFL7_00725	zg2615 (Acetoin(diacetyl) reductase)	putative dehydrogenase	85%
BBFL7_00726	zg2623 (KdgF)	pectin degradation protein	49%
BBFL7_00727	zg2624 (AlyA3)	putative alginate lyase	38%
BBFL7_00745	Zg2614 (KgdK)	gluconate kinase	42%
BBFL7_00743	Zg2617 (GntR)	transcriptional regulator GntR	44%
<b><i>Sphingobacteria</i></b>			
<b><i>Pedobacter saltans</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Pedsa_1191	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	63%
Pedsa_1192	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	47%
Pedsa_1193	zg4326	coagulation factor 5/8 type domain protein	35%
Pedsa_1194	zg2010	hypothetical protein	29%
Pedsa_1195	zg2623 (KdgF)	Cupin 2 conserved barrel domain protein	61%
Pedsa_1196	zg2624 (AlyA3)	Heparinase II/III family protein	42%
Pedsa_1197	zg2617 (GntR-like)	GntR domain protein	41%
Pedsa_0628	zg2621 (TBDR)	TonB-dependent receptor plug	32%
Pedsa_0629	zg2620 ( <i>SusD-like</i> )	RagB/SusD domain protein	30%
Pedsa_0630	no homolog	hypothetical protein	-
Pedsa_0631	zg4132 (AlyA6)	TonB-dependent receptor	38%
Pedsa_0632	zg4132 (AlyA6)	hypothetical protein	40%
Pedsa_2718	Zg2614 (KgdK)	PfkB domain protein	36%
<b><i>Cytophagia</i></b>			
<b><i>Rhodothermus marinus</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Rmar_1157	zg2617 ( <i>GntR-like</i> )	GntR domain-containing protein	27%
Rmar_1158	no homolog	hypothetical protein	-
Rmar_1159	zg3153 ( <i>DgoA</i> )	2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase	35%
Rmar_1160	zg2614 (KdgK1)	PfkB domain-containing protein	53%
Rmar_1161	zg1297 (TBDR)	TonB-dependent receptor	32%
Rmar_1162	no homolog	hypothetical protein	-

Rmar_1163	<i>no homolog</i>	<i>hypothetical protein</i>	-
Rmar_1164	<i>no homolog</i>	<i>hypothetical protein</i>	-
Rmar_1165	zg4132 (AlyA6)	poly(beta-D-mannuronate) lyase	40%
Rmar_1166	zg2624 (AlyA3)	Heparinase II/III family protein	37%
Rmar_1167	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	51%
Rmar_1168	zg2623 (KdgF)	cupin 2 barrel domain-containing protein	50%
Rmar_1169	<i>no homolog</i>	<i>Heparinase II/III family protein</i>	-
Rmar_1170	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	52%
<b>Intestinal Bacteroides</b>			
<b>Bacteroides clarus</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
HMPREF9445_01939	zg2617 (GntR-like)	FCD domain protein	44%
HMPREF9445_01940	zg4169 (GyaR)	glyoxylate reductase	37%
HMPREF9445_01941	zg2622 (SDR)	oxidoreductase, short chain dehydrogenase/reductase family	54%
HMPREF9445_01942	zg2616 (Sugar permease)	transporter, major facilitator family	45%
HMPREF9445_01943	zg2623 (KdgF)	cupin domain protein	60%
HMPREF9445_01944	zg2624 (AlyA3)	heparinase II/III-like protein	45%
HMPREF9445_01945	zg2621 (TBDR)	TonB-dependent receptor plug domain protein	36%
HMPREF9445_01946	zg2620 (SusD-like)	SusD family protein	32%
HMPREF9445_01947	zg3520 (SiaE2)	<i>hypothetical protein</i>	34%
HMPREF9445_01948	zg4132 (AlyA6)	<i>hypothetical protein</i>	38%
HMPREF9445_01628	Zg2614 (KdgK)	kinase, PfkB family	56%
<b>Bacteroides eggertii</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
BACEGG_03245	zg4132 (AlyA6)	<i>hypothetical protein</i>	38%
BACEGG_03246	zg3520 (SiaE2)	<i>hypothetical protein</i>	34%
BACEGG_03247	zg2620 (SusD-like)	<i>hypothetical protein</i>	32%
BACEGG_03248	zg2621 (TBDR)	<i>hypothetical protein</i>	35%
BACEGG_03249	zg2624 (AlyA3)	<i>hypothetical protein</i>	45%
BACEGG_03250	zg2623 (KdgF)	<i>hypothetical protein</i>	60%
BACEGG_03251	zg2616 (Sugar permease)	<i>hypothetical protein</i>	45%
BACEGG_03252	zg2622 (SDR)	<i>hypothetical protein</i>	54%
BACEGG_03253	zg4169 (GyaR)	<i>hypothetical protein</i>	38%
BACEGG_03254	zg2617 (GntR-like)	<i>hypothetical protein</i>	44%
BACEGG_01441	Zg2612 (NRAMP)	<i>hypothetical protein</i>	38%
BACEGG_01788	Zg2614 (KdgK)	<i>hypothetical protein</i>	55%
<b>Bacteroides sp. 2_2_4</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
BSCG_01543	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	54%
BSCG_01544	zg2616 (Sugar permease)	major facilitator superfamily transporter MFS_1	46%
BSCG_01545	zg2623 (KdgF)	conserved hypothetical protein	54%
BSCG_01546_01547	zg2624 (AlyA3)	conserved hypothetical protein	44%
BSCG_01548	zg4132 (AlyA6)	Poly(beta-D-mannuronate) lyase	37%
BSCG_01549	<i>no homolog</i>	<i>predicted protein</i>	-
BSCG_01550	zg2620 (SusD-like)	conserved hypothetical protein	29%
BSCG_01551	zg2621 (TBDR)	conserved hypothetical protein	32%
BSCG_01552	zg2617 (GntR-like)	predicted protein	38%
BSCG_03910	Zg2614 (KdgK)	2-dehydro-3-deoxygluconokinase	54%
<b>Bacteroides sp. 3_1_23</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
HMPREF9010_00433	zg2622 (SDR)	oxidoreductase, short chain dehydrogenase/reductase family	54%
HMPREF9010_00434	zg2616 (Sugar permease)	hexuronate transporter	46%
HMPREF9010_00435	zg2623 (KdgF)	pectin degradation protein KdgF	54%
HMPREF9010_00436	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	42%
HMPREF9010_00437	zg4132 (AlyA6)	alginate lyase	37%
HMPREF9010_00438	<i>no homolog</i>	<i>putative IPT/TIG domain protein</i>	-
HMPREF9010_00439	zg2620 (SusD-like)	putative outer membrane protein	29%
HMPREF9010_00440	zg2621 (TBDR)	putative outer membrane protein	32%
HMPREF9010_00441	zg2617 (GntR-like)	putative transcriptional regulator, GntR family	38%
HMPREF9010_04927	Zg2614 (KdgK)	kinase, PfkB family	54%
<b>Bacteroides sp. D2</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
BSGG_2567	zg2617 (GntR-like)	transcriptional regulator	38%
BSGG_2568	zg2621 (TBDR)	conserved hypothetical protein	31%

BSGG_2569	zg2620 (SusD-like)	conserved hypothetical protein	28%
BSGG_2570	no homolog	conserved hypothetical protein	-
BSGG_2571	zg4132 (AlyA6)	TonB-dependent receptor	37%
BSGG_2572	zg2624 (AlyA3)	conserved hypothetical protein	43%
BSGG_2573	zg2623 (KdgF)	conserved hypothetical protein	55%
BSGG_2574	zg2616 (Sugar permease)	major facilitator transporter	46%
BSGG_2575	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	55%

### Alphaproteobacteria

<i>Maricaulis maris</i>	<i>Z. galactanivorans</i> ortholog	Annotation	Sequence identity
Mmar10_0252	no homolog	transposase	-
Mmar10_0253	zg1355	TonB-dependent receptor	25%
Mmar10_0254	zg2617 (GntR-like)	GntR family transcriptional regulator	30%
Mmar10_0255	zg4132 (AlyA6)	Poly(beta-D-mannuronate) lyase	37%
Mmar10_0256	zg2624 (AlyA3)	heparinase II/III family protein	41%
Mmar10_0257	zg2623 (KdgF)	cupin 2 domain-containing protein	45%
Mmar10_0258	zg2616 (Sugar permease)	major facilitator transporter	46%
Mmar10_0259	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	62%
Mmar10_0260	zg2614 (KdgK1)	2-keto-3-deoxygluconate kinase	27%

<i>Hirschia baltica</i>	<i>Z. galactanivorans</i> ortholog	Annotation	Sequence identity
Hbal_2769	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	62%
Hbal_2770	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	46%
Hbal_2771	zg2624 (AlyA3)	Heparinase II/III	37%
Hbal_2772	zg4132 (AlyA6)	poly(beta-D-mannuronate) lyase	44%
Hbal_2773	zg1355 (TBDR)	TonB-dependent receptor	24%
Hbal_2774	zg2617 (GntR-like)	GntR family transcriptional regulator	29%
Hbal_2713	Zg2614 (KdgK)	PfkB domain protein	38%

<i>Brevundimonas subvibrioides</i>	<i>Z. galactanivorans</i> ortholog	Annotation	Sequence identity
Bresu_2894	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	58%
Bresu_2895	zg2616 (Sugar permease)	major facilitator superfamily MFS_1	47%
Bresu_2896	zg2624 (AlyA3)	Heparinase II/III	41%
Bresu_2897	zg4132 (AlyA6)	TonB-dependent receptor	40%
Bresu_2898	zg1355	TonB-dependent receptor	28%
Bresu_2899	zg2617 (GntR-like)	GntR family transcriptional regulator	27%
Bresu_3168	Zg2614 (KdgK)	PfkB domain protein	33%

### Gammaproteobacteria (Non Vibrio)

<i>Pseudoalteromonas haloplanktis</i>	<i>Z. galactanivorans</i> ortholog	Annotation	Sequence identity
PSHAa1739	zg2617 (GntR-like)	GntR family transcriptional regulator	29%
PSHAa1740	zg1355	TonB-dependent receptor	23%
PSHAa1741	no homolog	hypothetical protein	-
PSHAa1742	zg2116	fructose-1,6-bisphosphatase	38%
PSHAa1743	zg764	fructose-bisphosphate aldolase	64%
PSHAa1744	zg3153	4-hydroxy-2-oxoglutarate aldolase	28%
PSHAa1745	zg2614 (KdgK1)	ketodeoxygluconokinase	25%
PSHAa1746	zg2616 (Sugar permease)	hexuronate transport protein (MFS)	46%
PSHAa1747	zg2623 (KdgF)	pectin degradation protein	52%
PSHAa1748	zg2624 (AlyA3)	chondroitin AC/alginate lyase	43%
PSHAa1749	zg4132 (AlyA6)	alginate lyase precursor	44%

<i>Pseudoalteromonas atlantica</i>	<i>Z. galactanivorans</i> ortholog	Annotation	Sequence identity
Patl_3638	zg1355 (TBDR)	TonB-dependent receptor	25%
Patl_3639	zg4131 (AlyA5)	Poly(beta-D-mannuronate) lyase	34%
Patl_3640	zg4132 (AlyA6)	Poly(beta-D-mannuronate) lyase	46%
Patl_3641	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	43%
Patl_3642	no homolog	RbsD or FucU transport	-
Patl_3643	zg2612 (NRAMP)	natural resistance-associated macrophage protein	71%
Patl_3644	zg4757 (GntK2)	carbohydrate kinase	46%
Patl_3645	zg4131 (AlyA5)	alginate lyase	34%
Patl_3646	no homolog	phosphomannose isomerase-like	-

Patl_3647	zg2612 (NRAMP)	natural resistance-associated macrophage protein	37%
Patl_3648	zg4701 (UxuA2)	mannonate dehydratase	47%
Patl_3649	zg4623 (UxuB)	mannitol dehydrogenase Rossman-like	35%
Patl_3650	zg1355 (TBDR)	TonB-dependent receptor	27%
Patl_3651	zg2624 (AlyA3)	heparinase II/III-like	69%
Patl_3652	zg1697	cation diffusion facilitator family transporter	25%
Patl_3653	no homolog	hypothetical protein	-
Patl_3654	no homolog	NAD(P) transhydrogenase, beta subunit	-
Patl_3655	no homolog	putative NAD(P) transhydrogenase subunit alpha	-
Patl_3656	zg 2668 (AldI)	alanine dehydrogenase/PNT-like	23%
Patl_3657	zg1491 (ScrK1)	PfkB	26%
Patl_3658	zg2616 (Sugar permease)	major facilitator transporter	46%
Patl_3659	zg4132 (AlyA6)	Poly(beta-D-mannuronate) lyase	45%
Patl_3660	zg2617 (GntR-like)	GntR-like	28%
Patl_3661	zg2622 (SDR)	short-chain dehydrogenase/reductase SDR	56%
<b>Alteromonas macleodii</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
AmacA2_010100014539	zg1355 (TBDR)	TonB-dependent receptor	27%
AmacA2_010100014544	zg67 (GndA)	tartronate semialdehyde reductase	27%
AmacA2_010100014549	zg4132 (AlyA6)	alginate lyase precursor	45%
AmacA2_010100014554	zg2624 (AlyA3)	chondroitin AC/alginate lyase	42%
AmacA2_010100014559	zg2623 (KdgF)	pectin degradation protein	53%
AmacA2_010100014564	zg2616 (Sugar permease)	hexuronate transport protein (MFS)	47%
AmacA2_010100014569	zg2614 (KdgK1)	ketodeoxygluconokinase	23%
AmacA2_010100014574	zg3153 (dgoA)	4-hydroxy-2-oxoglutarate aldolase	27%
AmacA2_010100014579	zg2617 (GntR-like)	GntR-like protein	32%
AmacA2_010100014584	zg4131 (AlyA5)	cyclic nucleotide-binding domain-containing protein	23%
AmacA2_010100014589	zg1355 (TBDR)	TonB-dependent receptor	24%
AmacA2_010100014594	zg1708	Poly(beta-D-mannuronate) lyase	30%
AmacA2_010100014599	zg4131 (AlyA5)	putative alginate lyase	33%
<b>Saccharophagus degradans</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Sde_3274	zg4132 (AlyA6)	Poly(beta-D-mannuronate) lyase	29%
Sde_3275	zg4130 (AlyA4)	leucyl-tRNA synthetase class Ia	32%
Sde_3276	no homolog	hypothetical protein	-
Sde_3277	zg2780 (tpiA)	triosephosphate isomerase	41%
Sde_3278	no homolog	response regulator receiver domain-containing protein	-
Sde_3278	zg764 (fbaA)	fructose-bisphosphate aldolase	70%
Sde_3280	zg2614 (KdgK1)	2-keto-3-deoxygluconate kinase	27%
Sde_3281	zg2622 (SDR)	hypothetical protein	64%
Sde_3282	zg2616 (Sugar permease)	putative hexuronate transporter	51%
Sde_3283	zg2623 (KdgF)	diguanylate cyclase/phosphodiesterase	56%
Sde_3284	zg2624 (AlyA3)	hypothetical protein	42%
Sde_3285	zg4132 (AlyA6)	TonB-dependent receptor	49%
Sde_3286	zg4131 (AlyA5)	Poly(beta-D-mannuronate) lyase	27%
Sde_3287	zg2617 (GntR-like)	GntR family transcriptional regulator	29%
<b>Alteromonadales bacterium TW-7</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
ATW7_06243	zg4132 (AlyA6)	putative alginate lyase precursor	46%
ATW7_06248	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	41%
ATW7_06253	zg2623 (KdgF)	putative pectin degradation protein	50%
ATW7_06258	zg2616 (Sugar permease)	putative hexuronate transport protein	46%
ATW7_06263	zg2614 (KdgK1)	putative ketodeoxygluconokinase	21%
ATW7_06268	zg3153	putative 4-hydroxy-2-oxoglutarate aldolase	28%
ATW7_06273	zg764	fructose-bisphosphate aldolase	65%
ATW7_06278	zg2116	Fructose-1,6-bisphosphatase	38%
ATW7_06283	no homolog	hypothetical protein	-
ATW7_06288	zg4130 (AlyA4)	Poly(beta-D-mannuronate) lyase	22%
ATW7_06293	zg1355 (TBDR)	putative TonB-dependent receptor	25%
ATW7_06298	zg1355 (TBDR)	putative TonB-dependent receptor	26%

ATW7_02377	Zg2617 ( <i>GntR</i> )	putative regulatory protein, <i>GntR</i> family	31%
<b><i>Shewanella frigidimarina</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Sfri_3102	zg2617 ( <i>GntR-like</i> )	<i>GntR</i> domain-containing protein	30%
Sfri_3103	zg4132 ( <i>AlyA6</i> )	Poly(beta-D-mannuronate) lyase	47%
Sfri_3104	zg2624 ( <i>AlyA3</i> )	heparinase II/III family protein	40%
Sfri_3105	zg2623 ( <i>KdgF</i> )	cupin 2 domain-containing protein	47%
Sfri_3106	zg2616 ( <i>Sugar permease</i> )	major facilitator transporter	52%
Sfri_3107	zg2614 ( <i>KdgK1</i> )	ribokinase-like domain-containing protein	24%
<b><i>Stenotrophomonas maltophilia</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Smlt2599	zg2617 ( <i>GntR-like</i> )	putative <i>GntR</i> family transcriptional regulator	30%
Smlt2600	zg1355 ( <i>TBDR</i> )	putative <i>TonB</i> dependent receptor	21%
Smlt2601	zg4132 ( <i>AlyA6</i> )	putative poly (beta-D-mannuronate) lyase	42%
Smlt2602	zg2624 ( <i>AlyA3</i> )	putative alginate lyase protein	41%
Smlt2603	zg2616 ( <i>Sugar permease</i> )	putative hexuronate transporter	47%
Smlt2604	zg2622 ( <i>SDR</i> )	putative 3-oxoacyl-[acyl carrier protein] reductase	57%
Smlt2605	no homolog	putative GDSL-like lipase/acylhydrolase exported protein	-
Smlt2606	zg2614 ( <i>KdgK1</i> )	putative <i>PfkB</i> family carbohydrate kinase	37%
<b><i>Stenotrophomonas maltophilia R551-3</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Smal_2064	zg2617 ( <i>GntR-like</i> )	<i>GntR</i> domain-containing protein	29%
Smal_2065	zg1355 ( <i>TBDR</i> )	<i>TonB</i> -dependent receptor	21%
Smal_2066	zg4132 ( <i>AlyA6</i> )	alginate lyase precursor	40%
Smal_2067	zg2624 ( <i>AlyA3</i> )	Heparinase II/III family protein	40%
Smal_2068	zg2616 ( <i>Sugar permease</i> )	major facilitator superfamily protein	48%
Smal_2069	zg2622 ( <i>SDR</i> )	short-chain dehydrogenase/reductase <i>SDR</i>	59%
Smal_2070	no homolog	<i>G-D-S-L</i> family lipolytic protein	-
Smal_2071	zg2614 ( <i>KdgK1</i> )	<i>PfkB</i> domain-containing protein	36%
<b><i>Rahnella aquatilis</i> sp. Y9602</b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
Rahaq_2074	zg2624 ( <i>AlyA3</i> )	Heparinase II/III family protein	34%
Rahaq_2075	zg2622 ( <i>SDR</i> )	short-chain dehydrogenase/reductase <i>SDR</i>	38%
Rahaq_2076	zg2617 ( <i>GntR-like</i> )	<i>GntR</i> family transcriptional regulator	26%
Rahaq_2077	no homolog	<i>AraC</i> family transcriptional regulator	-
Rahaq_2078	zg2616 ( <i>Sugar permease</i> )	major facilitator superfamily <i>MFS_1</i>	24%
<b><i>Vibrio</i></b>			
<b><i>Vibrio sinoaloensis</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
VISI1226_00965	zg3153 ( <i>dgoA</i> )	<i>keto</i> -hydroxyglutarate-aldolase/ <i>keto</i> -deoxy-phosphogluconate aldolase	30%
VISI1226_00970	zg1491 ( <i>scrK1</i> )	2-dehydro-3-deoxygluconokinase	25%
VISI1226_00975	zg2623 ( <i>KdgF</i> )	pectin degradation protein	45%
VISI1226_00980	zg2624 ( <i>AlyA3</i> )	hypothetical protein	34%
VISI1226_00985	zg2624 ( <i>AlyA3</i> )	hypothetical protein	33%
VISI1226_00990	zg2282	hypothetical protein	33%
VISI1226_00995	zg2923 ( <i>norM1</i> )	putative multidrug resistance protein <i>norM</i>	24%
VISI1226_01000	zg2617 ( <i>GntR-like</i> )	putative transcriptional regulator	24%
<b><i>Vibrionales bacterium SWAT-3</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
VSWAT3_23599	zg3153 ( <i>dgoA</i> )	<i>keto</i> -hydroxyglutarate-aldolase/ <i>keto</i> -deoxy-phosphogluconate aldolase	28%
VSWAT3_23604	zg1491 ( <i>scrK1</i> )	ketodeoxygluconokinase	25%
VSWAT3_23609	zg2623 ( <i>KdgF</i> )	pectin degradation protein	44%
VSWAT3_23614	zg2624 ( <i>AlyA3</i> )	Heparinase II/III-like protein	34%
VSWAT3_23619	zg2624 ( <i>AlyA3</i> )	Heparinase II/III-like protein	34%
VSWAT3_23634	Zg2617 ( <i>GntR</i> )	Hypothetical <i>GntR</i> -family transcriptional regulator	27%
VSWAT3_23639	Zg2617 ( <i>GntR</i> )	Hypothetical <i>GntR</i> -family transcriptional regulator	35%
<b><i>Vibrio alginolyticus</i></b>	<b><i>Z. galactanivorans</i> ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
VMC_35160	zg2615 ( <i>Acetoin</i> (diacetyl) reductase)	conserved hypothetical protein	77%
VMC_35170	zg67 ( <i>gndA</i> )	conserved hypothetical protein	25%

VMC_35180	zg3605	conserved hypothetical protein	29%
VMC_35190	zg3153 ( <i>dgoA</i> )	Putative KHG/KDPG aldolase	25%
VMC_35200	zg1491 ( <i>scrK1</i> )	conserved hypothetical protein	24%
VMC_35210	zg2623 (KdgF)	conserved hypothetical protein	44%
VMC_35220	zg2624 (AlyA3)	hypothetical protein	34%
VMC_35230	zg2624 (AlyA3)	hypothetical protein	33%
VMC_35240	zg2618 (AlyA2)	hypothetical protein	25%
VMC_35250	zg4131 (AlyA5)	hypothetical protein	28%
VMC_35260	no homolog	hypothetical protein	-
VMC_35270	no homolog	hypothetical protein	-
VMC_35280	zg4131 (AlyA5)	hypothetical protein	34%
VMC_35290	zg2923 ( <i>norM1</i> )	conserved hypothetical protein	25%
VMC_35300	zg2617 (GntR-like)	hypothetical protein	27%
VMC_35310	zg2617 (GntR-like)	conserved hypothetical protein	27%
<b>Vibrio harveyi IDA3</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
VME_15190	zg2617 (GntR-like)	conserved hypothetical protein	26%
VME_15200	zg2615 (Acetoin(diacetyl) reductase)	conserved hypothetical protein	37%
VME_15210	zg2617 (GntR-like)	hypothetical protein	27%
VME_15220	zg2923 ( <i>norM1</i> )	conserved hypothetical protein	22%
VME_15230	zg4131 (AlyA5)	hypothetical protein	35%
VME_15370	zg4131 (AlyA5)	hypothetical protein	27%
VME_15380	zg2624 (AlyA3)	hypothetical protein	33%
VME_15390	zg2624 (AlyA3)	hypothetical protein	34%
VME_15400	zg2623 (KdgF)	hypothetical protein	35%
VME_15410	zg1491 ( <i>scrK1</i> )	conserved hypothetical protein	24%
VME_15420	zg3153 ( <i>dgoA</i> )	Putative KHG/KDPG aldolase	27%
<b>Vibrio sp. MED222</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
MED222_13495	zg3153 ( <i>dgoA</i> )	keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase	27%
MED222_13500	zg1491 ( <i>scrK1</i> )	2-dehydro-3-deoxygluconokinase	23%
MED222_13505	zg2623 (KdgF)	pectin degradation protein	46%
MED222_13510	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	34%
MED222_13515	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	34%
MED222_13535	Zg2617 (GntR)	putative transcriptional regulator	27%
MED222_13540	Zg2617 (GntR)	putative transcriptional regulator	26%
<b>Vibrio splendidus 12B01</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
V12B01_24199	zg3153 ( <i>dgoA</i> )	keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase	27%
V12B01_24204	zg1491 ( <i>scrK1</i> )	putative 2-dehydro-3-deoxygluconokinase	23%
V12B01_24209	zg2623 (KdgF)	pectin degradation protein	46%
V12B01_24214	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	33%
V12B01_24219	zg2624 (AlyA3)	putative chondroitin AC/alginate lyase	34%
V12B01_24224	zg3473	hypothetical protein	31%
V12B01_24229	zg2617 (GntR-like)	Hypothetical GntR-family transcriptional regulator	24%
V12B01_24259	zg4131 (AlyA5)	putative alginate lyase	29%
V12B01_24264	no homolog	hypothetical protein	-
V12B01_24269	no homolog	putative oligogalacturonate specific porin	-
V12B01_24274	zg4131 (AlyA5)	putative alginate lyase	34%
V12B01_24279	zg2923 ( <i>norM1</i> )	MATE efflux family protein	23%
V12B01_24284	zg2617 (GntR-like)	GntR-family transcriptional regulator	36%
V12B01_24289	zg2615 (Acetoin(diacetyl) reductase)	3-ketoacyl-CoA reductase PhaB	35%
V12B01_24294	zg2617 (GntR-like)	Hypothetical GntR-family transcriptional regulator	25%
V12B01_24254	Zg2618 (AlyA2)	lyase, putative	25%
<b>Vibrio splendidus LGP32</b>	<b>Z. galactanivorans ortholog</b>	<b>Annotation</b>	<b>Sequence identity</b>
VS_1242	zg3153 ( <i>dgoA</i> )	keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase	27%
VS_1243	zg1491 ( <i>scrK1</i> )	2-dehydro-3-deoxygluconokinase	24%
VS_1244	zg2623 (KdgF)	pectin degradation protein	46%
VS_1245	no homolog	hypothetical protein	-



VS_1246	zg2624 (AlyA3)	hypothetical protein	34%
VS_1247	zg2624 (AlyA3)	hypothetical protein	33%
VS_1250	Zg2617 (GntR)	putative transcriptional regulator	27%
VS_1250	Zg2617 (GntR)	putative transcriptional regulator	2%

**Table S6. Cloning strategy and expression results for recombinant alginate lyases.**

Target	Cloning strategy	Primers forward (top)-reverse (bottom)	Amplification	Soluble expression
AlyA1 cat. module	BglII / EcoRI	gggggggagatcttcgggtggtcgtccaccctc ccccccgaattcttaattgtgggttacgcttaggttttg	+	+
AlyA2	BamHI / MfeI	ggggggggatcctgccaagacagtaatgactggcag ccccccaattgctagtataacttctaaatcgtaaatttc	+	-
AlyA3	BamHI / XhoI	ggggggggatcccaagaacaccgagtctgatactg ccccccctcgagtaattgatttaataaattgatagggcc	-	-
AlyA4	BamHI / EcoRI	ggggggggatccaacgataaagggttacgtaaag ccccccgaattcttaatacatcaattgagtgggag	-	-
AlyA5	BamHI / EcoRI	ggggggggatcctgtaagacaaacctaggccactacg ccccccgaattctcactgggcttctggggcgcttc	+	+
AlyA6	BamHI / XhoI	ggggggggatccaacaaaaaacggtcagaacg ccccccctcgagtattgggaaattaccccaaatcg	+	-
AlyA7	BamHI / EcoRI	ggggggggatccgatgccttaggggttcatttgc ccccccgaattctcaactttataaatcatcttggaatc	+	+



# Chapitre 5

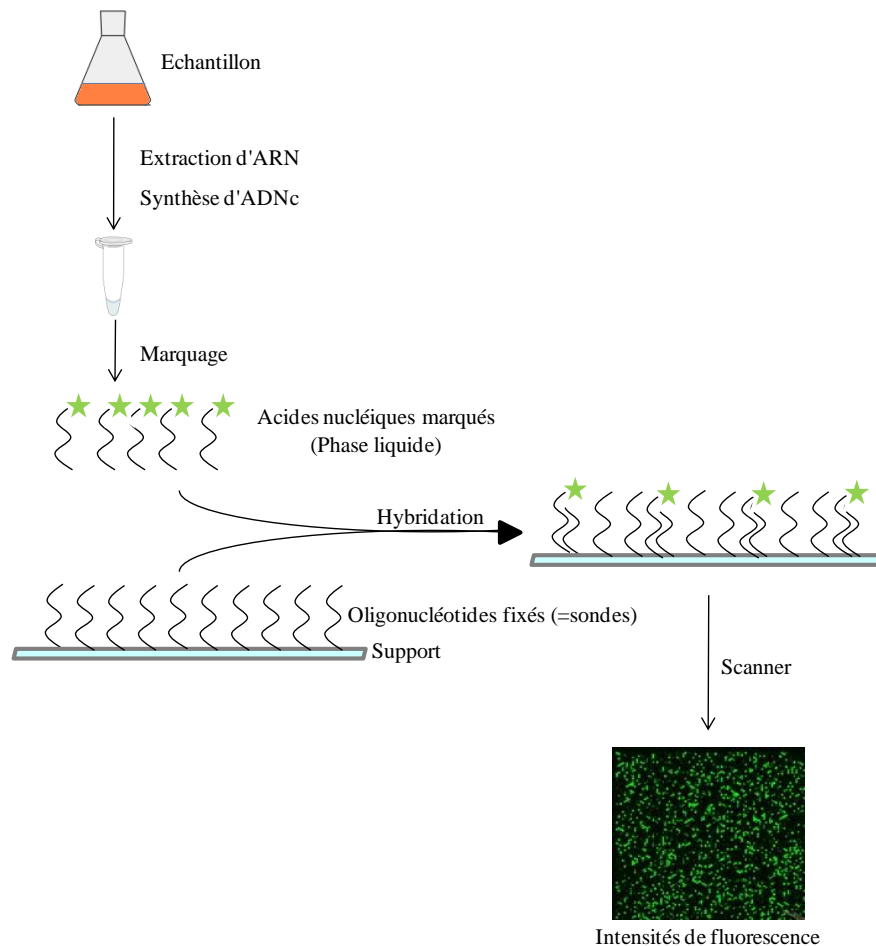
## Analyse transcriptomique à l'échelle du génome

---

## I- Introduction

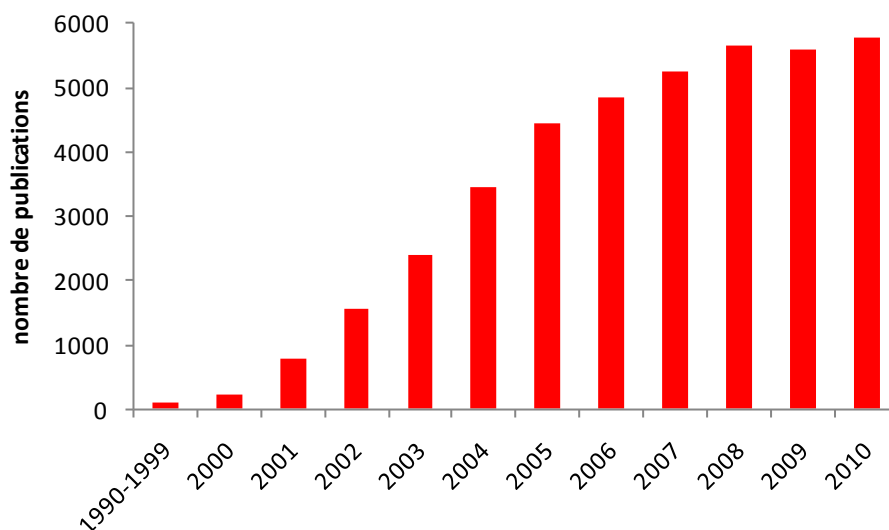
Les analyses présentées dans les Chapitres 3 et 4 ont permis de quantifier par RT-PCR quantitative l'expression de gènes prédits pour leur implication dans le métabolisme de l'alginate, en fonction des conditions de culture. Bien que très puissante et sensible, cette méthode ne permet pas de suivre en parallèle les niveaux de transcription d'un grand nombre de gènes. Par ailleurs, la nécessité de choisir un lot de gènes à tester empêche la découverte de nouveaux acteurs non prédits *a priori* dans le métabolisme étudié. Afin de mieux caractériser les processus cellulaires gouvernant la perception, la régulation du métabolisme et la dégradation de polysaccharides, il est donc nécessaire d'obtenir une "photographie" du transcriptome complet de *Z. galactanivorans* en présence de différentes sources de carbone.

Une des méthodes de choix pour l'analyse transcriptomique à l'échelle d'un génome entier, notamment chez les procaryotes, est l'utilisation de la technologie des puces à ADN, ou *microarrays* en anglais. Les puces consistent en une série d'oligonucléotides arrangés régulièrement sur un support (ex. une lame de verre). La mesure de l'expression implique l'hybridation de la puce avec un échantillon d'acides nucléiques marqués (Figure 5-1).



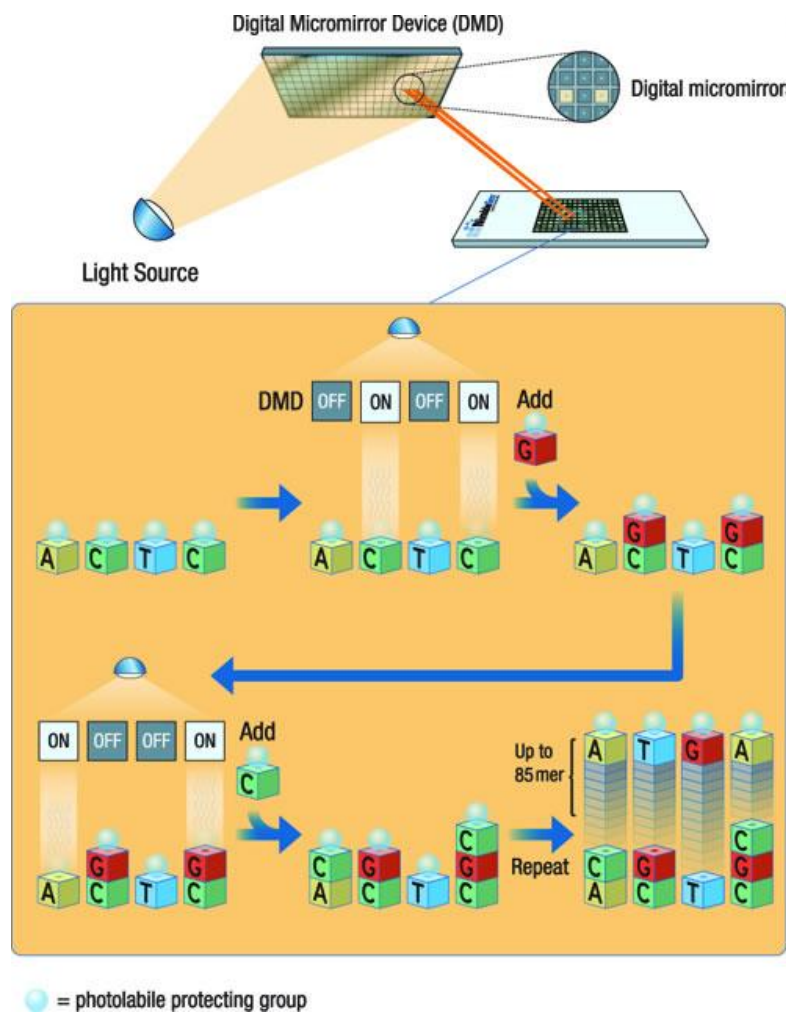
**Figure 5-1 : Schéma de principe d'une expérience transcriptomique utilisant des puces à ADN**

Les puces à ADN au sens large existent depuis les années 1970, sous forme de Dot-Blots réalisés manuellement sur des membranes, permettant l'analyse de l'expression sur une série d'acides nucléiques le plus souvent marqués à la radioactivité (Jordan 2002). Une avancée technologique majeure à la fin des années 1980 a été le développement de robots, rendant possible le dépôt régulier et compact de clones bactériens sur des filtres à "haute densité" (typiquement, 10000 spots sur un carré de 22 cm de côté). De tels filtres ont permis les premières analyses d'expression sur un grand nombre de gènes en parallèle (voir par exemple Brunet *et al.*, 1988). L'essor des techniques d'amplification par PCR a permis de déposer directement les fragments d'ADN sur les membranes. L'utilisation d'un marquage radioactif et de films de détection rendait cependant la méthode plus qualitative que quantitative. Le développement de marquages basés sur des molécules fluorescentes, et le dépôt robotisé de fragments d'ADN sur des lames de verre ont considérablement amélioré la technique des puces. Vers le milieu des années 1990, de véritables "*microarrays*" couvrant jusqu'à 50000 gènes ont pu être utilisés pour des analyses transcriptomiques (Schena *et al.*, 1995; DeRisi *et al.*, 1996; Shalon *et al.*, 1996). Grâce à la miniaturisation rapide des techniques, les premières puces couvrant la totalité de génomes procaryotes ont pu être développées pour *Mycobacterium tuberculosis* et *E. coli* (Behr *et al.*, 1999; Richmond *et al.*, 1999; Tao *et al.*, 1999). Depuis, l'utilisation de puces à ADN pour des expériences de transcriptomique à l'échelle de génomes entiers (procaryotes et eucaryotes) s'est largement répandue, tant dans le milieu académique que dans l'industrie (Figure 5-2).



**Figure 5-2 : Nombre annuel de publications traitant de puces à ADN depuis les années 1990**  
Recherche effectuée sur Pubmed avec le terme "*Microarray*".

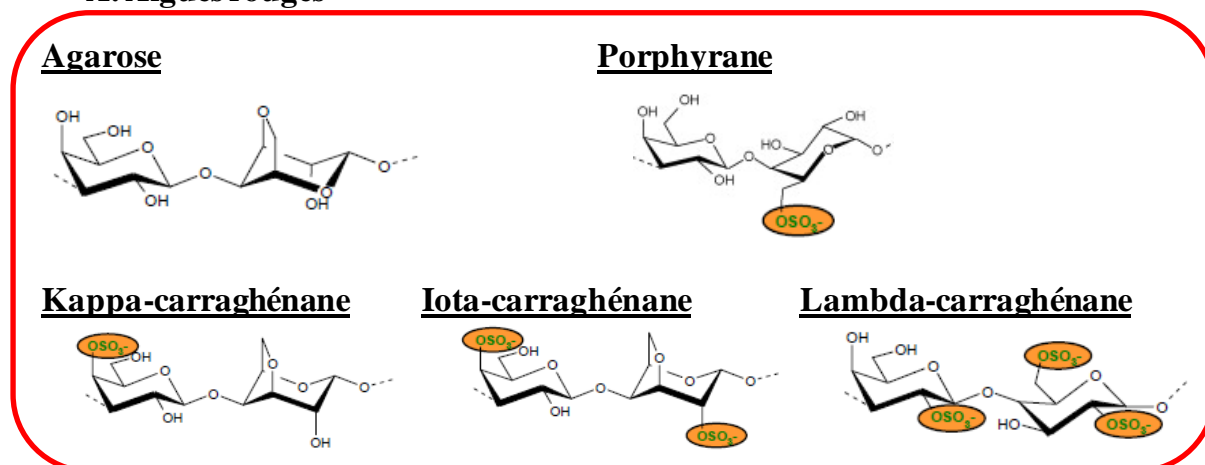
Dans le cadre du développement de *Z. galactanivorans* comme organisme modèle, il a été décidé d'initier une étude transcriptomique à l'échelle du génome. La technologie choisie est celle proposée par la compagnie Roche Nimblegen, basée sur la synthèse d'oligonucléotides de 60 bases sur des lames (Figure 5-3). Une lame de verre est dérivée de manière covalente avec une molécule *linker* exposant un groupement photolabile protecteur. A l'aide de miroirs pilotés par ordinateur, un faisceau lumineux est dirigé spécifiquement sur les emplacements à dé-protéger. Le nucléotide à insérer est alors incubé sur la lame, et se lie à l'oligomère en formation. La répétition de ces étapes permet la synthèse de sondes jusqu'à 85 bases, avec un niveau de miniaturisation très élevé (jusqu'à 385000 sondes par lame).



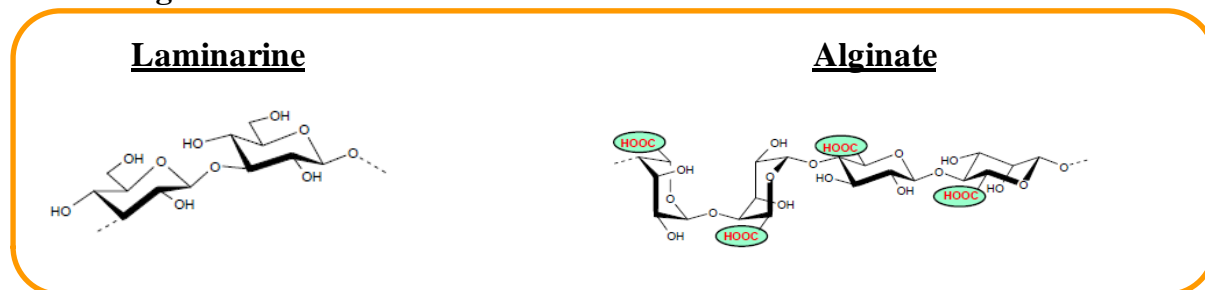
**Figure 5-3 : Principe de la synthèse de puces à ADN par photochimie**  
schéma du site Nimblegen  
<http://www.nimblegen.com/company/technology/>

Afin de déchiffrer les voies métaboliques impliquées dans la dégradation de polysaccharides marins chez *Z. galactanivorans*, l'étude présentée ici s'est concentrée sur la réponse à 8 sources de carbone différentes. Il s'agit du glucose, de deux polysaccharides d'algues brunes (alginate et laminarine) et de cinq polysaccharides d'algues rouges (agar, porphyrane, iota-, lambda- et kappa-carraghénanes). Ces différents polymères présentent des structures chimiques variées, tant au niveau du motif de répétition, des groupements de substitution que du type de liaison glycosidique (Figure 5-4).

### A. Algues rouges



### B. Algues brunes



**Figure 5-4 : Structure des polysaccharides utilisés dans cette étude**

**A.** Unité de répétition idéale disaccharidique de galactanes d'algues rouges. Le motif de base est composé de deux résidus galactose liés alternativement en  $\beta(1,4)$  et  $\alpha(1,3)$ . Le résidu galactose lié en  $(1,4)$  est de configuration D dans les carraghénanes et L dans les agars (agarose, porphyrane). Ce squelette est modifié par la présence de ponts 3,6 anhydro ou de groupements sulfate. **B.** Structure de polysaccharides d'algues brunes. La laminarine est un  $\beta(1,3)$  glucane. L'alginate est formé de deux résidus uroniques, l'acide  $\beta$ -D-mannuronique et l'acide  $\alpha$ -L-guluronique, liés par des liaisons 1,4.



La caractérisation du transcriptome de *Z. galactanivorans* devrait permettre de mettre en évidence des gènes dont l'expression répond spécifiquement à la présence de ces différents substrats, et donc de mieux comprendre les phénomènes de régulation et de dégradation enzymatique mis en jeu. Une telle approche a déjà été utilisée sur d'autres bactéries dégradant des polysaccharides de plantes terrestres (Chhabra *et al.*, 2003; Dodd *et al.*, 2010; Vanfossen *et al.*, 2011). Par ailleurs, compte-tenu de la grande proportion de gènes de fonction inconnue dans le génome de *Z. galactanivorans*, les résultats devraient permettre de sélectionner de futurs candidats pour des études structurales et fonctionnelles.

## II- Matériels et Méthodes

### II-1 Construction de la puce oligonucléotidique

Le dessin de la puce « à façon » 4\*72k a été réalisé par la compagnie PartnerChip (<http://www.partnerchip.fr/>), après discussion de nos besoins. Ce format de puce comporte 4 cadrans (soit 4 zones d'hybridation indépendantes par lame) contenant chacun 72000 sondes de 60 nucléotides. La séquence du génome de *Z. galactanivorans* a été utilisée pour la sélection des sondes. Douze sondes différentes ont été recherchées dans chacune des 4738 régions codantes. Afin de compléter la puce, des sondes ont également été recherchées dans les 4151 régions intergéniques avec un espacement moyen de 14 nucléotides. La construction de la puce a été mise en œuvre par la compagnie Roche Nimblegen.

### II-2 Culture de *Z. galactanivorans* et préparation des ARN

Des précultures de la souche *Z. galactanivorans* DsijT ont été réalisées en milieu Zobell 2216E (Zobell 1941) une nuit à 20°C et 180 rpm. La souche a ensuite été cultivée en triplicat, sous agitation à 180 rpm à 20°C en milieu minéral minimum (Thomas *et al.*, 2011). Ce milieu est composé, pour un litre, de : 50 mM Tris-HCl pH 8,0 ; 24,7 g NaCl ; 6,3 g MgSO<sub>4</sub>·7H<sub>2</sub>O ; 4,6 g MgCl<sub>2</sub>·H<sub>2</sub>O ; 2 g NH<sub>4</sub>Cl ; 0,7 g KCl ; 0,6 g CaCl<sub>2</sub> ; 200 mg NaHCO<sub>3</sub> ; 100 mg K<sub>2</sub>HPO<sub>4</sub> ; 50 mg extrait de levure et 20 mg FeSO<sub>4</sub>·7H<sub>2</sub>O. Huit sources de carbone différentes ont été testées : glucose, alginate, laminarine, agar, porphyrane, kappa-, iota- et lambda- carraghénanes. Chaque composé a été rajouté individuellement à la concentration finale de 2 g.L<sup>-1</sup> dans le milieu minimum. Le volume des cultures a été de 5 mL, sauf pour les conditions agar, iota- et lambda- carraghénanes (20 mL), toujours en respectant un ratio liquide/air de 1/10. Les volumes de réactifs dans la suite de cette partie sont donnés pour 5 mL de culture et ont été adaptés pour les cultures de 20 mL.

En fin de phase exponentielle, les cellules ont été collectées par centrifugation à 4000 rpm pendant 10 minutes. Les culots cellulaires ont été repris dans 200 µL de réactif Trizol (Sigma), transférés en tubes eppendorf, congelés dans l'azote liquide et conservés à -80°C. Les tubes ont été transférés directement du congélateur dans un bain-marie à 65°C pendant 15 minutes, en vortexant régulièrement jusqu'à la lyse des cellules. Cinquante microlitres de chloroforme ont été ajoutés. Après homogénéisation manuelle et incubation 5 minutes à

température ambiante, les tubes ont été centrifugés à 9000 rpm pendant 15 minutes à 4°C. La phase aqueuse a été récupérée et les ARN totaux ont été purifiés sur mini-colonne RNeasy (Qiagen), avec traitement à la DNase sur la colonne, en suivant le protocole du fournisseur. Les ARN ont été élués dans 15 µL d'eau RNase-free et analysés au spectrophotomètre (Nanodrop, ND-1000) afin de vérifier leur concentration (déduite de l'absorbance à 260 nm), la teneur en protéines (estimée par le ratio  $A_{260}/A_{280}$ ) et la contamination par des solvants ou des sels (estimée par le ratio  $A_{260}/A_{230}$ ). L'intégrité des ARN a été vérifiée en analysant 0,5 µL de chaque échantillon par électrophorèse sur gel d'agarose à 0,8%, après révélation au bromure d'éthidium.

### II-3 Synthèse des ADN complémentaires double brin

Les ADN complémentaires (ADNc) ont été synthétisés en utilisant le kit Superscript Double Stranded cDNA (Invitrogen) selon le protocole du fournisseur. La synthèse du premier brin a été réalisée par transcription inverse (RT) 1 h à 42°C dans un volume final de 20 µL en présence de 10 µg d'ARN totaux, 5 µM d'amorces hexamériques aléatoires, 1X du tampon commercial, 10 mM dithiothréitol, 500 µM de chaque dNTP et 400 U d'enzyme Superscript II RT. Les produits de cette réaction ont alors été utilisés comme matrice pour la synthèse du second brin d'ADNc, dans un volume final de 150 µL et après addition de 30 nmol de dNTP, 10 U de DNA ligase, 40 U de DNA polymérase I, 2 U de RNase H et 1 X de tampon commercial. Les milieux réactionnels ont été incubés 2 h à 16°C, puis complétés par 10 U de T4 DNA polymérase et ré-incubés 5 min à 16°C. Les échantillons ont été amenés à une concentration finale de 30 mM EDTA, puis traités par 4 µg de RNase A pendant 10 min à 37°C. Les protéines ont été éliminées par une étape d'extraction par un volume de phénol : chloroforme : alcool isoamylique (v/v, 25 :24 :1) suivie d'une centrifugation à 12000 g pendant 5 min. La phase aqueuse a été récupérée, complétée par 0,1 volume d'acétate d'ammonium 7,5 M et 7 µL de glycogène (5 mg.mL<sup>-1</sup>). Les ADNc ont été précipités par ajout de 2 volumes d'éthanol absolu froid et centrifugation 20 min à 12000 g et 4°C. Les culots d'ADNc ont été dessalés par deux lavages successifs avec 500 µL d'éthanol 80% froid et centrifugation 20 min à 12000 g et 4°C. Enfin, les ADNc ont été séchés sous vide (Speedvac) et repris dans 15 µL d'eau *nuclease-free*. Leur concentration a été déterminée au spectrophotomètre à 260 nm (Nanodrop, ND-1000) avant un envoi au prestataire de service PartnerChip.

## II-4 Marquage et hybridation des échantillons d'ADNc

Ces étapes ont été réalisées par la compagnie PartnerChip. La qualité des ADNc a été contrôlée par analyse des profils au Bioanalyser Agilent avec le kit RNA 6000 Nano. Le marquage des ADNc à la cyanine 3 a été effectué à partir de 1 µg d'ADNc en présence de fragment de Klenow (2 h à 37°C). Les échantillons d'ADNc marqués ont été précipités à l'isopropanol et quantifiés au spectrophotomètre à 260 nm. 2 µg ont été hybridés par cadran, en présence d'un cocktail d'hybridation dans une station Nimblegen (16 h à 42°C). Les puces à ADN ont été lavées dans trois bains successifs, séchées, et scannées pour mesurer la fluorescence de la cyanine 3 (scanner MS200, Roche-Nimblegen).

## II-5 Traitement et analyses statistiques des données transcriptomiques

Les fichiers de données brutes ont été fournis par PartnerChip et implémentés dans le logiciel ANAIS (Analysis of NimbleGen Arrays Interface) développé par l'INRA (<http://anais.versailles.inra.fr>) (Simon and Biot 2010). En utilisant cette interface, les données ont été normalisées par cadran et entre cadrans. La normalisation intra-cadran (algorithme *bg.adjust* dans le package *affy* de R) vise à corriger le bruit de fond. Dans cette méthode, l'intensité mesurée  $I_{obs}$  est considérée comme égale à la somme d'un signal spécifique  $S$  (de distribution exponentielle) et d'un bruit de fond  $B$  (de distribution normale). L'algorithme cherche à estimer le signal  $S$ , corrigé du bruit de fond.

$$I_{obs} = S + B \Rightarrow S = I_{obs} - B$$

La normalisation inter-cadrans a été réalisée par la méthode des quantiles (fonction *normalize.quantiles* dans le package *affy* de R). Le but de cette méthode est de rendre la distribution des intensités des sondes identique pour chaque cadran (Bolstad *et al.*, 2003). Le design de la puce prévoit plusieurs sondes pour une même cible (génique ou intergénique), réparties le long de la cible. Les valeurs d'expression pour chaque cible ont été estimées par la médiane des intensités mesurées pour les sondes correspondantes.

L'estimation du nombre de gènes exprimés sur chaque puce a également été réalisée dans ANAIS. Cette détection statistique est basée sur le calcul d'une p-value et estime la probabilité qu'un gène soit exprimé plus fortement que les sondes aléatoires incluses dans la puce (Archer and Reese 2010).

Les analyses de groupement sur les données normalisées ont été réalisées dans le logiciel MeV 4.6.1, en utilisant un algorithme hiérarchique agglomératif à liaison moyenne (HCL, *average linkage*) et une estimation de la distance par le coefficient de corrélation de Pearson.

Les tests statistiques de Student et ANOVA ont été réalisés dans le logiciel R 2.11.1, avec correction de la p-value en q-value par l'algorithme de Benjamini-Hochberg qui permet de prendre en compte les tests multiples (Benjamini and Hochberg 1995). La recherche des gènes différentiellement exprimés dans une condition donnée a été réalisée, en utilisant l'expression dans la condition "glucose" comme référence. Les variations d'expression (*fold-change*, FC) ont été calculées pour chaque gène selon la formule suivante :

$$FC = \frac{\text{expression moyenne condition X}}{\text{expression moyenne condition Glucose}}$$

L'expression des gènes remplissant les deux critères suivants a été considérée comme variable :

- une variation d'expression d'au moins un facteur 2 par rapport au glucose  
soit :  $FC \geq 2$  ou  $FC \leq 0,5$
- un test de Student significatif avec un FDR de 5% (FDR : *False Discovery Rate*)  
soit :  $q\text{-value} \leq 0,05$

La recherche de catégories fonctionnelles *Gene Ontology* (GO Terms) sur-représentées dans les listes de gènes différentiellement exprimés a été réalisée avec le logiciel BiNGO 2.42 implémenté dans Cytoscape 2.7.0 (Maere *et al.*, 2005), en utilisant un test hypergéométrique corrigé par la méthode de Benjamini-Hochberg, et un FDR de 10%. Ce test permet de détecter des catégories fonctionnelles dont la proportion est significativement plus grande dans la liste de gènes donnée que dans l'ensemble du génome.

### III- Résultats et Discussion

#### III-1 Une puce à ADN pangénomique pour *Z. galactanivorans*

Dans l'optique de caractériser les régulations métaboliques liées à l'utilisation de polysaccharides chez *Z. galactanivorans*, une puce à ADN permettant l'étude du transcriptome à l'échelle du génome a été élaborée à partir des données issues du séquençage complet de son génome. Douze sondes différentes de 60 nucléotides ont été recherchées pour les 4738 gènes (Tableau 5-1).

Nombre de séquences codantes de départ	4738
Nombre de séquences avec au moins une sonde	4482
Nombre de séquences sans sonde	256
Nombre total de sondes "géniques"	52389

**Tableau 5-1 : Résultat de la recherche des sondes dans les régions codantes**

Pour certaines régions codantes, aucune sonde répondant aux critères de Nimblegen n'a pu être dessinée. Dans le cas de 7 gènes, toutes les sondes trouvées étaient identiques à celle d'un gène paralogue et ont donc été abandonnées. Certaines régions trop courtes n'ont pas permis de trouver les 12 sondes recherchées. Finalement, les sondes choisies couvrent 95% des gènes de *Z. galactanivorans*, avec un nombre moyen de 11,6 sondes par gène. Sur les 256 gènes non représentés sur la puce, 90% codent des protéines de fonction inconnue. Les annotations des autres gènes absents de la puce sont données dans le Tableau 5-2. On remarque par exemple que 5 gènes codant des protéines ribosomales et 6 gènes codant des transposases n'ont pas permis le dessin de sondes oligonucléotidiques. Ceci vient du fort taux de similitude entre ces gènes paralogues.

Gene ID	Annotation
zobellia_81	Ribosomal protein L31
zobellia_291	Xre-type transcriptional regulator
zobellia_595	Ribosomal protein L28
zobellia_761	PorT protein
zobellia_767	Ribosomal protein S15
zobellia_874	Carbohydrate binding module, family CBM48
zobellia_900	Endonuclease
zobellia_990	Transposase, family 12
zobellia_1049	N-acetylneuraminase lyase
zobellia_1291	Alkyltransferase-like protein 1
zobellia_2048	Single-stranded DNA-binding protein
zobellia_2139	Peptidyl-tRNA hydrolase
zobellia_2256	Cytochrome cbb3 oxidase, CcoQ subunit
zobellia_2519	Ribosomal protein S17
zobellia_2609	CsbD-like protein
zobellia_2786	Transposase, family 12
zobellia_3015	Heat shock protein 15
zobellia_3037	Ribosomal protein L35
zobellia_3062	Transposase (pseudogene), family 8
zobellia_3284	Transposase, family 11
zobellia_3422	Transposase, family 11
zobellia_3586	Helix-turn-helix, AraC type
zobellia_3831	Gliding motility permease, ABC-2 type exporter
zobellia_3921	Addiction module antidote protein
zobellia_4005	Transposase, family 12
zobellia_4016	Glyoxalase superfamily protein

**Tableau 5-2 : Listes des gènes de fonction connue non représentés sur la puce Nimblegen**

Le format de puce Nimblegen choisi (4\*72k) peut accueillir jusqu'à 72000 sondes. Afin de compléter les positions libres sur la puce, des sondes spécifiques des régions intergéniques du génome de *Z. galactanivorans* ont été recherchées. La quantification du taux de transcription de zones séparant des gènes successifs devrait permettre de mieux prédire l'existence d'opérons, et d'analyser les éventuels événements de transcription alternative en fonction des conditions de culture. Des sondes de 50 à 75 nucléotides ont été recherchées dans les régions intergéniques, avec un espacement moyen de 14 nucléotides entre sondes. Le bilan de cette recherche est donné dans le Tableau 5-3.

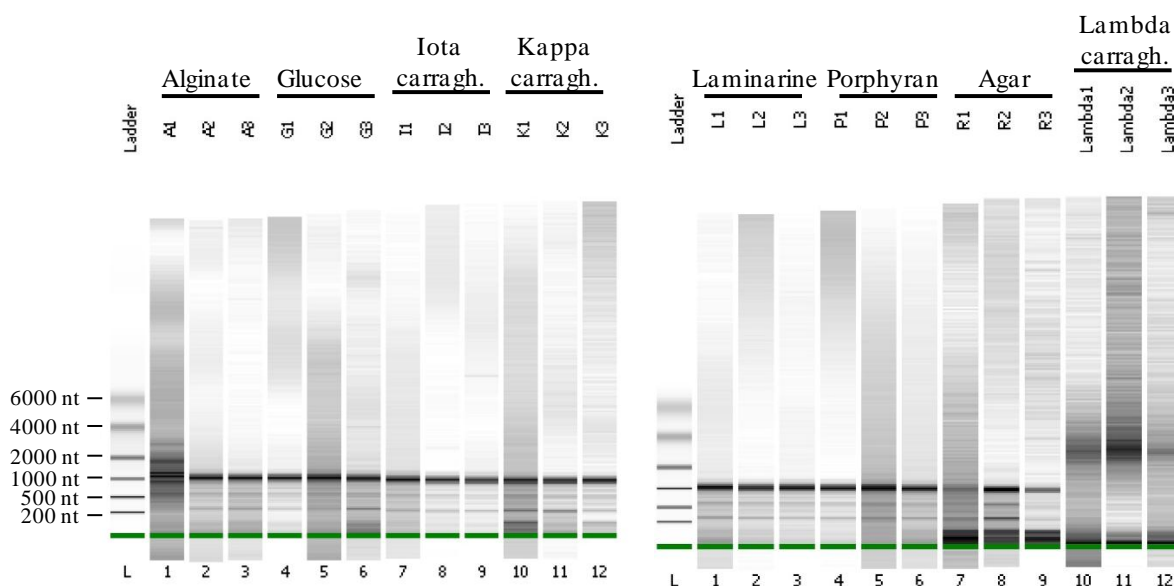
Nombre de séquences intergéniques de départ	4151
Nombre de séquences avec au moins une sonde	2595
Nombre de séquences sans sonde	1556
Espacement moyen des sondes	14 nt
Nombre total de sondes intergéniques	19435

**Tableau 5-3 : Résultat de la recherche de sondes dans les régions intergéniques**

## III-2 Contrôles qualité et validation des données transcriptomiques

### III-2-a Contrôles des échantillons biologiques

Les ARNs totaux extraits de *Z. galactanivorans* ont été analysés au spectrophotomètre afin de vérifier qu'ils remplissaient les critères définis par Nimblegen, à savoir : une concentration supérieure à  $1 \mu\text{g} \cdot \mu\text{L}^{-1}$ , un rapport  $A_{260}/A_{280}$  supérieur à 1,8 (dénotant une faible contamination par des protéines) et un rapport  $A_{260}/A_{230}$  supérieur à 1,8 (dénotant l'absence de contamination par des sels ou des solvants). Après les étapes de transcription inverse et de synthèse du deuxième brin, les ADNc des 24 échantillons biologiques (8 conditions en triplicat) répondaient bien aux critères de concentration et de pureté définis par Nimblegen ( $C \geq 100 \text{ ng} \cdot \mu\text{L}^{-1}$ ;  $A_{260}/A_{280} \geq 1,8$ ;  $A_{260}/A_{230} \geq 1,8$ ). Les profils des ADNc ont également été analysés (Figure 5-5).



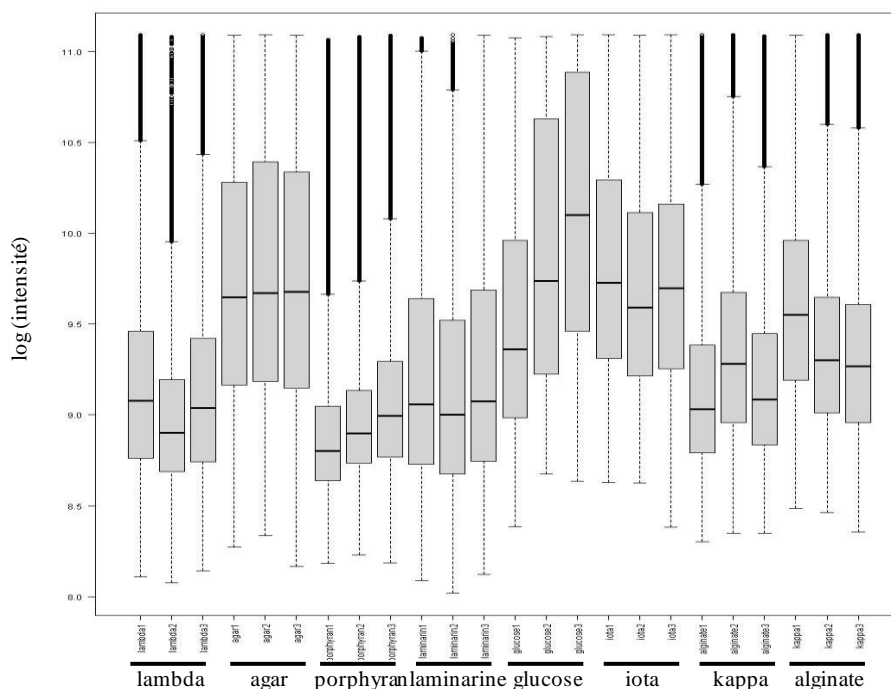
**Figure 5-5 : Analyse des profils des ADNc double-brin avant le marquage**

On constate que la plupart des échantillons ont un profil similaire. Cependant, les échantillons provenant de cultures en lambda-carraghénane semblent très différents, avec un profil atypique, ce qui peut dénoter un problème lors de la synthèse des ADNc. Les échantillons A1, R1, R2 et R3 ont également un profil particulier, mais de qualité tout à fait exploitable pour Nimblegen. Tous les échantillons d'ADNc (y compris "lambda") ont été marqués à la cyanine 3 et hybridés sur les puces à ADN.



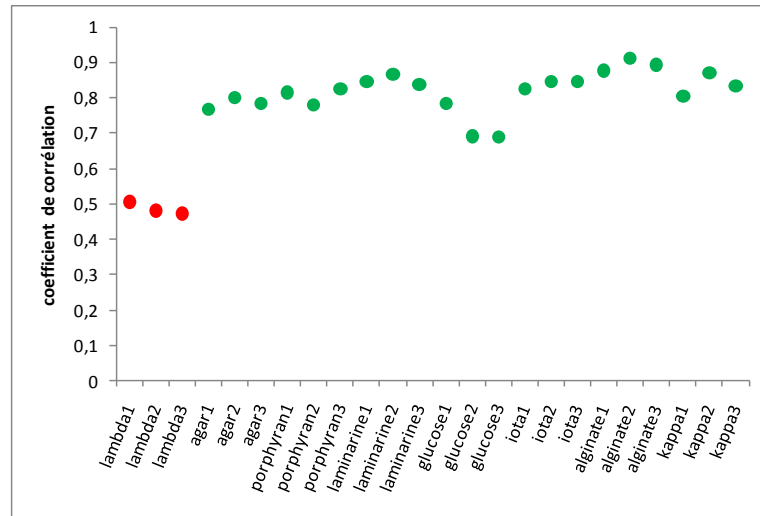
### III-2-b Contrôles qualité des données brutes

La distribution des données brutes issues du passage des puces au scanner a été contrôlée (Figure 5-6). On constate des différences d'intensité du signal mesuré selon les puces. Ces différences devront être corrigées par la normalisation inter-cadrons.



**Figure 5-6 : Graphique en boîte à moustaches des données brutes**

Afin de vérifier la cohérence des données, les coefficients de corrélation de Pearson entre les valeurs de chaque échantillon et l'échantillon médian (moyenne des 24 valeurs d'expression par gène) ont été calculés et sont représentés sur la Figure 5-7. Ce type de graphique permet de repérer des échantillons présentant une distribution atypique des signaux. On remarque que les trois échantillons provenant de culture en lambda-carraghénane n'ont qu'une faible corrélation avec l'échantillon médian calculé et se différencient clairement des autres. Ce phénomène est sans doute à relier avec les profils atypiques obtenus pour les ADNc de ces trois échantillons et peut poser un problème quant à leur utilisation ultérieure dans le jeu de données.



**Figure 5-7 : Coefficient de corrélation de Pearson des données brutes avec l'échantillon médian**

### III-2-a Contrôle des données normalisées

La distribution des données après normalisation par la méthode des quantiles est représentée sur la Figure 5-8. La normalisation a bien permis d'homogénéiser la distribution des signaux sur toutes les puces, en leur donnant des moyennes et des variances similaires. Cette étape permet de corriger les effets liés à l'expérimentation (biais dans les quantités d'ARN ou d'ADNc utilisées, efficacités de marquage ou d'hybridation différentes selon les échantillons, différences de détection de la fluorescence, etc.) et non à des variations biologiques. La normalisation inter-puces est basée sur plusieurs postulats dont : (i) la majorité des gènes sur une puce ont une expression constante et (ii) globalement l'expression est régulée autant positivement que négativement. Ces hypothèses sont assez bien vérifiées sur des puces pangénomiques telles que celle utilisée dans cette étude.

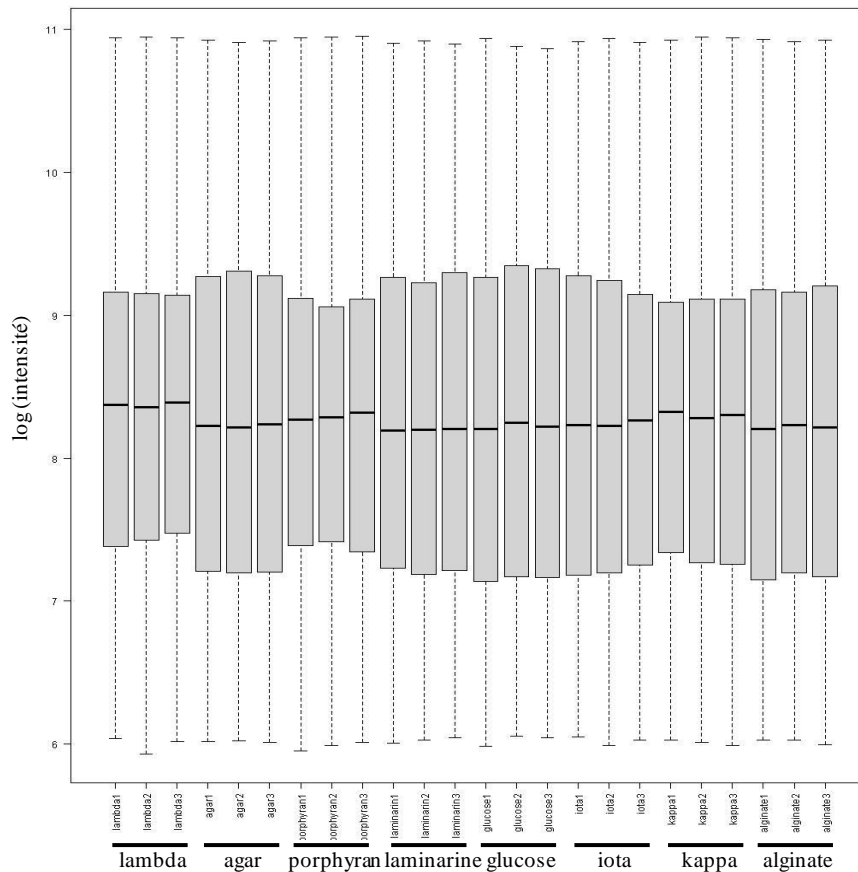
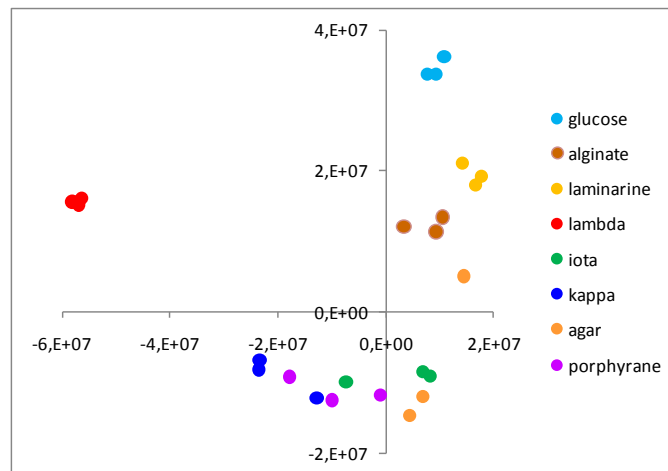


Figure 5-8 : Graphique en boîtes à moustaches des données normalisées

Une étude du comportement des trois répliquats biologiques pour une même condition a été réalisée en Analyse en Composante Principale (PCA). Les résultats sont représentés sur la Figure 5-9.



**Figure 5-9 : Représentation graphique de l'analyse PCA sur les données normalisées, selon les deux axes principaux**

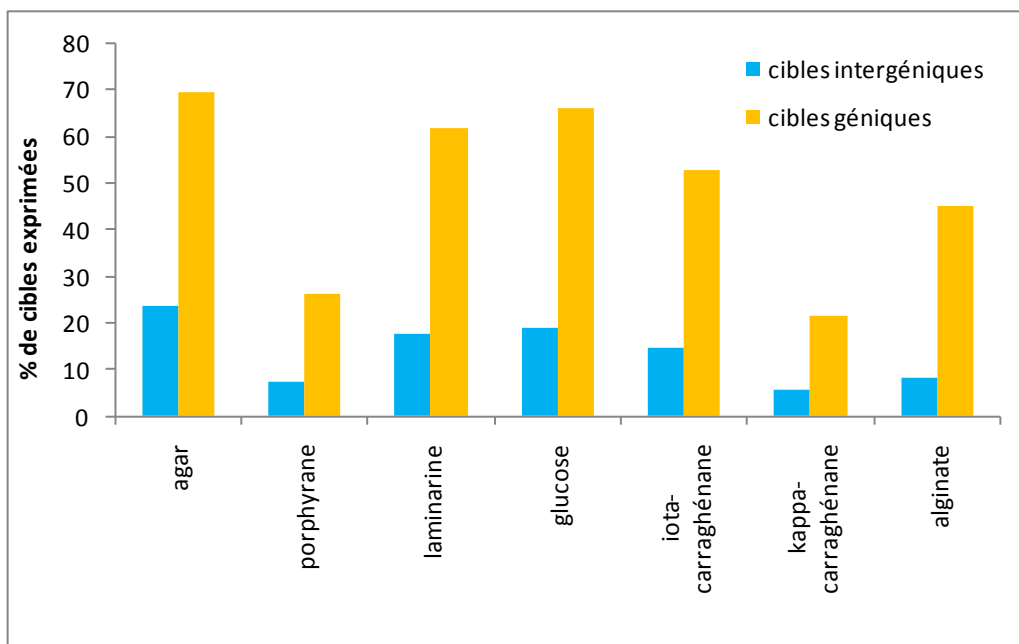
Sur ce graphique, on remarque que les différents répliquats biologiques d'une même condition sont relativement proches. Ceci dénote une bonne reproductibilité des valeurs d'expression des gènes et montre également que les variations dues à des phénomènes biologiques sont plus importantes que les biais expérimentaux potentiels. On remarque de plus que les trois répliquats obtenus à partir de cultures en lambda-carraghénane sont très éloignés de tous les autres et semblent encore une fois avoir un comportement particulier. Les trois répliquats "glucose" sont également à part, ce qui peut dénoter des différences majeures du métabolisme lors de l'utilisation de monosaccharides vs. polysaccharides.

La proportion de gènes exprimés a été estimée sur chaque cadran, en considérant que les cibles dont le signal ne dépassait pas celui des sondes aléatoires n'étaient pas transcrites. Seulement 5% des cibles (géniques + intergéniques) ont été détectées comme significativement exprimées dans la condition "lambda-carraghénane", contre une moyenne de 35% pour toutes les autres conditions ( $p$ -value < 0.01) Cette proportion semble extrêmement faible, et montre très probablement un grave défaut des trois répliquats concernés. En résumé, les échantillons provenant de cultures en lambda-carraghénane présentent un profil d'ADNc anormal, une faible corrélation avec l'échantillon médian, se démarquent nettement de tous les autres et n'ont conduit qu'à très peu d'hybridations spécifiques sur les puces. Par conséquent, il a été décidé d'abandonner l'analyse de cette condition, dont les résultats risqueraient d'être trop entachés d'erreur. Les intensités correspondant à ces trois répliquats ont été supprimées du jeu de données brutes, qui a ensuite été normalisé à nouveau.

### III-3 Analyse globale de l'expression dans les différentes conditions

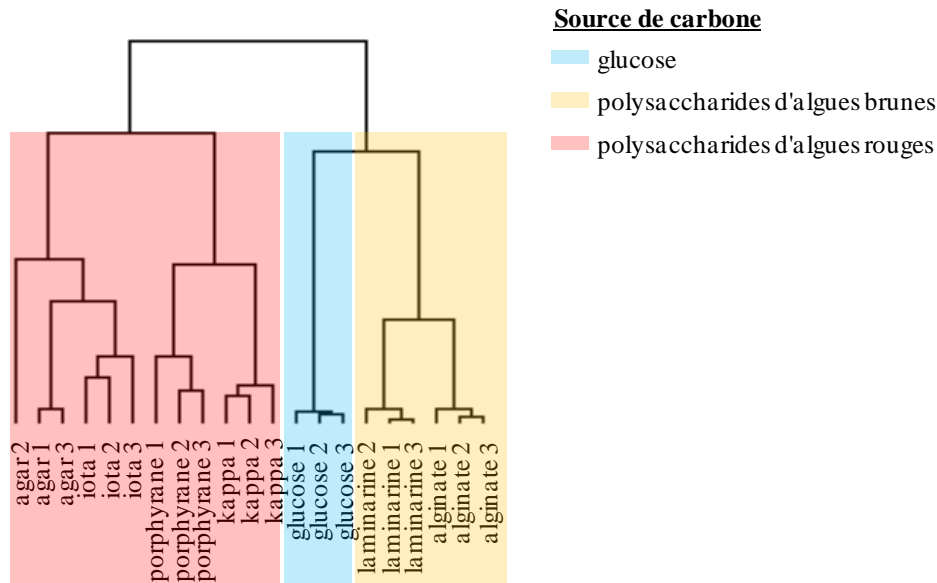
A la suite des étapes présentées précédemment, le jeu de données final utilisé dans cette étude comprend les valeurs d'intensités de 2595 cibles intergéniques et 4483 gènes, dans 7 conditions de culture en triplicats.

Dans un premier temps, le pourcentage de cibles significativement exprimées dans chaque condition a été estimé (Figure 5-10). Comme attendu, la proportion de cibles intergéniques exprimées est faible (entre 5% et 23%) et toujours plus basse que celle des cibles dans les régions codantes. On peut donc s'attendre à ce que les zones intergéniques exprimées soient présentes sur des ARN messagers polycistroniques. La détection de telles zones devraient faciliter la prédiction de nouveaux opérons dans le génome de *Z. galactanivorans*. On constate d'autre part des disparités importantes de transcription des régions codantes selon les conditions de culture. La proportion de gènes transcrits semble assez faible lorsque les bactéries sont cultivées en présence de porphyrane ou de kappa-carraghénane (26% et 21%, respectivement). En revanche, cette proportion est plus élevée si les cellules sont en présence de glucose ou d'agar, par exemple (66% et 69%, respectivement). Ces écarts importants montrent à quel point le transcriptome de *Z. galactanivorans* est influencé par la source de carbone mise à disposition.



**Figure 5-10 : Pourcentage de cibles significativement exprimées dans les différentes conditions de culture (p-value < 0.01)**

Une analyse de groupement a été effectuée sur les différentes puces, afin de définir des groupes de conditions ayant un profil similaire. Le dendrogramme généré est représenté sur la Figure 5-11.



**Figure 5-11 : Analyse de groupement du transcriptome des sept conditions de culture**

Cette analyse confirme tout d'abord que les triplicats biologiques se regroupent de façon satisfaisante, ce qui avait déjà été observé sur le graphique de l'analyse en composantes principales (Figure 5-9). D'autre part, et de façon plus inattendue, il semble possible de discriminer différents groupes hiérarchiques selon la provenance de la source de carbone ajoutée dans le milieu de culture pour *Z. galactanivorans*. Ainsi, on peut construire trois grands groupes : un groupe pour les cultures en présence de glucose (utilisé comme référence dans la suite de l'étude) et deux groupes bien distincts pour les cultures dont la source de carbone provient soit d'algues rouges, soit d'algues brunes. Un tel regroupement était déjà visible sur l'analyse en PCA (Figure 5-9). Le transcriptome de *Z. galactanivorans* est donc plus similaire lorsque les cellules sont cultivées sur des polysaccharides du même phylum d'algues (rouges vs. brunes), bien que de structures chimiques complètement différentes. Il faut ici noter que les sources de carbone utilisées dans l'expérience sont des polysaccharides purifiés et non des extraits pariétaux bruts. On peut donc s'affranchir d'une contamination d'une des sources de carbone par d'autres composés provenant des algues utilisées lors de l'extraction. Le résultat de l'analyse de groupement pourrait ainsi indiquer que des régulations communes du transcriptome existent chez *Z. galactanivorans* en fonction de l'origine du

substrat dégradé. On peut faire l'hypothèse de l'existence d'un "programme" de dégradation adapté soit aux algues brunes, soit aux algues rouges. Il faut cependant remarquer qu'au sein du groupe des cultures en présence de polysaccharides d'algues rouges, on ne distingue pas ceux provenant d'algues agarophytes (agar et porphyrane) des carraghénophytes (iota- et kappa-carraghénanes).

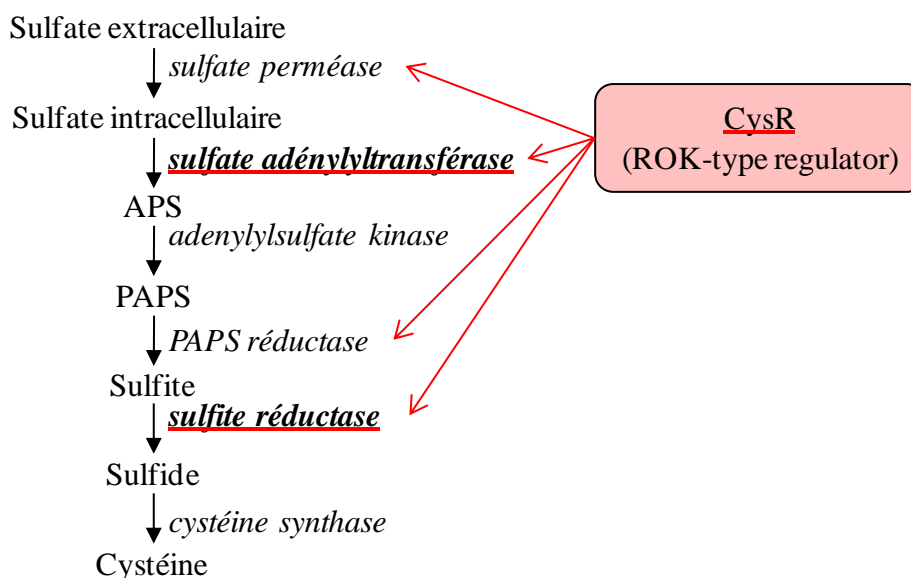
Avant de rechercher les gènes dont l'expression change en fonction des conditions de culture, il est intéressant de considérer ceux qui sont fortement transcrits quelle que soit la source de carbone. En effet, de tels gènes peuvent nous renseigner sur les voies métaboliques fortement exprimées indépendamment de la source de carbone chez *Z. galactanivorans*. Les gènes non-différentiellement exprimés ont été détectés par ANOVA sur toutes les conditions ( $q\text{-value} > 0,1$ ). Cette absence de différence significative entre conditions autorise au calcul de la moyenne de leur expression toutes conditions confondues. Les soixante premiers gènes (correspondants aux 10 premiers centiles les plus exprimés) ont été considérés: 21 d'entre eux sont de fonction inconnue; les annotations des autres gènes sont reportées dans le Tableau 5-4 (page suivante). Le paragraphe suivant présente plus de détails sur certains de ces gènes.

Gene ID	Catégorie fonctionnelle	Annotation
zobellia_651	Métabolisme protéique	Translation elongation factor Tu
zobellia_726	Métabolisme protéique	Sec-independent protein translocase
zobellia_910	Métabolisme protéique	Glutamate synthase [NADPH] small chain
zobellia_987	Métabolisme protéique	GTP-binding protein Era
zobellia_2506	Métabolisme protéique	Ribosomal protein S12
zobellia_4404	Métabolisme protéique	Chorismate synthase
zobellia_4547	Métabolisme protéique	Carboxypeptidase, family M20
zobellia_4723	Métabolisme protéique	2-Isopropylmalate synthase
zobellia_552	Méthyltransférase	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
zobellia_1293	Méthyltransférase	tRNA (guanine-N(7)-)-methyltransferase
zobellia_2174	Méthyltransférase	Adenine-specific DNA methyltransferase
zobellia_2496	Méthyltransférase	23S rRNA (guanosine-2'-O-)-methyltransferase
zobellia_199	Acquisition de nutriments	SusD/RagB family lipoprotein
zobellia_1942	Acquisition de nutriments	TonB-dependent Receptor
zobellia_2849	Acquisition de nutriments	TonB-dependent Receptor
zobellia_1125	Métabolisme central	Methylmalonyl-CoA mutase small subunit
zobellia_3591	Métabolisme central	Glutamate dehydrogenase
zobellia_4607	Métabolisme central	Dihydropyruvate dehydrogenase, E3 component of pyruvate dehydrogenase complex
zobellia_919	Métabolisme de monosaccharides	D-tagatose 3-epimerase
zobellia_2175	Métabolisme de monosaccharides	Xylose isomerase-like TIM barrel protein
zobellia_2591	Métabolisme de monosaccharides	Aldose 1-epimerase family protein
zobellia_841	Métabolisme du sulfate	Sulfite reductase [ferredoxin]
zobellia_842	Métabolisme du sulfate	Sulfate adenylyltransferase subunit 1
zobellia_843	Métabolisme du sulfate	Sulfate adenylyltransferase subunit 2
zobellia_114	Gestion du stress	Superoxide dismutase [Cu-Zn]
zobellia_3484	Gestion du stress et autres fonctions	Universal stress protein
zobellia_1989	Métabolisme de polysaccharides	Polysaccharide lyase, family PL12
zobellia_4265	Métabolisme de polysaccharides	Iota-carrageenase, family GH82
zobellia_840	Synthèse de porphyrines	Uroporphyrinogen-III C-methyltransferase
zobellia_2232	Synthèse de porphyrines	Oxygen-independent coproporphyrinogen III oxidase
zobellia_3488	Efflux de calcium	Calcium-transporting P-type ATPase
zobellia_862	Homéostasie du pyrophosphate	Pyrophosphate-energized proton pump
zobellia_3676	Métabolisme des purines	GTP pyrophosphokinase
zobellia_1608	Synthèse d'acides gras	3-oxoacyl-[acyl-carrier-protein] synthase III
zobellia_3501	Transduction du signal	ROK-type transcriptional regulator
zobellia_4499	Transport / Adhésion	OmpA-like lipoprotein
zobellia_1289	Inconnue	NifU family protein
zobellia_1505	Inconnue	Carbonate dehydratase
zobellia_4519	Inconnue	Prohibitin-like protein

**Tableau 5-4 : Listes des gènes de *Z. galactanivorans* fortement exprimés dans toutes les conditions**  
Seuls les gènes ayant une annotation connue et appartenant aux 10% les plus exprimés sont donnés.



Comme attendu, on retrouve un grand nombre de gènes impliqués dans les fonctions de base des cellules, notamment le métabolisme des protéines, des acides nucléiques, des sucres simples, le métabolisme central (autour du cycle de Krebs) ou l'acquisition de nutriments. On remarque qu'un cluster de gènes impliqués dans le métabolisme du sulfate est fortement exprimé dans les 7 conditions testées. Ce cluster comprend *zobellia\_841*, codant une sulfite réductase, et *zobellia\_842/843*, codant les deux sous-unités d'une sulfate adénylyltransferase. Les microorganismes doivent réduire le sulfate  $\text{SO}_4^{2-}$  en sulfide par un processus nécessitant de l'énergie, afin d'incorporer le soufre dans les molécules organiques. La sulfate adénylyltransferase (ou ATP sulfurylase) catalyse l'activation du sulfate par un ATP pour produire le 5'-phosphosulfate (APS), dérivé essentiel à l'assimilation du soufre à partir de sulfate inorganique. La sulfite réductase catalyse quant à elle la réduction du sulfite  $\text{SO}_3^{2-}$  en sulfide  $\text{S}^{2-}$ . Le rôle prépondérant de ces deux enzymes pour l'assimilation du sulfate chez les bactéries est représenté sur la Figure 5-12. Des gènes candidats codant les autres enzymes de la voie ont été trouvés chez *Z. galactanivorans*. La concentration en sulfate dans l'eau de mer est connue pour être élevée, de l'ordre de 25 à 30 mM, ce qui fait de cet ion une source de soufre abondante et facile d'accès en milieu marin (Giordano *et al.*, 2005). Ceci peut permettre d'expliquer la forte expression de ces gènes impliqués dans l'assimilation du sulfate chez une bactérie marine.



**Figure 5-12 : Schéma simplifié de l'assimilation du sulfate inorganique chez les bactéries**

Les enzymes dont les gènes sont fortement exprimés dans toutes les conditions testées sont soulignés en rouge. APS : 5'-phosphosulfate; PAPS : 3'-phosphoadénylyl sulfate. Les flèches rouges symbolisent une activation de la transcription. D'après Schmidt et Jager (1992) et Ruckert *et al.* (2008).

Il est également intéressant de noter qu'un autre gène fortement exprimé, *zobellia\_3501*, code un régulateur transcriptionnel de la famille ROK. Cette famille de protéines regroupe principalement des kinases à sucres et des régulateurs de la transcription d'opérons du catabolisme des sucres. C'est cette dernière fonction qui a été prédite lors de l'annotation du génome de *Z. galactanivorans*. Cependant, un membre de cette même famille, CysR, est connu comme régulateur transcriptionnel de l'assimilation du sulfate chez *Corynebacterium glutamicum* (Ruckert *et al.*, 2008). Une hypothèse alternative pourrait donc être que la protéine codée par *zobellia\_3501* contrôle l'assimilation du sulfate chez *Z. galactanivorans*, ce qui concorderait avec la forte expression d'autres gènes participant à cette voie métabolique. Cette hypothèse est étayée par le fait que l'ORF en question se trouve juste en amont de deux gènes codant des sulfatases (*zobellia\_3502* et *zobellia\_3503*).

Deux gènes codant des polysaccharidases sont fortement exprimés quelle que soit la source de carbone, suggérant une absence de régulation au niveau transcriptionnel. Le premier, *zobellia\_1989*, code une enzyme de la famille CAZy PL12, dont les seuls membres caractérisés à ce jour sont des héparine-sulfate lyases (Cantarel *et al.*, 2009). Néanmoins, *Z. galactanivorans* n'est pas capable d'utiliser l'héparane sulfate (T. Barbeyron, communication personnelle), suggérant que cette enzyme agit sur un autre polysaccharide uronique. Le deuxième gène, *zobellia\_4265*, code une iota-carraghénase (CgiA1) de la famille GH82. Cependant, lors d'une recherche de galactane-hydrolases produites par *Z. galactanivorans*, il a été montré que l'activité iota-carraghénase était détectable dans le surnageant de culture seulement après induction par du carraghénane (Potin 1992). Trois hypothèses non exclusives peuvent être envisagées. Tout d'abord, l'expérience de Potin (1992) a été réalisée en milieu Zobell (fort apport de tryptone) supplémenté en iota-carraghénane, ce qui est différent d'une culture en milieu minimum. Deuxièmement, l'activité de CgiA1 pourrait être contrôlée au niveau post-transcriptionnel, par un mécanisme non-élucidé mais dépendant de la présence de carraghénane. Enfin, *Z. galactanivorans* possède deux autres iota-carraghénases (CgiA2 et CgiA3), dont l'expression pourrait être régulée au niveau transcriptionnel.

L'ORF *zobellia\_3484*, aussi fortement exprimée, est annotée "Universal Stress Protein" (Usp). Cette famille est nommée d'après la protéine UspA d'*E. coli*, dont l'expression est fortement induite lors de différents stress tels que le manque de carbone ou d'azote, l'arrivée en phase stationnaire, le choc thermique, l'exposition à des oxydants ou à des métaux, etc. (Nystrom and Neidhardt 1994). Les membres de la famille UspA sont très nombreux, et sont généralement considérés comme des marqueurs d'un stress cellulaire et d'un arrêt de la

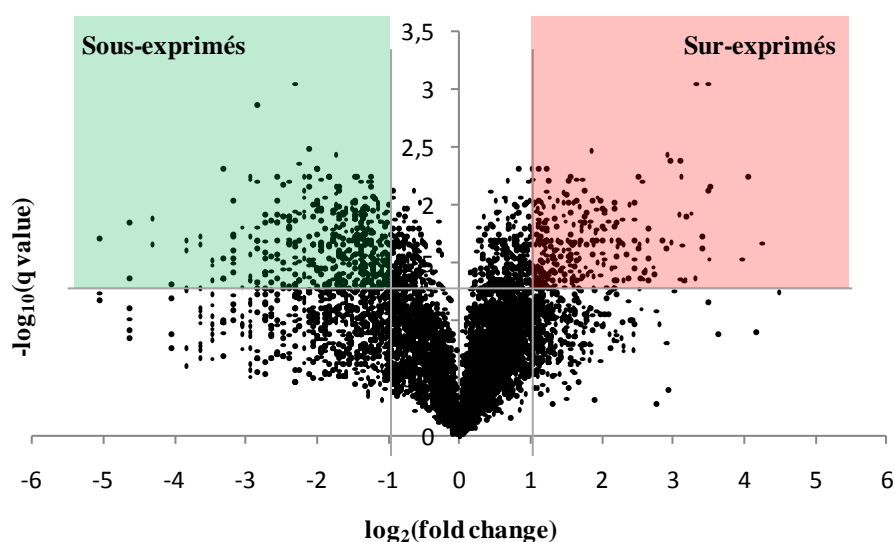
croissance bactérienne (Kvint *et al.*, 2003). Au premier abord, il peut donc paraître surprenant de retrouver un membre de cette famille UspA parmi les gènes les plus fortement exprimés dans l'expérience présentée ici. Si l'annotation est bonne, cela signifierait que les conditions de culture utilisées induisent un stress généralisé chez *Z. galactanivorans*. Cependant, des études de protéines Usp-like ont montré que certains membres de la famille pourraient avoir évolué vers des fonctions différentes et non-liées au stress. Ainsi, la bactérie *E. coli* possède 6 protéines de la famille UspA (UspA, -C, -D, -E, -F, -G). Parmi celles-ci, il a été prouvé que UspC n'est pas impliquée dans la résistance au stress, mais est essentielle, comme UspE, à la mobilité cellulaire (Nachin *et al.*, 2005). De même, UspF et UspG n'ont que peu d'effet sur la défense face au stress oxydant, mais favorisent l'adhésion cellulaire par rapport à la mobilité. Chez *Porphyromonas gingivalis*, qui appartient au phylum des *Bacteroidetes* comme *Z. galactanivorans*, le gène *uspA* est certes requis pour la résistance au stress mais également pour la formation d'un biofilm (Chen *et al.*, 2006). L'expression d'*uspA* est induite plus de mille fois dans le biofilm par rapport à une croissance en phase planctonique. Le génome de *Z. galactanivorans* contient 10 gènes avec un domaine homologue à la famille UspA. Il est donc probable que les membres de cette famille multigénique aient aussi diverses fonctions, non obligatoirement liées au stress. Une hypothèse quant au gène *zobellia\_3484*, trouvé fortement exprimé quelle que soit la source de carbone, pourrait être son implication dans l'adhésion cellulaire et la formation de biofilm. En effet, les membres du genre *Zobellia* sont fréquemment retrouvés associés à des surfaces (notamment des surfaces d'algues) où ils ont tous la capacité de se mouvoir par glissement (Barbeyron *et al.*, 2001; Nedashkovskaya *et al.*, 2004). De plus, *Z. galactanivorans* a été utilisée comme bactérie modèle pour sa capacité à former des biofilms (Salaün 2009).

Après cette analyse globale, je m'intéresserai principalement dans la suite du chapitre à la régulation du transcriptome en présence d'alginate. La détection de gènes sur- ou sous-exprimés devrait permettre d'approfondir la compréhension de la dégradation de ce polysaccharide par *Z. galactanivorans*. Cependant, je présenterai également quelques éléments de la régulation en présence de laminarine, afin d'illustrer l'existence d'un "programme" transcriptomique de dégradation des algues brunes.

### III-4 L'alginate induit une régulation massive du transcriptome

#### III-4-a Détection des gènes différentiels

Les gènes de *Z. galactanivorans* dont l'expression est différente en présence d'alginate par rapport au glucose ont été détectés comme présenté dans la section Matériels et Méthodes. Le principe d'identification peut être représenté graphiquement sur un Volcano-Plot (Figure 5-13). Cette approche a permis d'identifier 259 cibles sur-exprimées (régions géniques et intergéniques confondues) et 421 cibles sous-exprimées en présence d'alginate, réparties comme indiqué dans le Tableau 5-5.



**Figure 5-13 : Diagramme en Volcano-Plot de la comparaison des valeurs d'expression en présence d'alginate ou de glucose**

	Cibles sur-exprimées	Cibles sous-exprimées	% régulé
régions codantes	226	369	13,3
régions intergéniques	33	52	3,3
intervalle du <i>fold-change</i>	2,00 – 19,10	0,03 – 0,50	

**Tableau 5-5 : Répartition des cibles différentielles détectées lors de la comparaison de l'expression en présence d'alginate (référence = glucose)**

% régulé : proportion de cibles régulées par rapport au nombre total de cibles présent sur la puce

L'utilisation d'alginate comme source de carbone par *Z. galactanivorans* implique une régulation de l'expression de 595 gènes par rapport à la condition glucose, soit 13% de son génome. Les listes des gènes différentiellement exprimés sont reportées dans les Annexes. On

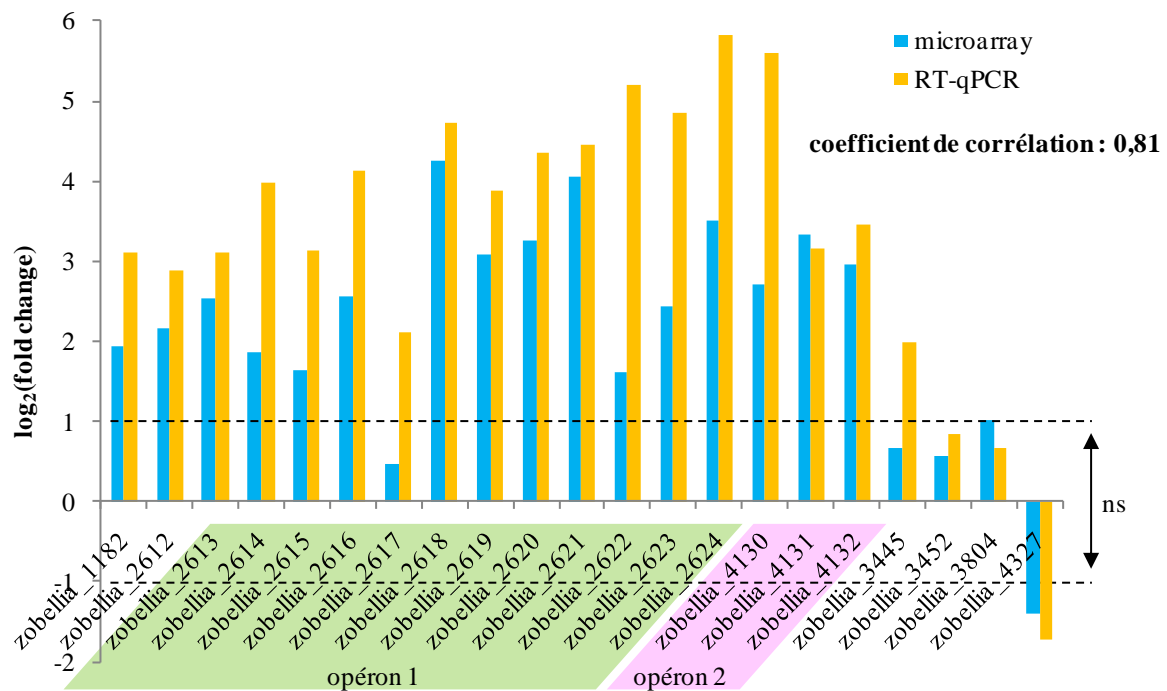
constate qu'un grand nombre de gènes est sous-exprimé, ce qui peut indiquer que les cellules adaptent leur métabolisme en se spécialisant vers la dégradation de l'alginate, et en réprimant d'autres voies. L'expression de 226 gènes est induite en présence d'alginate, et on peut penser que leurs fonctions sont liées à une des étapes de la perception et/ou de la dégradation du polysaccharide. Le nombre de gènes impliqué est important, et montre que les régulations mises en jeu doivent aller au-delà de la simple production de quelques enzymes supplémentaires. Ce nombre est par ailleurs cohérent avec une autre étude chez *Bacillus subtilis* (Ochiai *et al.*, 2007), dont la taille du génome (4112 gènes) est similaire à *Z. galactanivorans*. Chez cette bactérie détritivore, il a été montré que l'utilisation d'un polysaccharide pariétal de plantes terrestres, le rhamnogalacturonane, induit l'expression de 186 gènes (soit 4,5% contre 4,8% chez *Z. galactanivorans*).

Une grande partie des gènes de *Z. galactanivorans* dont l'expression est régulée par la présence d'alginate par rapport au glucose n'ont pas reçu de fonction prédite lors de l'annotation du génome. Ainsi, 115 des 226 gènes sur-exprimés (soit 50%) et 67 des 369 gènes sous-exprimés (soit 18%) sont de fonction inconnue. La proportion de gènes annotés est donc plus importante parmi ceux dont l'expression est diminuée. Ceci semble indiquer que les voies métaboliques concernées sont relativement mieux connues, ce qui a permis d'annoter plus de gènes. Parallèlement, cela montre qu'une grande partie des voies métaboliques responsables de la dégradation et de la perception de l'alginate reste à découvrir. Dans la suite de ce chapitre, je détaillerai les différentes voies répondant positivement ou négativement à la présence d'alginate, puis discuterai de l'éventualité d'un tronc commun de régulation par les polysaccharides d'algues brunes.

#### **III-4-b Comparaison avec les résultats obtenus en transcriptomique "ciblée"**

Cette étude transcriptomique à l'échelle du génome fait suite à une analyse plus ciblée portant sur un ensemble de gènes impliqués dans la dégradation de l'alginate (cf. Chapitre 4). L'étude ciblée a démontré que l'expression de 6 des 7 alginate lyases de *Z. galactanivorans* est induite par la présence du substrat, alors que celle de la septième enzyme est réprimée. Deux opérons alginolytiques dont la transcription est dépendante de la présence d'alginate dans le milieu ont par ailleurs été mis en évidence. La comparaison des résultats obtenus par cette approche ciblée d'une part et par l'hybridation de puces à ADN d'autre part permet d'évaluer la cohérence des données entre les deux méthodes. Comme représenté sur la Figure 5-14, les

taux de variation déterminés par les deux techniques suivent la même tendance, avec un bon coefficient de corrélation.

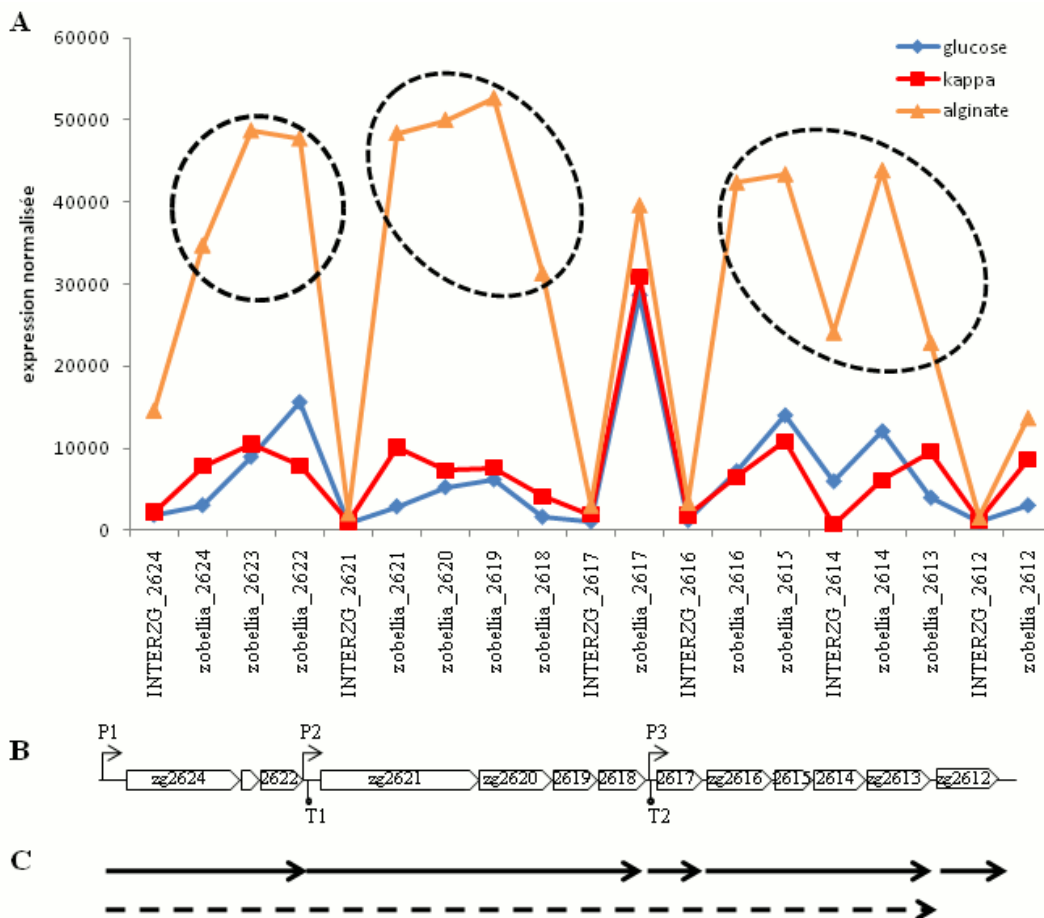


**Figure 5-14 : Comparaison des taux de variation obtenus en RT-qPCR et par les puces à ADN**

Les taux de variation (*fold change*) ont été calculés par rapport à la condition "glucose". Le coefficient de corrélation de Pearson entre les deux séries de résultats est donné sur le graphique. ns : intervalle de variation ne remplissant pas les critères définis dans la section Matériels et Méthodes.

De manière globale, les taux de variation trouvés semblent plus élevés en RT-qPCR qu'avec les puces à ADN, ceci étant dû à une plus grande sensibilité de la première méthode. L'analyse des puces à ADN conduit aux mêmes conclusions quant à l'expression des alginate lyases chez *Z. galactanivorans* : pour les enzymes AlyA1 à AlyA6 (zobellia\_1182, 2618, 2624, 4130, 4131 et 4132), l'expression est induite par la présence du substrat, alors qu'elle est réprimée pour AlyA7 (zobellia\_4327). De même, l'étude à l'échelle du génome confirme l'induction par l'alginate de la transcription des deux opérons alginolytiques identifiés. On constate cependant que le gène zobellia\_2617, qui code un facteur de transcription de la famille GntR, n'a pas été détecté comme différentiel sur les puces à ADN. En RT-qPCR, ce gène montre déjà le plus faible taux de variation de tous ceux de l'opéron 1 par rapport aux cultures en glucose. Ceci est sans doute à relier avec la forte expression constitutive de ce gène (cf. Chapitre 4), qui explique en partie qu'il ne soit pas détecté comme différentiel ici, bien que fortement transcrit.

L'approche ciblée présentée dans le Chapitre 4 avait par ailleurs permis de mettre en évidence des phénomènes de transcription sous forme d'ARN messagers polycistroniques. La présence sur la puce à ADN de sondes dirigées contre les régions intergéniques du génome de *Z. galactanivorans* permet de le vérifier. Comme présenté sur la Figure 5-15, trois régions de la zone comprise entre les gènes *zobellia\_2624* et *zobellia\_2612* sont transcrites en ARNm polycistroniques. Pour deux d'entre elles, les bornes des zones transcrites coïncident parfaitement avec les promoteurs et terminateurs prédits dans le Chapitre 4 (de P1 à T1 et de P2 à T2). Pour la zone *zobellia\_2616-2613*, il semble qu'un promoteur supplémentaire non-prédit existe en amont du gène *zobellia\_2616*. On remarque encore une fois sur ce graphique la forte expression du gène *zobellia\_2617* (codant un facteur de transcription *GntR-like*) quelles que soient les conditions de culture. Ce gène semble transcrit de manière isolée majoritairement, comme en témoigne le faible niveau d'expression des deux régions intergéniques flanquantes. Cependant, l'existence de longs ARN messagers montrée dans le Chapitre 4 (jusqu'à 18 kb) indique que l'ARN polymérase peut parfois franchir les terminateurs T1 et/ou T2 sans arrêter la transcription.

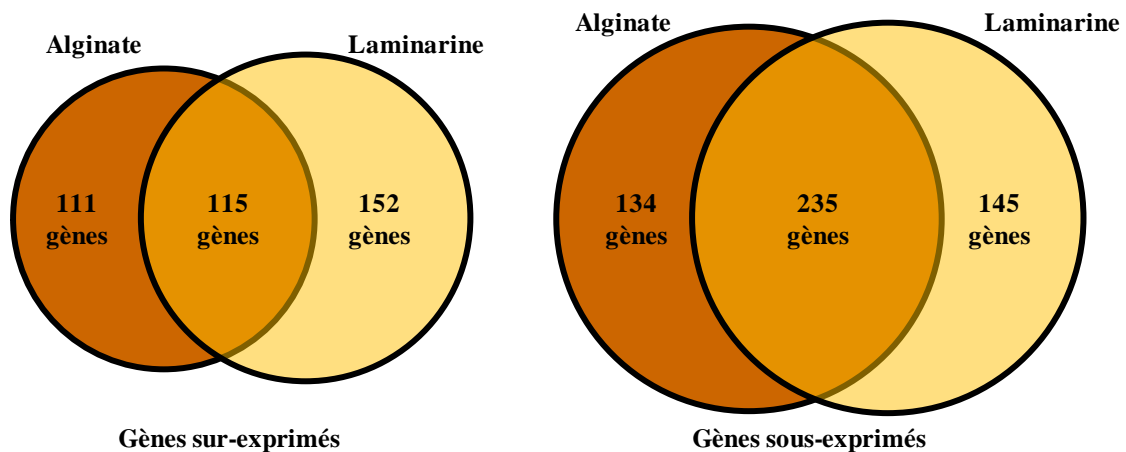


**Figure 5-15: Détection d'ARNm polycistroniques par l'approche *microarray***

A. Niveaux d'expression des gènes (*zobellia\_X*) et régions intergéniques (INTERZG\_X) en présence de glucose (losanges bleus), alginate (triangles orange) et kappa-carraghénane (carrés rouges, donné pour comparaison) comme seule source de carbone. B. Promoteurs (P) et terminateurs Rho-indépendants (T) prédits dans cette région (cf. Chapitre 4). C. ARN messagers correspondants. En trait pointillé : ARNm synthétisé lorsque l'ARN polymérase franchit les terminateurs sans se détacher.

### III-4-c Alginate et Laminarine : des régulations communes

L'analyse de groupement hiérarchique a montré que les échantillons provenant de culture en présence d'alginate et de laminarine ont des profils d'expression proches (cf. III-3). La recherche de gènes différentiels communs dans les deux conditions le confirme. En effet, ces deux polysaccharides d'algues brunes (l'un de paroi, l'alginate, et l'autre de stockage, la laminarine) régulent l'expression d'un cœur de gènes communs (Figure 5-16).



**Figure 5-16 : Représentation en diagramme de Venn des gènes dont l'expression est régulée en présence d'alginate ou de laminarine**

Ce résultat suggère que des cascades de signalisation communes sont mises en jeu lors de la détection d'alginate ou de laminarine *in vitro*. En termes de comportement de la bactérie dans son environnement naturel, ceci signifierait que *Z. galactanivorans* est en quelque sorte capable de détecter le type d'algue présent dans son milieu (ici, algue brune) et d'adapter son métabolisme en fonction. Les gènes dont la régulation est spécifique à une condition coderaient dans ce cas des protéines dont la fonction est étroitement liée à l'utilisation du polysaccharide en question. Dans la suite de ce chapitre, je n'aborderai pas dans le détail les gènes répondant spécifiquement à la laminarine. Je discuterai principalement les gènes dont l'expression est régulée par l'alginate, afin de dresser un schéma de l'utilisation de ce polysaccharide à la fois *in vitro* et dans le contexte biologique de la paroi d'une algue brune.



### III-5 Les gènes surexprimés en présence d'alginate

#### III-5-a Recherche de catégories fonctionnelles répondant à l'alginate

Comme mentionné plus haut, la présence d'alginate dans le milieu de culture induit l'expression de 226 gènes (Annexes 1 et 2). Une recherche des ontologies (termes GO) significativement sur-représentées dans ce lot de gènes a été effectuée afin de définir des processus biologiques prépondérants en présence d'alginate (Tableau 5-6).

GO Term	p-value <sup>a</sup>	Description	nombre de gènes
50794	0.07	regulation of cellular process	16
160	0.02	two-component signal transduction system (phosphorelay)	9
42122	0.00	alginic acid catabolic process	6
45488	0.06	pectin metabolic process	2

**Tableau 5-6 : Détection de termes GO sur-représentés parmi les gènes surexprimés en présence d'alginate**

Seules les ontologies des nœuds terminaux du réseau obtenu sont données.

a : test hypergéométrique, après correction de Benjamini-Hochberg

La sur-représentation des termes GO 50794 et 160 montre qu'une part importante du transcriptome de *Z. galactanivorans* en présence d'alginate est dédiée à la perception de son environnement et la transduction de signaux. Ceci permet aux cellules d'adapter leur métabolisme par rapport au substrat présent. La nature des senseurs et des régulateurs impliqués sera discutée par la suite, dans l'optique de mieux comprendre les cascades de signalisation mises en jeu. D'autre part, l'enrichissement du lot de gènes en terme GO 42122 montre bien la spécialisation du métabolisme des cellules vers la dégradation de l'alginate. Les 6 gènes retrouvés dans cette catégorie sont ceux codant les alginates lyases déjà mentionnés dans la partie précédente.

Il peut sembler plus surprenant de trouver une sur-représentation de l'ontologie GO 45488, qui suggérerait une induction d'un métabolisme lié à la pectine. Il faut noter qu'un des deux gènes assignés à cette catégorie, *zobellia\_2623*, appartient à l'opéron alginolytique et code un homologue de la protéine KdgF de *Dickeya dadantii* (auparavant *Erwinia chrysantemi*). Chez cette bactérie phytopathogène, la présence de pectine induit l'expression d'enzymes de dégradation, les pectate lyases (Hugouvieux-Cotte-Pattat *et al.*, 1996). Le gène *kdgF* appartient à un cluster de 4 gènes impliqués dans la dégradation de la pectine et code une

petite protéine (12,4 kDa) n'ayant aucune homologie avec des protéines de fonction connue et ne dégradant pas la pectine (Condemine and Robertbaudouy 1991). Une invalidation du gène *kdgF* par mutation réduit la réponse de la bactérie à la présence de pectine. En conséquence, KdgF pourrait jouer un rôle dans la perception et/ou la transduction du signal, mais non dans la dégradation de la pectine proprement dite. La localisation de *zobellia\_2623* dans un opéron alginolytique, et son induction par l'alginate, montrent que la famille des protéines KdgF pourrait avoir divergé pour participer à des voies cataboliques différentes. Le deuxième gène assigné à l'ontologie GO 45488 et dont l'expression est induite par l'alginate a été annoté en tant que pectate lyase de la famille PL1 (*pelA1*, *zobellia\_3770*). *Z. galactanivorans* possède un deuxième gène avec la même annotation (*pelA2*), mais il a été montré que la bactérie ne dégrade pas la pectine (Barbeyron T., communication personnelle). En conséquence, les données présentées ici suggèrent que la protéine codée par *zobellia\_3770* n'a pas la pectine pour substrat. Une hypothèse alternative pourrait être son implication dans la dégradation d'alginate ou d'oligo-alginate, ce qui serait une première pour la famille PL1 qui ne contient actuellement que des pectate ou pectine lyases.

L'approche *Gene Ontology* a montré que les cellules expriment un grand nombre d'acteurs de phénomènes de régulation et du catabolisme de l'alginate. Elle a été complétée par une classification manuelle des gènes (Figure 5-17). Certaines catégories fonctionnelles ont été mises en évidence, et leur implication potentielle dans la perception et le métabolisme de l'alginate est discutée ci-après. Les gènes induits par la laminarine ont été reportés sur le même schéma afin de souligner l'existence d'un lot de gènes partagé entre les deux conditions (cf. III-4-c).

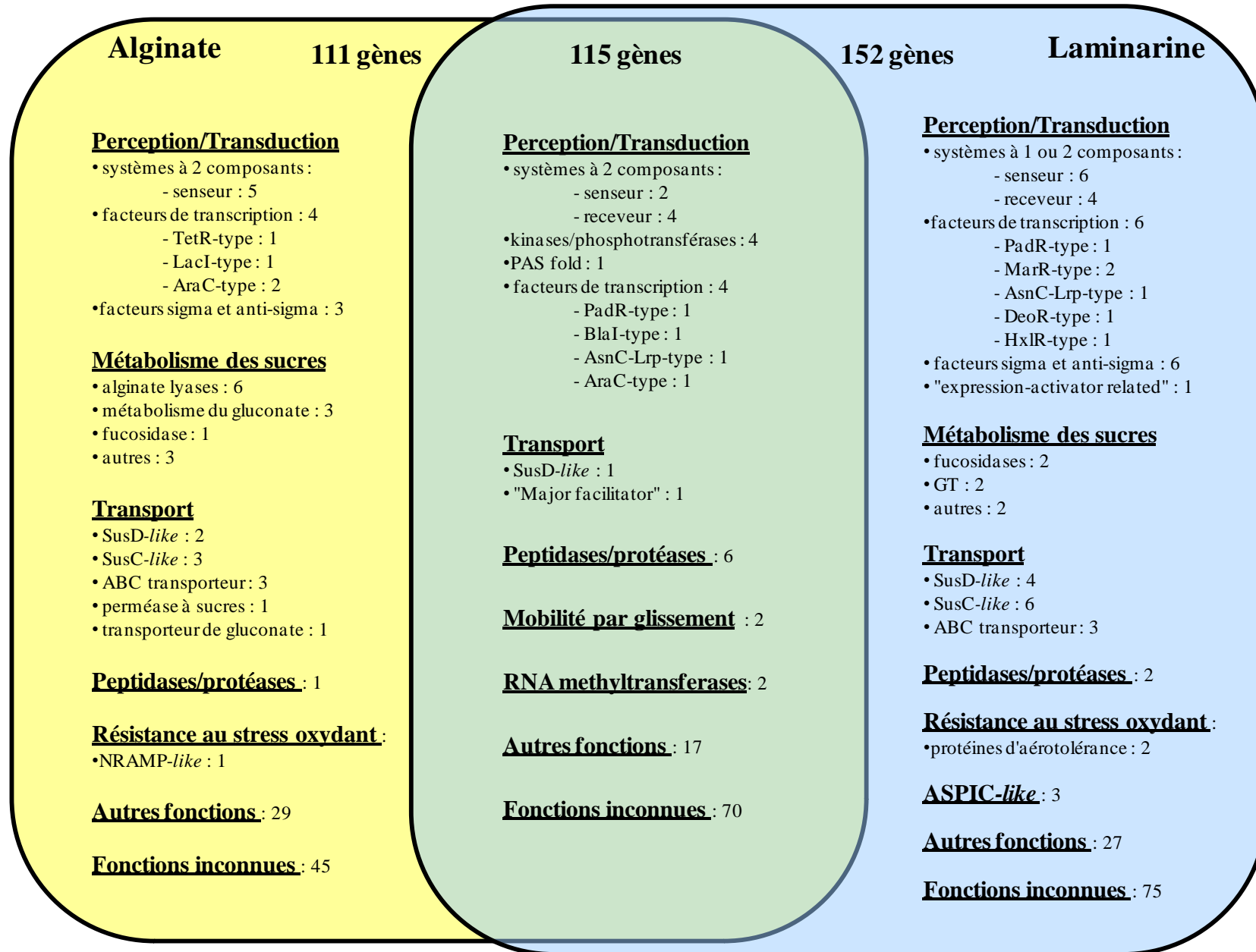
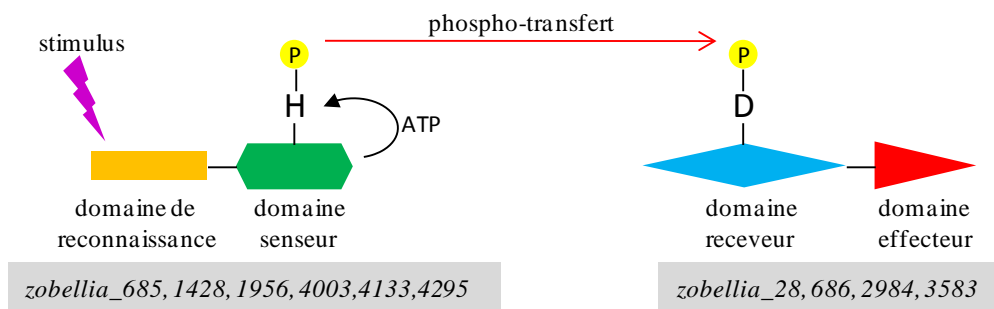


Figure 5-17 : Schéma des gènes sur-exprimés en présence d'alginate et de laminarine <sup>204</sup>  
 Le nombre de gènes inclus dans chaque catégorie est donné.

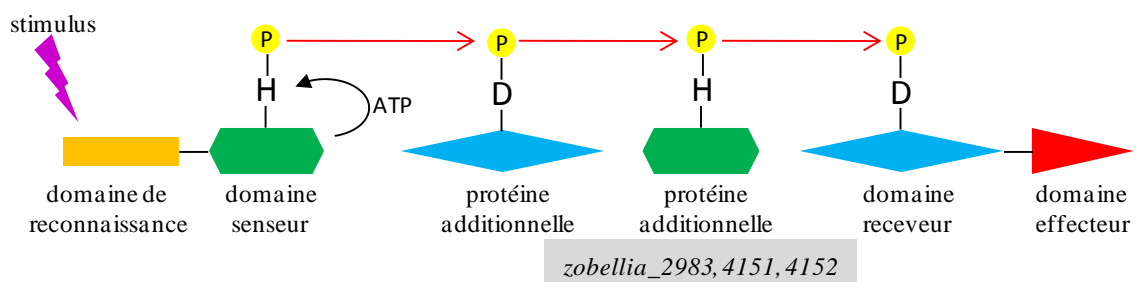
### III-5-b Perception de l'environnement et régulation

35 gènes sur-exprimés en présence d'alginate sont impliqués dans des processus de perception et de transduction du signal, soit 15% du nombre total (Figure 5-17). On constate qu'un grand nombre de ces gènes code des partenaires de systèmes à deux composants ou de transfert de phosphate ("*phosphorelay*"). Chez les bactéries, les signaux extracellulaires sont transmis majoritairement par des systèmes à deux composants (Cashin *et al.*, 2006). Ces systèmes comprennent un premier partenaire protéique composé d'un domaine de reconnaissance du signal et d'un domaine kinase senseur, et un deuxième partenaire composé d'un domaine receveur et d'un domaine effecteur (Hoch 2000; Stock *et al.*, 2000). L'organisation et le fonctionnement archétypique d'un tel système sont décrits dans la Figure 5-18a.

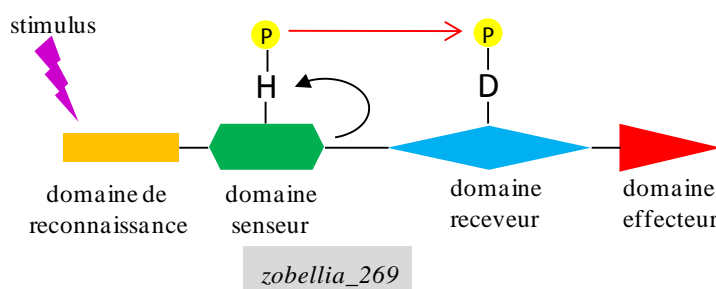
#### a) Système à deux composants



#### b) Système "phospho-relais"



#### c) Système à un composant



**Figure 5-18 : Schéma des systèmes bactériens de transduction du signal impliquant des transferts de phosphate**

Les gènes correspondants de *Z. galactanivorans* surexprimés en présence d'alginate sont reportés dans les encadrés gris. H : histidine. D : aspartate.

En réponse à un stimulus perçu par le domaine de reconnaissance, le domaine senseur du premier partenaire s'autophosphoryle sur une histidine. Le groupement phosphate est ensuite transféré sur un aspartate du domaine receveur du deuxième partenaire. Cette phosphorylation provoque une activation du domaine effecteur, qui peut se lier sur des séquences promotrices de l'ADN et induire la transcription de gènes cibles. Des systèmes alternatifs basé sur le même principe mais impliquant des partenaires supplémentaires existent (Figure 5-18b), et sont nommés systèmes "phospho-relais". Dans certains cas, les deux partenaires (senseur et receveur) ont fusionné en une seule protéine : on parle dans ce cas de système à un composant (Figure 5-18c). Chez *Z. galactanivorans*, 14 gènes codant des protéines impliquées dans de tels systèmes de transduction du signal sont surexprimés en présence d'alginate. Les partenaires identifiés sont donc probablement impliqués dans les phénomènes de régulation dépendants de la source de carbone.

Un autre senseur potentiel est surexprimé en présence d'alginate (Figure 5-17). Il s'agit du gène *zobellia\_1595*, qui code une protéine cytoplasmique adoptant un repliement de la famille PAS. Ce repliement PAS tire son nom des protéines PER, ARNT et SIM, où il a été mis en évidence pour la première fois. Il est fréquemment retrouvé dans des protéines de signalisation, notamment des senseurs de systèmes à deux composants (Szurmant *et al.*, 2007), participe à la liaison de petites molécules et à des interactions protéines-protéines. Les domaines PAS perçoivent des variations de facteurs environnementaux tels que la lumière, l'oxygène, le potentiel redox ou de petits ligands qui traversent la membrane et affectent le métabolisme cellulaire (Taylor and Zhulin 1999).

Les résultats montrent également que l'expression de 8 gènes codant des facteurs de transcription augmente en présence d'alginate (Figure 5-17). Les différentes familles concernées sont présentées dans le Tableau 5-7.

Famille	n° PFAM	Action	Exemples de voies régulées identifiées	gènes induits avec alginate
AraC	PF00165	Activateurs	Dégradation de molécules organiques, pathogénicité, réponse au stress, etc... <sup>a</sup>	zobellia_920 zobellia_4009 zobellia_4664
AsnC-Lrp	PF01037	Activateurs Répresseurs	Métabolisme des acides aminés et de l'azote <sup>b</sup>	zobellia_4436
BlaI-MecI-CopY	PF03965	Répresseurs	Résistance aux pénicillines, homéostasie du cuivre <sup>c</sup>	zobellia_2269
LacI	PF00356	Répresseurs	Dégradation de sources de carbone <sup>d</sup>	zobellia_2224
PadR	PF03551	Répresseurs	Réponse au stress par les acides phénoliques <sup>e</sup>	zobellia_1768
TetR	PF00440	Activateurs Répresseurs	Résistance aux drogues, différentes voies cataboliques, biosynthèse d'antibiotiques, stress osmotique, pathogénicité, etc... <sup>f</sup>	zobellia_701

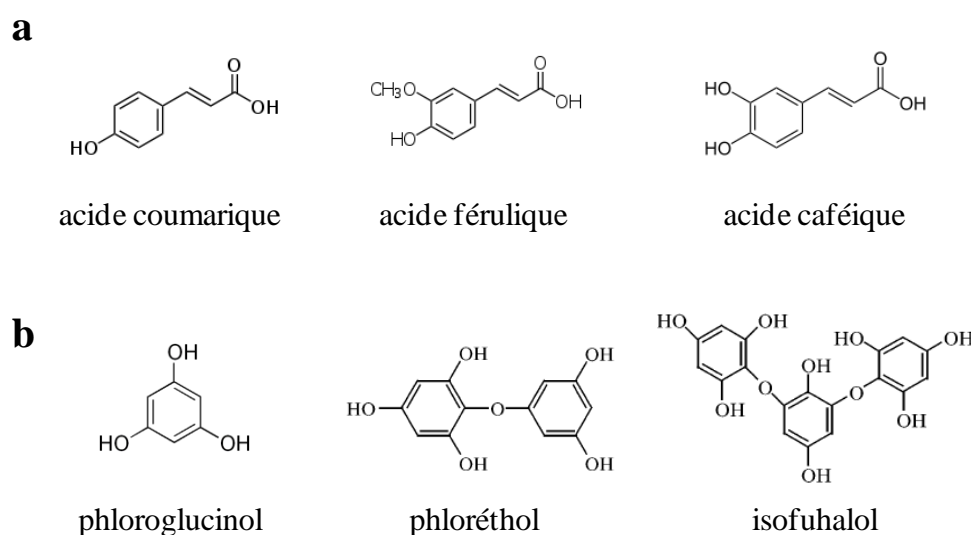
**Tableau 5-7 : Familles de régulateurs transcriptionnels dont l'expression est induite en présence d'alginate**

a : Gallegos *et al.* (1997); b : Friedberg *et al.* (2001); c : McKinney *et al.* (2001); d : Weickert and Adhya (1992); e : Gury *et al.* (2004); f : Ramos *et al.* (2005). Tableau réalisé avec l'aide de la base de données accessible en ligne : <http://centre.bioinformatics.zj.cn/proTF/index.php> (Bai *et al.*, 2010)

Parmi les 6 familles identifiées, AraC et LacI sont bien connues pour réguler des mécanismes de dégradation de sucres. La famille AraC regroupe un grand nombre d'activateurs de la transcription (Gallegos *et al.*, 1997), notamment impliqués dans la dégradation de l'arabinose (Greenblatt and Schleif 1971), du rhamnose (Tobin and Schleif 1990), du mélibiose (Caswell *et al.*, 1992) et du cellobiose (Plumbridge and Pellegrini 2004). Le membre le plus connu de la famille LacI est le répresseur de l'opéron *lac* chez *E. coli* (Gilbert and Muller-Hill 1966), mais d'autres membres sont impliqués par exemple dans l'utilisation du galactose (Geanakopoulos and Adhya 1997), du maltose (Reidl *et al.*, 1989) et du fructose (Saier and Ramseier 1996). La fonction de ces régulateurs (familles AraC et LacI) dépend de leur capacité à fixer un effecteur qui, dans les cas présentés ici, est le monosaccharide à utiliser. Cette interaction induit un changement de conformation de la protéine, modifiant ses propriétés de liaison aux promoteurs des gènes du régulon. Les régulateurs des familles LacI et AraC mis en évidence lors de la réponse à l'alginate chez *Z. galactanivorans* pourraient donc avoir pour effecteur un produit de dégradation du polysaccharide.

Les familles de régulateurs TetR et BlaI sont connues pour intervenir dans le contrôle de voies métaboliques très diverses (Tableau 5-7). Il est donc plus difficile d'interpréter en quoi des membres de ces familles pourraient jouer un rôle dans la réponse de *Z. galactanivorans* à l'alginate. La surexpression d'un membre de la famille AsnC-Lrp, connue pour le contrôle du métabolisme des acides aminés, est sans doute à relier avec la répression de la synthèse de certains d'entre eux en présence d'alginate (cf. section suivante).

Le dernier facteur de transcription dont l'expression est corrélée avec la présence d'alginate appartient à la famille PadR. Les membres caractérisés de cette famille de régulateurs interviennent dans la réponse au stress par les composés phénoliques (Barthelmebs *et al.*, 2000). Les acides phénoliques, ou acides hydroxycinnamiques substitués, sont des composés abondants chez les végétaux. Les structures chimiques de certains d'entre eux sont données sur la Figure 5-19a.



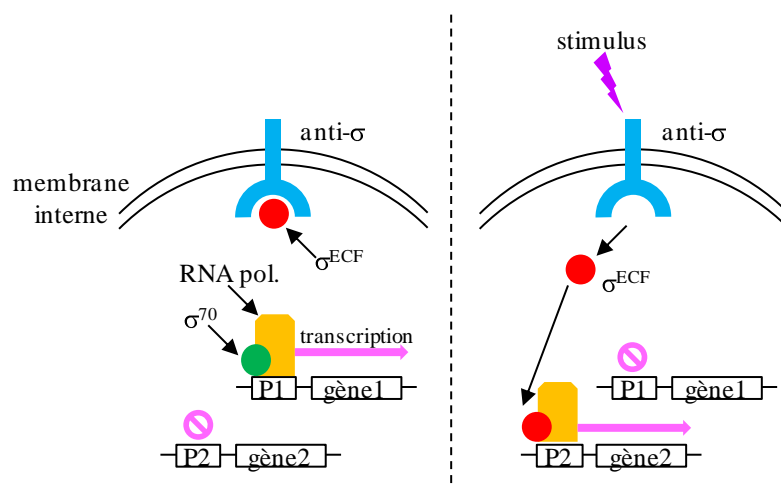
**Figure 5-19 : Structure chimique de quelques composés phénoliques de végétaux**

a. Composés phénoliques de plantes terrestres. b. Composés phénoliques d'algues brunes.

Ces composés phénoliques ont une activité anti-microbienne démontrée, et interviennent dans les réponses de défense des plantes face aux pathogènes (Nicholson and Hammerschmidt 1992). Certains micro-organismes peuvent adapter leur métabolisme suite à la perception de ces composés, via l'action du facteur de transcription PadR. Les régulations qui s'en suivent conduisent notamment à l'expression de l'enzyme PAD (Phenolic Acid Decarboxylase) qui convertit ces substrats en dérivés moins toxiques (Barthelmebs *et al.*, 2000; Gury *et al.*, 2004; Tran *et al.*, 2008). De la même manière, les algues brunes produisent des composés phénoliques dérivés du phloroglucinol (Figure 5-19b) qui jouent un rôle lors

des réactions de défense face à une agression microbienne (Amsler and Fairhead 2006). Il est donc possible que des systèmes de perception de ces composés phénoliques aient évolué chez les bactéries marines capables de dégrader les parois d'algues, et les membres de la famille PadR seraient de bons candidats. La présence de laminarine provoque la surexpression d'un second facteur de transcription de la famille PadR chez *Z. galactanivorans* (Figure 5-17). Je propose l'hypothèse suivante quant à leur rôle dans cette réponse. Suite à la perception de la présence des polysaccharides dans son environnement, *Z. galactanivorans* activerait non seulement des systèmes de dégradation efficaces, mais se préparerait également à répondre aux composés de défense émis par les algues. Cette hypothèse est réminiscente avec l'induction du gène *zobellia\_2612* qui code un transporteur Nramp et dont l'implication dans la résistance aux espèces réactives de l'oxygène émises par les algues a déjà été discutée (cf. Chapitre 4).

Finalement, dans la catégorie relative à la perception et la transduction du signal, deux gènes codant des facteurs sigma de type ECF et un gène codant un facteur anti-sigma sont surexprimés en présence d'alginate. Les facteurs sigma bactériens sont des composants majeurs du complexe ARN polymérase, dont ils déterminent la spécificité de fixation à un promoteur. La plupart des bactéries contiennent un facteur sigma primaire ( $\sigma^{70}$ ) pour la transcription de gènes constitutifs (dits "de ménage"), et différents facteurs sigma alternatifs activés par des conditions spécifiques (Helmann 2002). C'est le cas des facteurs sigma de type ECF (Extra-Cytoplasmique Function). Le mécanisme de régulation de l'expression fait le plus souvent intervenir une protéine transmembranaire nommée anti- $\sigma$ , qui séquestre le facteur  $\sigma^{ECF}$  en l'absence de stimulus environnemental (Figure 5-20). Suite à la perception du signal, le facteur anti- $\sigma$  libère  $\sigma^{ECF}$ , qui peut alors interagir avec le complexe ARN polymérase et modifier la transcription des gènes cibles (Helmann 1999).



**Figure 5-20 : Principe de la régulation par un système sigma/anti-sigma**



### III-5-c Les protéines de transport

En réponse à l'alginate, et suite aux différents phénomènes de régulation décrits précédemment, *Z. galactanivorans* exprime un grand nombre de protéines impliquées dans le transport de molécules organiques (Figure 5-17). Douze gènes ont été classés dans la catégorie "Transport", soit 5% du nombre total de gènes surexprimés.

Parmi ceux-ci, on retrouve des gènes homologues à *susC* et *susD*, qui appartiennent au locus d'utilisation de l'amidon (Starch Utilization System) de *Bacteroides thetaiotaomicron* (Anderson and Salyers 1989). Comme présenté dans le Chapitre 2, les protéines membranaires SusC et SusD interviennent dans la fixation du polymère d'amidon à la surface de la bactérie, qui est alors dégradé par une amylase membranaire (Martens *et al.*, 2009). Les produits de dégradation pénètrent ensuite dans le périplasme via le canal formé par SusC qui adopte un repliement en tonneau beta ( $\beta$ -barrel). Chez les membres du phylum des *Bacteroidetes*, des gènes *susC* et *susD-like* ont été trouvés dans des locus d'utilisation d'autres polysaccharides que l'amidon (cf. Chapitre 2). Différentes études ont montré que l'expression de ces gènes est induite par la présence du substrat (Spence *et al.*, 2006; Dodd *et al.*, 2010). Par exemple, le xylane induit l'expression de 10 gènes *susC*- ou *susD-like* chez *Prevotella bryantii* (Dodd *et al.*, 2010). Les résultats de l'étude présentée ici vont dans le même sens. Dans le Chapitre 4, un locus d'utilisation de l'alginate comprenant un système similaire à SusCD avait déjà été mis en évidence. Les deux gènes correspondants (*zobellia\_2620* et *2621*) sont bien retrouvés surexprimés dans l'analyse transcriptomique à l'échelle du génome, tout comme un gène codant une perméase à sucres (*zobellia\_2616*). Les résultats suggèrent par ailleurs que d'autres protéines homologues à SusC et SusD non prises en considération auparavant pourraient participer à la dégradation de l'alginate.

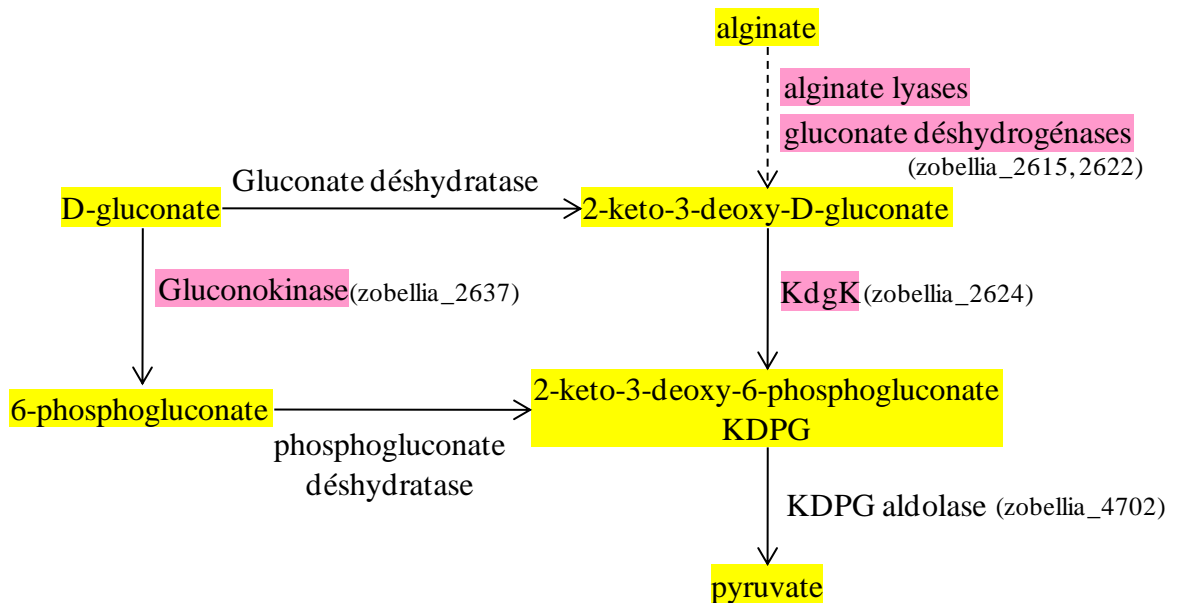
L'expression de trois gènes codant des composants d'ABC-transporteurs est également induite en présence d'alginate. Les ABC-transporteurs (pour ATP Binding Cassette) utilisent l'énergie libérée par l'hydrolyse d'ATP pour transporter de manière active différents types de composés. Ils sont notamment connus pour l'import de mono- et d'oligosaccharides résultants de la dégradation enzymatique de polymères (Schneider 2001; Eitinger *et al.*, 2011). Ainsi, des systèmes ABC prennent notamment en charge les produits de dégradation de la pectine (Hugouvieux-Cotte-Pattat *et al.*, 2001), de la chitine (Saito and Schrempf 2004) ou de la cellulose (Koning *et al.*, 2001). Chez *Sphingomonas spp.*, un système ABC transporteur peut même importer un polymère d'alginate intact (Murata *et al.*, 2008). Il est donc possible que les

3 composants d'ABC-transporteurs répondant à la présence d'alginate chez *Z. galactanivorans* soient capables de transférer des oligo-alginates du périplasme vers le cytoplasme.

### III-5-d Le catabolisme des sucres

La présence d'alginate induit l'expression de gènes codant pour 13 protéines impliquées dans le catabolisme des sucres (Figure 5-17). Il est intéressant de remarquer qu'aucun de ces gènes n'est régulé de façon commune par l'alginate et la laminarine. Ceci suggère que les voies cataboliques régulées sont bien spécifiques de la source de carbone mise à disposition.

Parmi ces gènes, on retrouve ceux qui codent les 6 alginate lyases dont il a déjà été question dans le Chapitre 4. D'autre part, 3 gènes surexprimés codent des enzymes du métabolisme du gluconate, qui est un intermédiaire important dans la voie d'Entner-Doudoroff menant à la formation de pyruvate (Entner and Doudoroff 1952). Il a par ailleurs été montré que l'assimilation des monosaccharides provenant de la dégradation de l'alginate aboutit à la formation d'un intermédiaire de cette voie, le 2-céto-3-désoxy-D-gluconate (KDPG) (Preiss and Ashwell 1962b). Nos résultats montrent que ces mêmes voies métaboliques sont impliquées dans la dégradation de l'alginate chez *Z. galactanivorans* (Figure 5-21).



**Figure 5-21 : Place des dérivés du gluconate dans le catabolisme de l'alginate**

Les enzymes dont les gènes sont surexprimés en présence d'alginate chez *Z. galactanivorans* sont surlignées en rouge.

Quatre autres gènes codant des CAZymes sont surexprimés en présence d'alginate. Le gène *zobellia\_3464* code une  $\alpha$ -L-fucosidase de la famille GH29. Ce gène appartient à un cluster codant notamment trois sulfatases, un système SusCD-like, une fuconate déshydratase et un importeur de fucose. On peut penser que cette région génomique est dédiée à la dégradation de fucane, un polymère de fucose sulfaté présent dans la paroi des algues brunes (Li *et al.*, 2008). Les fucosidases catalysent la libération d'une unité fucose à partir de l'extrémité non-réductrice du polysaccharide (Tanaka and Sorai 1970; Berteau *et al.*, 2002). La surexpression d'une telle enzyme en présence d'alginate purifié suggère l'existence de régulations croisées. La perception d'un polysaccharide donné d'algues brunes par *Z. galactanivorans* activerait des voies cataboliques pour d'autres composés abondants chez ces mêmes algues. Il est intéressant de remarquer que l'expression de deux autres gènes de fucosidases est induite en présence de laminarine, et que la même explication peut être envisagée.

Il est plus difficile d'avancer des hypothèses sur l'induction de l'expression des trois autres gènes classés dans la catégorie "Métabolisme des sucres". Le gène *zobellia\_3446* code une Carbohydate Esterase de la famille CE10. Les membres de cette famille catalysent la dé O- ou N-acétylation de sucres substitués (Cantarel *et al.*, 2009). De manière intéressante, ce gène est situé juste en aval de *uglA1* (*zobellia\_3445*), dont l'expression avait été montrée comme induite en présence d'alginate lors de l'étude en RT-PCR quantitative (Chapitre 4). Le gène *zobellia\_3770* est annoté comme pectate lyase de la famille PL1. Comme cela a déjà été mentionné, *Z. galactanivorans* ne semble pas capable de dégrader la pectine, ce qui suggère que la fonction de ce gène pourrait être différente. Enfin, *zobellia\_4267* code AgaC, une  $\beta$ -agarase de la famille GH16 (Jam *et al.*, 2005). Néanmoins, cette agarase est très divergente (Hehemann *et al.*, 2010a) et n'est pas retenue sur une colonne de sépharose (Jam *et al.*, 2005), suggérant que l'agar ne soit pas son substrat naturel.

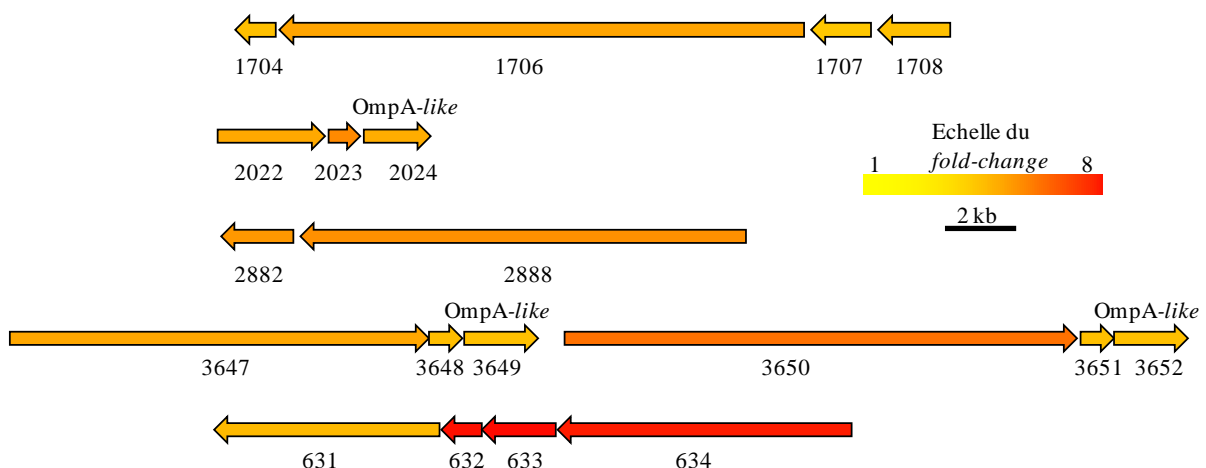
### III-5-e La mobilité par glissement : lien avec la sécrétion et la dégradation

Deux gènes potentiellement impliqués dans la mobilité par glissement répondent à la présence d'alginate et de laminarine dans le milieu de culture (Figure 5-17). Les ORF *zobellia\_1773* et *zobellia\_1806* codent des homologues des protéines de *Flavobacterium johnsoniae* SprE (similitude = 61%) et GldH (similitude = 63%), respectivement. Chez cette

espèce, des études par mutagenèse ont montré que GldH et SprE sont requises pour la mobilité par glissement (MacBride and Samarasinghe, non publié; McBride *et al.*, 2003). Par ailleurs, des mutants dans des protéines nécessaires au glissement (dont GldH) ne peuvent plus dégrader la chitine (McBride *et al.*, 2003; Braun and McBride 2005; Nelson and McBride 2006). Il a été récemment démontré que les protéines des familles Gld et Spr sont des composants d'un système original de sécrétion de protéines PorSS chez *Porphyromonas gingivalis*, conservés au sein du phylum des *Bacteroidetes* (Sato *et al.*, 2010). Chez *F. johnsoniae*, il a été proposé que ce système PorSS soit impliqué à la fois dans l'assemblage d'un complexe protéique nécessaire au glissement et dans la sécrétion de chitinases extracellulaires (Rhodes *et al.*, 2010). L'induction de zobellia\_1773 et zobellia\_1806 en présence d'alginate suggère qu'ils pourraient coder des protéines d'un complexe analogue au système de sécrétion PorSS.

### III-5-f Gènes de fonction inconnue : un réservoir de protéines à étudier

La moitié des gènes dont l'expression est induite en présence d'alginate sont de fonction inconnue. Cela signifie qu'aucun homologue caractérisé n'a pu être trouvé dans les banques de données. Ils représentent donc une source potentielle de nouvelles familles de protéines impliquées dans la dégradation de polysaccharides. Certains de ces gènes sont regroupés en clusters, dont quelques exemples sont donnés sur la Figure 5-22.



**Figure 5-22 : Clusters de gènes de fonction inconnue surexprimés en présence d'alginate**

Le taux de variation (*fold-change*) entre alginate et glucose est représenté pour chaque gène selon une échelle de couleur.

De manière préliminaire, des similitudes de séquences ont été recherchées au sein des protéines codées par ces gènes de fonction inconnue. Par alignement de séquences, deux familles de gènes paralogues ont ainsi pu être mises en évidence (Figure 5-23, page suivante). Les cinq protéines de la famille 1 s'alignent bien sur toute la longueur de leur séquence. Dans la famille 2, la similitude de séquence ne concerne que les régions C-terminales des 10 protéines.

Une recherche par BLAST a montré que les protéines de la famille 1 semblent uniques au phylum des *Bacteroidetes*, et particulièrement représentées chez les *Flavobacteriia*. Après interrogation de la banque de données CDD au NCBI (Marchler-Bauer *et al.*, 2011), on détecte un domaine conservé de fonction inconnue de la superfamille DUF 3308 dans les cinq protéines mises en évidence ici.

Mis à part pour *zobellia\_2761*, l'expression des protéines classées dans les familles 1 et 2 est induite à la fois en présence d'alginate ou de laminarine. Même si leur fonction reste mystérieuse, ces protéines paraissent donc être des acteurs clés de la dégradation de polysaccharides par *Z. galactanivorans*. Elles sont par conséquent des cibles de choix pour des études fonctionnelles ultérieures.

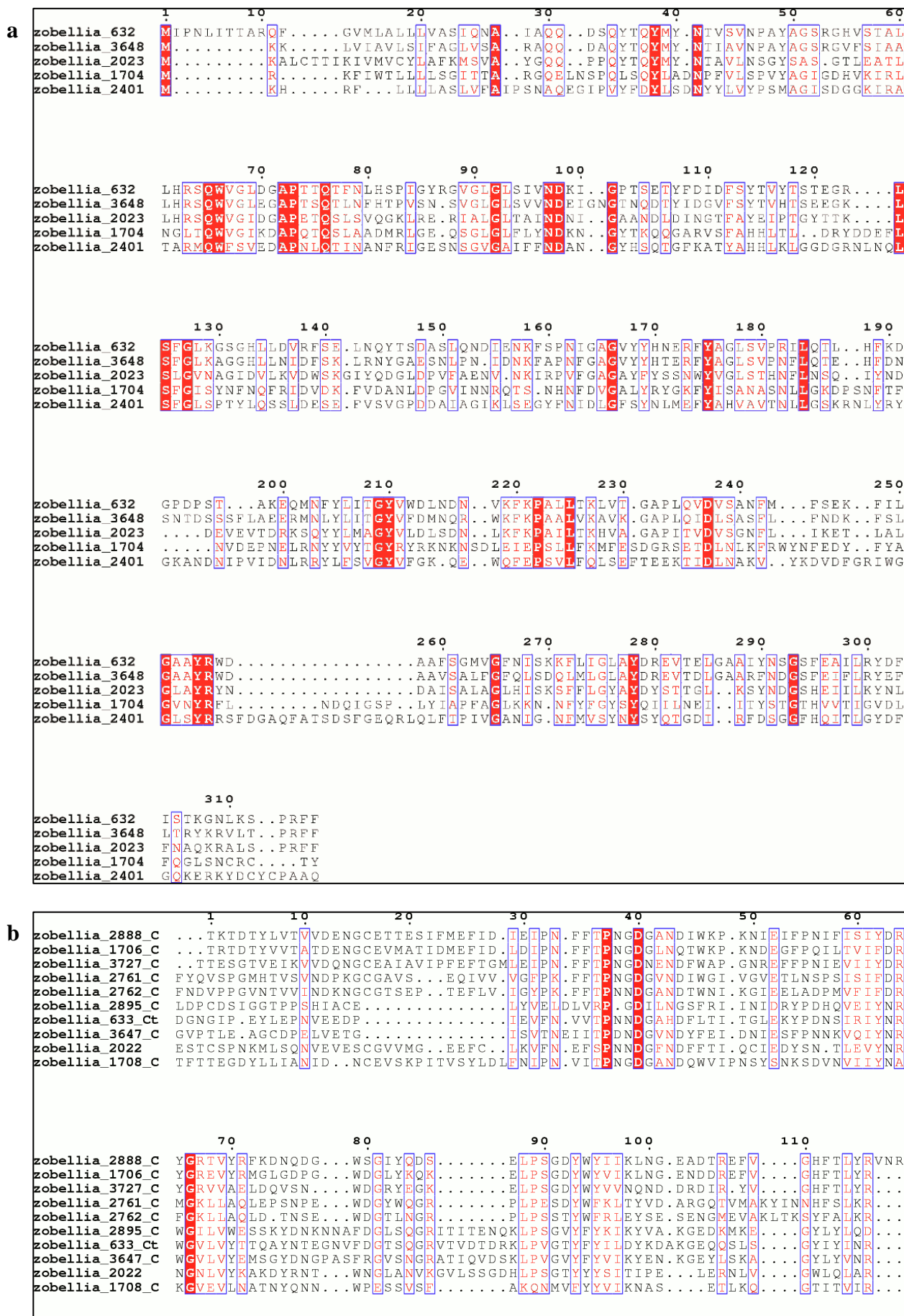


Figure 5-23 : Alignement de séquences de protéines de fonction inconnue surexprimées en présence d'alginate  
a. Famille 1. b. Famille 2, uniquement les régions C-terminales des protéines.

### III-6 Les gènes sous-exprimés en présence d'alginate

En présence d'alginate, l'expression de 369 gènes est réprimée, dont 69 sont de fonction inconnue (Annexe 2). Comme je l'ai déjà mentionné, la faible proportion de gènes de fonction inconnue indique que les voies sous-exprimées en présence d'alginate comme seule source de carbone (par rapport au glucose) sont relativement bien caractérisées. Dans le cadre de cette thèse, la finalité de l'analyse du transcriptome était de mieux définir les voies conduisant à la dégradation de l'alginate. En conséquence, je ne discuterai pas dans le détail toutes les voies sous-exprimées en présence d'alginate par rapport au glucose, comme j'ai pu le faire pour les gènes sur-exprimés.

Une recherche des ontologies (termes GO) significativement sur-représentées dans le lot de gènes réprimés en présence d'alginate a été effectuée (Tableau 5-8, page suivante). Elle montre que lors de l'utilisation d'alginate comme source de carbone, *Z. galactanivorans* réprime un certain nombre de voies du métabolisme basal, notamment le métabolisme énergétique. Ainsi, les voies de la glycolyse, des pentoses phosphate et le cycle de Krebs sont moins actives en présence d'alginate que de glucose comme seule source de carbone. De même, l'expression d'un nombre important de gènes codant des enzymes impliquées dans la synthèse d'acides aminés est réprimée. Ceci tend à montrer une réallocation des ressources vers la production d'acteurs moléculaires permettant une dégradation la plus efficace possible des polysaccharides présents dans l'environnement, au détriment d'autres voies. Il est intéressant de constater que le même phénomène a été observé chez les bactéries utilisant la laminarine (Tableau 5-8). Parmi les 369 gènes sous-exprimés en présence d'alginate, on en compte 6 (soit 1,6%) codant des protéines impliquées dans la régulation de la transcription (Systèmes à deux composants, régulateur *GntR-like*, etc.). En comparaison, 11,9% des gènes surexprimés en présence d'alginate par rapport au glucose ont été classés dans la catégorie Perception/Transduction (Figure 5-17). Ceci pourrait indiquer que la majorité des régulations liées à la présence de polysaccharide dans le milieu est contrôlée via des inductions. L'absence du substrat résulterait plutôt en une absence d'induction qu'en une répression active. Cette absence de répression catabolique avait déjà été remarquée dans le Chapitre 4.

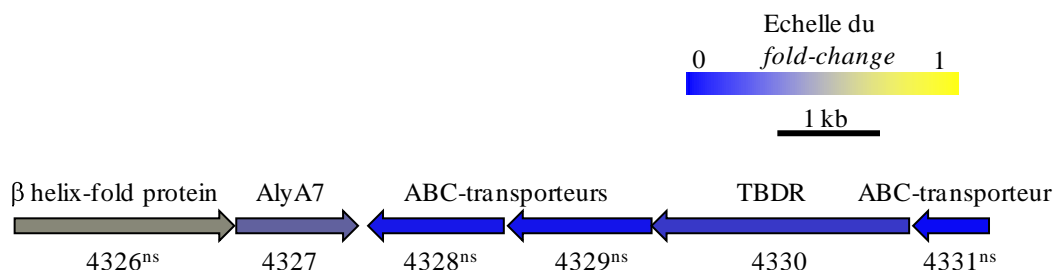
GO ID	Fonction	Alginate	Laminarine
<b>Métabolisme énergétique</b>			
6096	glycolysis	X	
6098	pentose-phosphate shunt	X	
6099	tricarboxylic acid cycle	X	
6183	GTP biosynthetic process		X
15986	ATP synthesis coupled proton transport	X	X
19647	formaldehyde assimilation via ribulose monophosphate cycle	X	X
<b>Métabolisme des acides aminés</b>			
162	tryptophan biosynthetic process	X	X
6525	arginine metabolic process	X	X
6544	glycine metabolic process		X
6561	proline biosynthetic process	X	X
6564	L-serine biosynthetic process		X
9067	aspartate family amino acid biosynthetic process	X	X
9082	branched chain family amino acid biosynthetic process	X	X
9095	aromatic amino acid family biosynthetic process, prephenate pathway		X
<b>Traduction</b>			
6418	tRNA aminoacylation for protein translation	X	X
<b>Métabolisme des nucléotides</b>			
6221	pyrimidine nucleotide biosynthetic process		X

**Tableau 5-8 : Détection de termes GO sur-représentés parmi les gènes sous-exprimés en présence d'alginate ou de laminarine**

Seules les ontologies des nœuds terminaux du réseau obtenu sont données.



L'analyse par puces à ADN confirme la répression du gène codant l'alginate lyase AlyA7 (zobellia\_4327) qui avait été montrée par RT-PCR quantitative (cf. Chapitre 4). Par ailleurs, elle montre qu'*alyA7* fait partie d'un cluster de gènes sous-exprimés en présence d'alginate (Figure 5-24).



**Figure 5-24 : Région génomique autour d'*alyA7*, sous-exprimée en présence d'alginate**

Le taux de variation (*fold-change*) entre alginate et glucose est représenté pour chaque gène selon une échelle de couleur.

ns : gène dont le *fold-change* est inférieur à 1, mais qui ne remplit pas le critère de q-value < 0,05.

Par ailleurs, il est intéressant de remarquer que l'expression d'un locus d'utilisation de l'amidon (zobellia\_369 à zobellia\_372) est nettement plus faible en présence d'alginate que de glucose. Ce locus est homologue au Starch Utilization System de *B. thetaiotaomicron* présenté dans le Chapitre 2. De même, l'expression d'une oligo-1,6-glucosidase de la famille GH13 (zobellia\_530) est réprimée en absence de glucose. Cette enzyme catalyse l'hydrolyse des liaisons glycosidiques dans les oligosaccharides résultant de la dégradation de l'amidon ou du glycogène. Là encore, on peut supposer que la faible expression de ces gènes en présence d'alginate évite à *Z. galactanivorans* de synthétiser des enzymes de dégradation de glucanes lorsqu'ils sont absents.

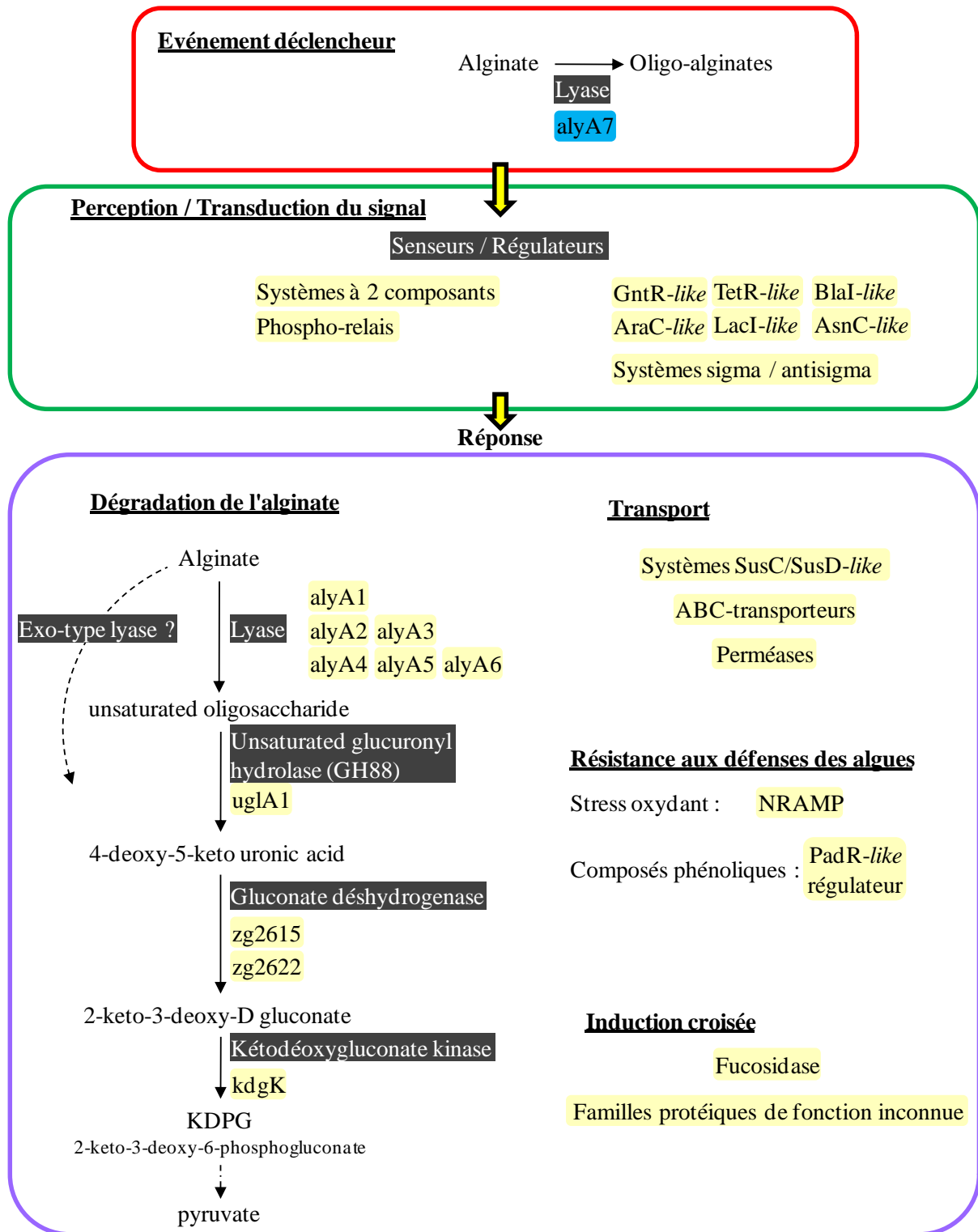
Enfin, la répression de l'expression du gène zobellia\_155 en présence d'alginate est particulièrement intéressante. En effet, ce gène code une C-alpha-formylglycine-generating protein. Cette enzyme convertit les sulfatases nouvellement synthétisées sous forme inactive en leur forme active, en modifiant une sérine ou une cystéine du site actif en C $\alpha$ -formylglycine (FGly). Le génome de *Z. galactanivorans* contient une seule copie de ce gène. En conséquence, l'enzyme qu'il code est cruciale pour l'activation des sulfatases. Comme je l'ai mentionné dans l'introduction de cette thèse, un grand nombre de polysaccharides marins est sulfaté, et leur dégradation nécessite le retrait des groupements sulfate. *Z. galactanivorans* possède 72 sulfatases qui pourraient y contribuer. La dégradation de l'alginate, un

polysaccharide non-sulfaté, ne nécessite pas l'action de telles sulfatases. La répression du gène codant l'enzyme qui active les sulfatases apparaît comme une stratégie d'adaptation à la source de carbone. Il peut apparaître plus efficace pour la bactérie de réprimer simultanément l'action de toutes les sulfatases par l'absence de la modification post-traductionnelle qui crée la formylglycine. On peut imaginer que dès qu'un polysaccharide sulfaté est détecté dans le milieu, le gène *zobellia\_155* est de nouveau transcrit, et l'enzyme qu'il code va activer les sulfatases déjà présentes. Cette stratégie semble ainsi permettre une régulation plus rapide du métabolisme du catabolisme de substrats sulfatés qu'une synthèse *de novo* d'un grand nombre de sulfatases.

#### IV- Vers un nouveau modèle de régulation

L'analyse globale par puces à ADN du transcriptome de *Z. galactanivorans* en présence d'alginate a en grande partie confirmé les résultats obtenus en RT-PCR quantitative sur une vingtaine de gènes. Par ailleurs, cette approche à l'échelle du génome entier a permis de mettre en évidence de nouveaux acteurs de la perception, de la régulation et de la dégradation impliqués dans les réponses à l'alginate. Dans le Chapitre 4, un premier modèle de régulation avait été proposé. En y intégrant les nouvelles données, je propose un modèle plus complet, qui pourra servir de base pour une validation fonctionnelle ultérieure (Figure 5-25, page suivante).

L'alginate lyase AlyA7, exprimée en absence d'alginate, pourrait jouer le rôle "d'enzyme-sentinelle" pour la cellule et réaliser les premiers cycles de dégradation produisant les oligo-alginates. Ces oligosaccharides pourraient alors être perçus comme signal par la bactérie et déclencher une réponse spécifique. Cette première phase d'amorçage de la réponse nécessite l'expression minimale de systèmes de perception, permettant de déclencher la réponse proprement dite. Une des premières phases de la réponse est l'induction de systèmes de perception et de régulation de l'expression génétique qui permettent des changements majeurs du transcriptome, illustrés par la forte proportion de gènes régulés en présence d'alginate (13% du génome). Les cellules réorientent leur métabolisme pour dégrader activement le polysaccharide, en activant l'expression d'enzymes des différentes étapes du catabolisme aboutissant à la production de KDPG, un intermédiaire du cycle d'Entner-Doudoroff. En parallèle, des systèmes de transports sont exprimés et probablement responsables de l'import des produits de dégradation dans le cytoplasme. Parmi ces systèmes de transport, on retrouve un grand nombre de protéines SusC ou SusD-like, dont on sait qu'elles permettent aussi de fixer les substrats à la cellule. Ce phénomène évite sans doute la perte des oligo-alginates générés dans le milieu extracellulaire. Il semble que simultanément à cette réponse, *Z. galactanivorans* se prépare à résister activement aux réactions de défense des algues, comme la production d'espèces réactives de l'oxygène ou la libération de composés phénoliques. Ceci pourrait résulter d'une coévolution façonnée par les interactions entre algues et bactéries. Enfin, la perception d'alginate semble également déclencher des voies métaboliques liées à d'autres polysaccharides d'algues brunes, tels que la laminarine ou le fucane. Ces "inductions croisées" permettraient à *Z. galactanivorans* d'utiliser au maximum les sources de carbone disponibles lors de la dégradation d'une algue brune.



**Figure 5-25 : Modèle de la réponse à l'alginate chez *Z. galactanivorans***

Modèle basé sur les résultats présentés dans ce chapitre et le Chapitre 4. Les protéines dont l'expression est régulée aux différentes étapes sont notées dans les cadres jaunes (surexpression en présence d'alginate par rapport au glucose) et bleu (sous-expression).



## Chapitre 6

# Clonage et surexpression à moyen débit de protéines impliquées dans la dégradation de l'alginate

---

## I- Introduction

### I-1 Contexte

L'annotation du génome de *Zobellia galactanivorans* a permis de prédire les fonctions de 61% de ses 4738 gènes. Cette annotation manuelle s'est basée sur l'existence de similitudes de séquences des gènes ou modules avec des protéines homologues de fonction biologique et/ou biochimique caractérisée. Cette approche, bien que beaucoup plus juste que l'annotation automatique, n'en reste cependant que prédictive. Une validation post-génomique est donc indispensable afin de confirmer et/ou compléter l'annotation du génome de *Z. galactanivorans*. D'autre part, l'avènement des techniques de séquençage à haut débit a permis de réduire drastiquement les coûts et les délais, et abouti à un dépôt massif de données. A titre d'exemple, le site NCBI recense à ce jour 1457 génomes microbiens complets, et 3918 en cours (<http://www.ncbi.nlm.nih.gov/gate1.inist.fr/genomes/lproks.cgi>, 01/03/11). Plus le nombre de gènes de fonction validée augmentera dans les bases de données, plus l'annotation de nouvelles séquences pourra être précise. Dans ce cadre, le clonage et l'expression en système hétérologue des gènes d'intérêt permettent d'obtenir des quantités de protéines relativement élevées, facilitant les études biochimiques et structurales. Des projets de production de protéine à haut débit et de génomique structurale ont ainsi vu le jour pour tenter de combler l'écart grandissant entre le nombre de séquences protéiques connues et celles pour lesquelles les fonctions et/ou structures sont caractérisées (Matte *et al.*, 2003; Chruszcz *et al.*, 2010).

L'étude approfondie du système alginolytique de *Z. galactanivorans* a mis en évidence un certain nombre de gènes potentiellement impliqués dans la dégradation de l'alginate (cf. Chapitres 4 et 5). Le modèle proposé, basé sur des données génomiques et transcriptomiques, nécessite cependant une validation fonctionnelle pour les différentes protéines. En effet, ces techniques apportent des informations sur les voies métaboliques et les réseaux de régulation auxquels les gènes appartiennent, mais non sur la fonction biochimique précise des différents acteurs. Dans ce contexte, j'ai entrepris une stratégie de clonage et d'expression hétérologue des différentes cibles identifiées, en vue de leur caractérisation biochimique et structurale. Les résultats, combinés à ceux de génomique et de transcriptomique, devraient apporter un

éclairage sur les fonctions biologiques de chacun des gènes, et sur leur implication dans le métabolisme de l'alginate chez *Z. galactanivorans*.

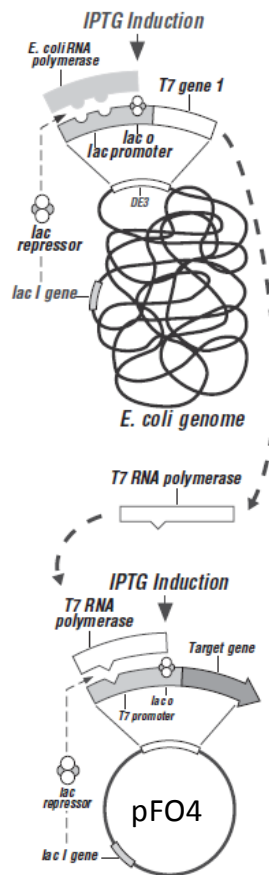
## I-2 Philosophie de la méthode de surexpression à moyen débit

En s'inspirant des techniques utilisées par les projets des grands centres de génomique structurale, une stratégie de clonage et d'expression de protéines à moyen débit a été développée au sein du laboratoire (Dabin 2008; Groisillier *et al.*, 2010). Elle consiste à traiter en parallèle un grand nombre de cibles d'intérêt. Afin de maximiser les chances d'obtenir des protéines exprimées et solubles, deux types de vecteur sont utilisés.

Le vecteur pGEX-4-T1 permet le clonage du gène d'intérêt en aval d'une séquence une glutathion-S-transférase. La protéine recombinante peut ainsi être purifiée par chromatographie d'affinité au glutathion. Le vecteur pGEX-4-T1 contient le promoteur *tac* en amont du gène d'intérêt. Ce promoteur hybride fort est composé de la région -35 du promoteur *trp* et de la région -10 du promoteur *lacUV5* (Deboer *et al.*, 1983). L'expression des cibles recombinantes est réprimée par la protéine LacI, et peut être induite par ajout d'IPTG ou de lactose. pGEX-4-T1 peut donc être transformé dans les cellules d'expression *E. coli* BL21 classiques.

Le deuxième vecteur utilisé dans cette stratégie à moyen débit est un dérivé du plasmide pET-15b nommé pFO4 (don de M. Cygler). Le gène d'intérêt est ici cloné en aval d'une séquence codant 6 histidines. La protéine recombinante, étiquetée en N-terminal par ce peptide hexahistidine, peut être purifiée par chromatographie d'affinité au nickel. L'expression des protéines fusions dans le vecteur pFO4 est sous le contrôle du promoteur nommé *T7lac* (Figure 6-1). Ce plasmide contient une séquence opératrice *lac* juste en aval du promoteur du phage T7. Ce type de vecteur doit donc être transformé dans une souche d'*E. coli* lysogénisée par le prophage DE3, telle que BL21(DE3), qui contient la séquence du gène de la T7 RNA polymérase en aval d'un promoteur *lacUV5*. En absence d'induction par le lactose ou l'IPTG, le répresseur LacI agit à deux endroits: (i) sur le promoteur *lacUV5* du chromosome bactérien pour bloquer la transcription de la T7 RNA polymérase et (ii) sur le promoteur *T7lac* du vecteur pour bloquer la transcription du gène cible par les éventuelles traces de T7 RNA polymérase produite malgré la répression. Cette stratégie de "double-contrôle" permet de limiter au maximum l'expression de fuite de certaines protéines recombinantes qui pourraient s'avérer toxiques pour la souche d'expression.





**Figure 6-1 : Principe de l'expression dans le vecteur pFO4.**

adapté de la brochure Novagen pET system manual.

Cette stratégie de clonage dans deux vecteurs différents, ainsi que le traitement en parallèle d'un nombre important de cibles d'intérêt, permet de maximiser les chances d'obtenir des protéines recombinantes pour des études ultérieures. En effet, l'expression, le repliement et donc la solubilité des protéines peut varier en fonction du type de vecteur utilisé. Ce traitement à moyen débit nécessite une normalisation des procédures utilisées à chaque étape, telles que l'analyse bioinformatique, le dessin des amorces, la préparation des vecteurs, etc., afin de minimiser les risques de biais inhérents à chacune des cibles. A l'arrivée, ce type de stratégie permet souvent d'obtenir un assez grand nombre de protéines surexprimées. Il est alors nécessaire de faire des choix quant à celles à produire à plus grande échelle et à étudier biochimiquement et structurellement.

## II- Matériels et Méthodes

### II-1 Analyses bioinformatiques des séquences

La présence de peptides signaux et la position du site de clivage ont été prédites en utilisant les logiciels SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP/>) et LipoP 1.0 (<http://www.cbs.dtu.dk/services/LipoP/>). SignalP combine deux algorithmes basés sur des réseaux neuronaux et des automates de Markov à états cachés (Hidden Markov Model, HMM) entraînés sur des jeux de séquences eucaryotes ou de bactéries Gram-négatives ou positives (Emanuelsson *et al.*, 2007). LipoP est basé sur un algorithme HMM et prédit les peptides signaux lipoprotéiques chez les bactéries Gram-négatives (Juncker *et al.*, 2003). Le serveur identifie les lipoprotéines (signal clivé par la signal peptidase SPaseII), les protéines clivées par la signal peptidase SPaseI, les protéines cytoplasmiques et les protéines transmembranaires. La prédiction d'hélices transmembranaires a été approfondie en utilisant l'outil TMHMM 2.0 (Sonnhammer *et al.*, 1998; Krogh *et al.*, 2001) [<http://www.cbs.dtu.dk/services/TMHMM/>].

Afin de définir d'éventuels modules, les séquences protéiques ont été analysées par différents outils bio-informatiques. Les modules conservés ont été prédits par les outils Pfam (<http://pfam.jouy.inra.fr/hmmsearch.shtml>), CDD (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) and ScanProsite (<http://www.expasy.org/tools/scanprosite/>). La présence éventuelle de modules de fonction inconnue et non recensés dans les bases de données a été détectée par des recherches BLAST contre la base UNIPROT. Les délimitations fines des différents modules protéiques ont ensuite été définies manuellement par la méthode HCA (Hydrophobic Cluster Analysis, <http://mobylye.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py?form=HCA>) (Gaboriaud *et al.*, 1987; Callebaut *et al.*, 1997).

### II-2 Dessin des amorces oligonucléotidiques

Afin de définir la stratégie de clonage, les séquences nucléotidiques ont été scannées dans DNAMAN 1 pour la présence de site de restriction des endonucléases BamHI (G\*GATCC, \* représente le site de clivage), BglII (A\*GATCT), EcoRI (G\*AATTC), MfeI (C\*AATTG) et XhoI (C\*TCGAG). Les enzymes utilisables pour une cible donnée sont celles dont le site de restriction est absent de la séquence. Lorsque cela était possible, la stratégie préférentielle a

été l'utilisation de BamHI / EcoRI. En cas d'incompatibilité, l'utilisation des enzymes BglII, MfeI et XhoI a été envisagée.

Les amorces nucléotidiques ont été dessinées selon le schéma suivant:

1) Amorce sens:



- 6 bases G
- Site de restriction (BamHI ou BglII)
- Séquence de la cible à cloner (sans le codon d'initiation qui est apporté par le vecteur)

2) Amorce anti-sens:



- 6 bases C
- Site de restriction (EcoRI ou MfeI ou XhoI)
- TTA, qui apporte dans la séquence le codon stop TAA (le plus fréquent chez *E.coli*)
- Séquence complémentaire inverse de la région 3' de la cible à cloner

Les longueurs des amorces ont été ajustées afin d'obtenir une unique température de semi-hybridation  $T_m = 70^\circ\text{C}$ . Le détail des séquences des amorces dessinées est donné dans le Tableau 6-1.

N°cible	N°gène	Modularité	Fonction potentielle	Taille des fragments géniques clonés (pb)	Stratégie de clonage	amorce sens amorce anti-sens
1	zg4327	prot. entière	alginate lyase	1101	BamHI EcoRI	ggggggGGATCCgatgccttaggggtttcattttg ccccccGAATTCtcaaactttataaatcatcttggaatc
2	zg4327	module 1	module catalytique	702	BamHI EcoRI	ggggggGGATCCgatgccttaggggtttcattttg ccccccGAATTCttaacctattcggtgggtcgatac
3	zg4327	module 2	inconnue	264	BamHI EcoRI	ggggggGGATCCcaaaaccctagtgatactagcgatac ccccccGAATTCtcaaactttataaatcatcttggaatc
4	zg3452	-	oligo-alginate lyase	2304	BamHI EcoRI	ggggggGGATCCaaaagaacacttctatatatattttcttg ccccccGAATTCttattccgtgatttttagttggatc
5	zg3445	-	unsaturated glucuronyl hydrolase	1149	BamHI EcoRI	ggggggGGATCCtgtgcttcacaaaagaagaacgg ccccccGAATTCttattgttttatagaaagccccaatttc
6	zg3804	-	unsaturated glucuronyl hydrolase	1170	BamHI EcoRI	ggggggGGATCCtgtaccttataaaaagaaggactaaac ccccccGAATTCttaattttcatgatactgttctttataacc
7	zg2623	-	inconnue	348	BamHI EcoRI	ggggggGGATCCgaaagttttgatcaagtaagag ccccccGAATTCttatttagtcataaaatcttccctg
8	zg2622	-	dehydrogenase/reductase	750	BamHI EcoRI	ggggggGGATCCgatttaaaaggaaaattgcccgttgc ccccccGAATTCttatgaaaaggcgagtcgccattg
9	zg2619	prot. entière	inconnue	798	BamHI EcoRI	ggggggGGATCCtgtgaagacaccttcgagtacg ccccccGAATTCttaagggtgttcttctactggagc
10	zg2619	module 1	inconnue	372	BamHI EcoRI	ggggggGGATCCtgtgaagacaccttcgagtacg ccccccGAATTCttagtcaagacaactgtcatcgac
11	zg2619	module 2	CBM	429	BamHI EcoRI	ggggggGGATCCagtagggtgttggagaattgaag ccccccGAATTCttaagggtgttcttctactggagc
12	zg2615	-	acetoin reductase	696	BamHI EcoRI	ggggggGGATCCaaaagattagggcaagatgggtatac ccccccGAATTCttagattcctaaagcttgcccacc
13	zg4130	prot. entière	alginate lyase	2256	BamHI EcoRI	ggggggGGATCCaacgataaagggtttacgtaaag ccccccGAATTCttaatacatcaatttgagtgaggag
14	zg4130	module 1	module catalytique	1341	BamHI EcoRI	ggggggGGATCCaacgataaagggtttacgtaaagaatag ccccccGAATTCttattgttccgcccctggagttccc
15	zg4130	module 2	module catalytique	918	BamHI EcoRI	ggggggGGATCCaacgcggtcggtgccattg ccccccGAATTCttaatacatcaatttgagtgaggag
16	zg4131	-	alginate lyase	1011	BamHI EcoRI	ggggggGGATCCtgtaaagacaaacctaaaggccactacg ccccccGAATTCtactgggtctctggggcgcttc
26	zg2618	-	alginate lyase	858	BamHI MfeI	ggggggGGATCCtgccaagacagtaatgacttggcag ccccccCAATTGctagtgtataacttctaatacgtaaaatttc

27	zg2617	-	facteur de transcription	726	BamHI MfeI	ggggggGGATCCaaactagaaatattgaccaaacacg ccccccCAATTGttatccttttccaattttaatttttacg
28	zg2613	-	phosphotransferase	1164	BamHI MfeI	ggggggGGATCCgttataagcaccgtagctaaaatttttc ccccccCAATTGttacacttctccggatagaattttgttg
31	zg1182	prot. entière	alginate lyase	1275	BglII EcoRI	ggggggAGATCTcccgtatTTTTccttgcccttatgcc ccccccGAATTCttaattgtgggttacgcttaggtttttg
32	zg1182	module 1	CBM 32	501	BglII EcoRI	ggggggAGATCTcccgtatTTTTccttgcccttatgc ccccccGAATTCttaattgtgggttacgcttaggtttttg
33	zg1182	module 2	module catalytique	777	BglII EcoRI	ggggggAGATCTtccgggtggttcgtccacccttc ccccccGAATTCttaattgtgggttacgcttaggtttttg
34	zg2614	-	kinase à sucres	1038	BglII EcoRI	ggggggAGATCTaaaaaagtagtaaccttcggagagatc ccccccGAATTCttaccgcgctacccttcgggaagc
35	zg2620	-	SusD/RagB lipoprotéine	1374	BglII MfeI	ggggggAGATCTtgcgaaaaggacttcttgtctccc ccccccCAATTGttagtatcccggattttgactcattg
36	zg2620	module 1	SusD/RagB lipoprotéine	1173	BglII MfeI	ggggggAGATCTaaaagtagcgagggtgaagcagc ccccccCAATTGttagtatcccggattttgactcattg
39	zg2624	prot. entière	alginate lyase	2181	BamHI XhoI	ggggggGGATCCcaagaacacccgagtctgatactg ccccccCTCGAGttaattgattttaataaattgatagggcc
40	zg2624	module 1	module catalytique	1131	BamHI XhoI	ggggggGGATCCcaagaacacccgagtctgatactgac ccccccCTCGAGtttagccgttgccccatcgctaag
41	zg2624	module 2	module catalytique	474	BamHI XhoI	ggggggGGATCCgaccaagggggcgtggccatac ccccccCTCGAGtttaggcattcgaaccgacccttag
42	zg2624	module 3	inconnue	582	BamHI XhoI	ggggggGGATCCgccaaccaatacgccttcccttc ccccccCTCGAGttaattgattttaataaattgatagggcc
43	zg4132	prot. entière	alginate lyase	2268	BamHI XhoI	ggggggGGATCCcaacaaaaaaaaatacgggtcaagaacg ccccccCTCGAGttattgggaaattacgccccaatcg
44	zg4132	module 1	module catalytique	1251	BamHI XhoI	ggggggGGATCCcaacaaaaaaaaatacgggtcaagaacg ccccccCTCGAGtttagatattcggacttatggttacc
45	zg4132	module 2	module catalytique	1020	BamHI XhoI	ggggggGGATCCgccaccaaggaaaatacagggg ccccccCTCGAGttattgggaaattacgccccaatcg
46	zg2621	-	récepteur TonB dépendant	3063	BamHI BglII	ggggggGGATCCcaagacggttacacactgtcag ccccccAGATCTttaaaaaatcaagggttaagtcctatag
47	zg2616	-	perméase à sucres	1275	BamHI BglII	ggggggGGATCCaaagtcaaaggtttaagatgggtggg ccccccAGATCTttaactgactcaagttctacgggg
48	zg2612	-	transporteur à Mn/Fe	1266	BamHI BglII	ggggggGGATCCcttgaagagaaaccgggcttgc ccccccAGATCTctagaataattccagagggccac

**Tableau 6-1: Tableau récapitulatif des cibles clonées et des amorces oligonucléotidiques utilisées**

## II-3 Préparation du matériel biologique

### II-3-a Préparation des cellules DH5 $\alpha$ compétentes

#### Milieus et Solutions

##### 1) Milieu SOB

20 g.L<sup>-1</sup> tryptone, 5 g.L<sup>-1</sup> extrait de levure, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose. Préparer une solution sans magnésium ni glucose à autoclaver, puis ajouter ces composants filtrés à 0,2  $\mu$ m au dernier moment sous la hotte.

##### 2) Solution RF1

100 mM RbCl, 50 mM MnCl<sub>2</sub>, 10 mM CaCl<sub>2</sub>, 30 mM CH<sub>3</sub>CO<sub>2</sub>K, 15% glycérol  
Ajuster à pH 5,8 avec de l'acide acétique concentré. Filtrer sur 0,2  $\mu$ m.

##### 3) Solution RF2

10 mM RbCl, 75 mM CaCl<sub>2</sub>, 10 mM MOPS, 15% glycérol  
Ajuster à pH 6,8 avec NaOH. Filtrer sur 0,2  $\mu$ m.

#### Mode opératoire

Les cellules compétentes ont été préparées selon la méthode de Hanahan (1983). Une préculture (2 mL) de la souche DH5 $\alpha$  d'*E. coli* a été réalisée pendant une nuit à 37°C en milieu SOB. Un volume de 100 mL de milieu SOB préalablement chauffé à 37°C en erlenmeyer de 1 L a été inoculé avec 1 mL de préculture et incubé à 37°C et 180 rpm. Lorsque l'absorbance à 600 nm a atteint une valeur comprise entre 0,4 et 0,6 (environ 2 h), la culture a été refroidie sur glace pendant 10 min puis centrifugée à 1000 g à 4°C pendant 10 min. Le culot cellulaire a été repris dans un tiers du volume de culture initial (soit 30 mL) de solution RF1 préalablement refroidie à 4°C. La suspension a été incubée sur glace pendant 15 min puis centrifugée à 1000 g à 4°C pendant 10 min. Le culot cellulaire a été repris dans 8% du volume de culture initial (soit 8 mL) de solution RF2 à 4°C. La suspension contenant les cellules compétentes a été incubée 30 min sur glace puis répartie en aliquots de 300  $\mu$ L en tubes cryogéniques, congelée dans l'azote liquide et stockée à -80°C. La compétence des cellules a été testée en réalisant des transformations avec un vecteur intact portant un gène de résistance à l'ampicilline (pGEX et pFO4). Le nombre de colonies après croissance sur milieu

gélosé additionné d'ampicilline dépassait toujours  $10^6$  CFU /  $\mu\text{g}$  de vecteur, attestant de la bonne compétence des cellules DH5 $\alpha$ .

### II-3-b Préparation des vecteurs d'expression

Les plasmides pFO4 et pGEX-4T-1 ont été purifiés sur gradient de chlorure de césium. 10  $\mu\text{g}$  de chaque vecteur ont été linéarisés pendant 4 h à 37°C par action du couple d'endonucléases approprié dans 50  $\mu\text{L}$  de mélange réactionnel contenant 1X du tampon NEB adéquat, 100  $\mu\text{g}\cdot\text{mL}^{-1}$  BSA et 25 U de chaque enzyme (New England Biolabs). La linéarisation a été vérifiée en analysant 100 ng de plasmide en électrophorèse sur gel d'agarose. Les vecteurs ont ensuite été purifiés sur colonne Qiaquick (Qiagen) selon les instructions du fournisseur et élués dans 50  $\mu\text{L}$ . Les vecteurs ont été déphosphorylés par action de la Shrimp Alkaline Phosphatase (SAP, New England Biolabs) pendant 45 min à 37°C dans un mélange réactionnel contenant 1X du tampon fourni et 5 U d'enzyme. La SAP a été inactivée en chauffant 15 min à 65°C. Après une purification sur colonne Qiaquick (Qiagen), les solutions de vecteurs ont été dosées au Nanodrop et diluées à 2.5  $\text{ng}\cdot\mu\text{L}^{-1}$  en eau.

### II-3-c Extraction et purification de l'ADN génomique de *Z. galactanivorans*

#### 1) Lyse des cellules

La bactérie *Z. galactanivorans* a été cultivée à 20°C jusqu'à la phase stationnaire dans 250 mL de milieu Zobell (5  $\text{g}\cdot\text{L}^{-1}$  tryptone, 1  $\text{g}\cdot\text{L}^{-1}$  extrait de levure en eau de mer). La culture a été centrifugée 10 min à 5000  $\text{g}$ , puis le culot cellulaire a été repris dans un volume d'eau de mer filtrée. La suspension a été centrifugée lentement (2000  $\text{g}$ , 10 min) pour éliminer d'éventuels polysaccharides. Les volumes de solution ci-dessous sont donnés pour 10 g de cellules. Les cellules ont été reprises dans 25 mL de solution 1 (50 mM Tris-HCl pH 8,0, 25% saccharose), puis 5 mL de solution 2 (50 mM Tris-HCl pH 8,0, 5 mM EDTA, 50 mM NaCl) et 50 mg de lysozyme ont été ajoutés. La suspension a été incubée 10 min sur glace, puis 10 mL d'EDTA 500 mM ont été ajoutés. Après 10 min sur glace, la suspension a été amenée à 2% SDS final et agitée manuellement jusqu'à apparition de la viscosité.

## 2) Extraction de l'ADN génomique

L'extrait obtenu a été additionné de 25 mL de solution 3 (50 mM Tris-HCl pH 8,0, 100 mM EDTA, 100 mM NaCl) et 500  $\mu\text{g}\cdot\text{mL}^{-1}$  de protéinase K puis incubé 1 h à 50°C. Les contaminants protéiques ont été éliminés par une extraction avec 1 volume de phénol/chloroforme (1:1, v/v). La phase aqueuse a été amenée à 1 M de perchlorate de sodium et l'ADN a été extrait par 1 volume de chloroforme/alcool isoamylique (24:1, v/v). Après centrifugation et récupération de la phase aqueuse, l'ADN a été précipité par ajout de 2 volumes d'éthanol 100%, récupéré à l'aide d'une baguette de verre et resuspendu dans du tampon TE (10 mM Tris-HCl pH 8,0, 1 mM EDTA). La concentration a été estimée par mesure de l'absorbance à 260 nm au Nanodrop.

## 3) Purification de l'ADN sur gradient de chlorure de césium

La solution d'ADN a été amenée à 4,5 M de CsCl et 600  $\mu\text{g}\cdot\text{mL}^{-1}$  de bromure d'éthidium en tampon TE, puis délicatement coulée en tube à sceller qui a été centrifugé à 50000 g pendant 48 h à température ambiante. L'ADN a été prélevé à l'aide d'une seringue sous UV. Le bromure d'éthidium a été éliminé par des extractions successives avec un volume de butanol saturé en eau suivies d'une centrifugation (2 min, 14000 g), jusqu'à ce que la phase aqueuse inférieure ne fluoresce plus sous UV. L'ADN a alors été dilué dans 2 volumes d'eau et précipité par 2,5 volumes d'éthanol absolu 2 h à -20°C. La solution a été centrifugée 30 min à 10000 g et 4°C et le culot a été rincé dans 1 volume d'éthanol 70% froid. Après centrifugation, le culot d'ADN a été repris dans l'eau.

### **II-3-d Préparation des produits PCR pour le clonage**

Les réactions de PCR (Polymerase Chain Reaction) ont été réalisées dans un volume final de 50  $\mu\text{L}$  de mélange réactionnel comprenant 10 ng d'ADN génomique purifié, 1 U d'ADN polymérase Pfu (Promega), 1X du tampon commercial fourni, 400  $\mu\text{M}$  de chaque dNTP, 1 mM  $\text{MgCl}_2$  et 100 nM de chaque amorce. Les cycles de PCR ont été réalisés comme suit: 5 min à 94°C, suivi de 35 cycles de 30 s à 94°C, 30 s à la température d'hybridation choisie et 4 min à 72°C, pour finir par 10 min à 72°C. Deux températures d'hybridation ont été testées: 50°C et 60°C. Les produits d'amplification (5 $\mu\text{L}$ ) ont été analysés en électrophorèse sur gel d'agarose 0,8% en TAE (40 mM Tris-HCl pH 8,0, 20 mM acétate de sodium, 1 mM EDTA) afin de vérifier la présence d'une bande unique à la taille attendue. Les produits d'amplification ont été purifiés sur colonne Qiaquick (Qiagen) et élués dans 60  $\mu\text{L}$ .



Ils ont ensuite été digérés 4 h à 37°C par le couple d'endonucléases approprié dans un volume final de 30 µL de mélange réactionnel comprenant : 20 µL de produits PCR purifiés, 1X du tampon NEB adéquat, 100 µg.mL<sup>-1</sup> BSA et 10 U de chaque enzyme (New England Biolabs). Les produits PCR digérés ont enfin été purifiés sur colonne Qiaquick (Qiagen), élués dans 50 µL et conservés à -20°C.

## **II-4 Clonage**

### **II-4-a Ligature des produits PCR dans les vecteurs d'expression**

Les réactions de ligature ont été réalisées en plaque 96 puits dans un volume final de 10 µL de mélange réactionnel composé de : 0.4 ng de plasmide digéré (pFO4 ou pGEX-4-T1), 5 µL de produits PCR, 100 U de T4 DNA ligase (New England Biolabs) et 1X du tampon commercial. Les réactions ont été incubées une nuit à 13°C, puis congelées à -20°C.

### **II-4-b Transformation en souche de stockage et vérification du succès du clonage**

Les transformations ont eu lieu directement dans les puits des réactions de ligature auxquels ont été ajoutés 30 µL de cellules DH5α compétentes. L'étape de transformation a été réalisée par choc thermique. Après incubation 30 min sur un mélange glace / eau / sel, les réactions ont été transférées dans un bain-marie à 42°C pendant 45 s, puis déposées sur le mélange glacé pendant 10 min. 100 µL de milieu SOC (2% tryptone; 0,5% extrait de levure; 0,05% NaCl; 0,2% glucose) ont été ajoutés dans chaque puits. Les cellules transformées ont alors été incubées 35 min à 37°C et 900 rpm. 50 µL de cellules ont été étalés sur des boîtes de Petri en plaque de 12 puits contenant 2 mL de milieu LB-agar supplémenté d'ampicilline à 100 µg.mL<sup>-1</sup>. Les boîtes ont été incubées 12 h à 37°C.

Le succès des étapes de clonage et de transformation a été testé par réaction de PCR directement sur les colonies bactériennes obtenues, en utilisant des amorces nucléotidiques flanquant le site de clonage du vecteur (Tableau 6-2). Les réactions ont eu lieu dans un volume final de 10 µL composé de : 1X de mélange PCR Master Mix (Promega), 1 µM de chaque amorce et quelques cellules prélevées à partir d'une colonie à l'aide d'un cône stérile.

Vecteur	Amorce sens	Amorce Anti-sens
pFO4	GCA GCA GCC ACC ATC ACC ATC ACC	CCT TTC GGG CTT TGT TAG CAG CCG G
pGEX-4-T1	CCT CCA AAA TCG GAT CTG GTT CCG CG	CGA TGC GGC CGC TCG AGT CGA CCC G

**Tableau 6-2 : Amorces universelles des vecteurs de clonage**

Les cycles de PCR ont été réalisés selon la séquence suivante : 5 min à 94°C, suivi de 30 cycles de 30 s à 94°C, 30 s à 47°C et 3 min à 72°C, pour finir par 10 min à 72°C. Dix microlitres des réactions d'amplification ont été analysés en électrophorèse sur gel d'agarose 0,8%. L'obtention d'une bande unique à la taille attendue a permis la validation des colonies testées.

Pour chaque colonie validée, des cultures de 2 mL en milieu LB additionné d'ampicilline à 100 µg.mL<sup>-1</sup> ont été inoculées en plaque de 96 puits et incubées 12 h à 37°C sous agitation. 1 mL de ces cultures a servi à stocker les cellules DH5α transformées dans 25% glycérol en cryotubes à -80°C. Le reste a été centrifugé 15 min à 1500 rpm et utilisé pour l'extraction des plasmides sur un robot Tecan Genesis 150.

#### II-4-c Transformation en souche d'expression et vérification

Les plasmides recombinants pGEX-4-T1 et pFO4 obtenus à l'étape précédente ont été transformés dans les souches d'expression d'*E. coli* BL21 et BL21(DE3), respectivement. Les cellules compétentes (25 µL) ont été mélangées avec 2 µL de solution de plasmides et transformées selon le protocole précédemment décrit (II.4.b). Les colonies obtenues ont été criblées pour la présence de l'insert selon le protocole décrit. Les clones validés ont été stockés à -80°C en 25% glycérol.

#### II-5 Tests d'expression

Des précultures des souches d'expression recombinantes BL21 et BL21(DE3) (Novagen, USA) ont été réalisées en milieu LB additionné d'ampicilline à 100 µg.mL<sup>-1</sup> (12 h à 37°C). 30 µL de préculture ont été dilués dans 3 mL de milieu auto-inductible ZYP5052 (Studier 2005) avec ampicilline en plaque de 24 puits (contenance 15 mL). Après incubation 3 jours à 19°C et 1200 rpm, les cellules ont été centrifugées à 2500 g pendant 30 min. Les culots ont été congelés 1 h à -20°C, décongelés, puis repris dans 500 µL de tampon de lyse (25 mM Tris-HCl pH 7,5; 500 mM NaCl; 0,1 mg.mL<sup>-1</sup> DNase; 1 mg.mL<sup>-1</sup> lysozyme; une demi pastille de cocktail d'antiprotéases Complete [Roche] pour 50 mL de tampon). Les cellules ont été

incubées 1 h à 18°C sous agitation, puis 1 h à 4°C et centrifugées 20 min à 12000 g et 4°C. Les surnageants ont été récupérés (extrait soluble), et les culots ont été repris dans 200 µL de tampon de lyse additionné d'urée à 6 M (extrait insoluble). Chaque fraction a été analysée en électrophorèse en condition dénaturante sur gel de polyacrylamide 12% (SDS-PAGE) coloré au Bleu de Coomassie. Les fractions solubles (environ 500 µL) ont été purifiées sur minicolonnes Microspin His (vecteur pFO4) ou GST (pGEX-4-T1) (GE Healthcare Life Science). Pour la purification sur HisSpinTrap, les échantillons ont été amenés à 20 mM imidazole, puis purifiés selon le protocole du fournisseur avec adaptation du tampon de lavage (20 mM imidazole; 25 mM Tris-HCl pH 7,5; 500 mM NaCl) et d'élution (500 mM imidazole; 25 mM Tris-HCl pH 7,5; 500 mM NaCl). Pour la purification sur minicolonne GST, les échantillons ont été purifiés selon le protocole du fournisseur avec adaptation du tampon de lavage (25 mM Tris-HCl pH 7,5; 500 mM NaCl) et d'élution (10 mM glutathion réduit; 25 mM Tris-HCl pH 7,5; 500 mM NaCl). Dix microlitres des fractions purifiées ont été analysés en électrophorèse en condition dénaturante sur gel de polyacrylamide 12% (SDS-PAGE) coloré au Bleu de Coomassie.

### III- Résultats et Discussion

#### III-1 Choix des cibles à cloner et analyse de séquences

Le travail d'annotation du génome de *Zobellia galactanivorans* et l'analyse détaillée de la voie alginolytique (Chapitre 4) ont permis de mettre en évidence un certain nombre de gènes impliqués dans la dégradation de l'alginate. Nous avons choisi d'intégrer tous ces gènes dans le programme de clonage à moyen débit. Il s'agit des 3 alginate lyases de l'opéron à 6-kb, des 12 gènes de l'opéron à 15,9-kb, et de 6 gènes isolés dans le génome (*alyA1*, *alyA7*, *uglA1*, *uglA2*, *zg3452* et *zg2612*).

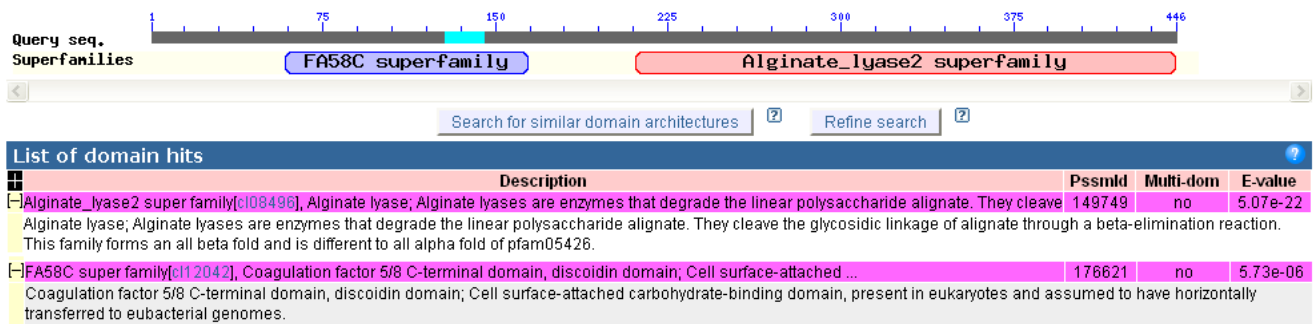
Une analyse des séquences protéiques d'intérêt a été entreprise afin de définir les limites des cibles à cloner. Dans un premier temps, une recherche d'éventuels peptides signaux a été entreprise. Chez les bactéries Gram-négatives, l'adressage de protéines vers le périplasme ou le milieu extracellulaire nécessite le plus souvent la présence d'une séquence signal clivable. Il existe au moins trois types de signal peptidases (SPases) responsables du clivage de différentes séquences d'adressage chez les bactéries (Paetzel *et al.*, 2002). Les protéines "classiques" (par opposition aux lipoprotéines) transitent soit par la voie SecYEG, soit par

la voie Tat (Twin Arginin Translocation) et sont clivées par la SPase I. La présence d'un peptide signal reconnu par la SPase I suggère donc que la protéine est au moins sécrétée dans le périplasme, puis éventuellement relâchée dans le milieu extracellulaire. Les lipoprotéines transitent par la voie Sec et sont clivées par la SPase II. Après clivage, trois acides gras attachés à la cystéine N-terminale permettent l'ancrage des lipoprotéines dans la membrane externe en général. La SPase IV clive les prépilines nécessaires à la synthèse du pilus de type IV et les protéines de l'appareil de sécrétion de type II. En revanche, la présence d'un peptide signal non clivé ou d'hélices transmembranaires suggère une localisation sur la membrane cytoplasmique. Les résultats de la recherche de peptides signaux dans les 21 séquences d'intérêt sont donnés dans le Tableau 6-3. Les séquences nucléotidiques codant les peptides signaux détectés ont été supprimées des cibles à cloner, afin de maximiser les chances d'obtenir une protéine soluble dans le cytoplasme d'*E.coli*.

n° gène	fonction prédite	Peptide signal	Hélices transmembranaires	Localisation probable
zg1182	AlyA1	SPase II	-	Membrane externe
zg2612	Transporteur Mn/Fe	-	11	Membrane cytoplasmique
zg2613	Pfp	-	-	Cytoplasmique
zg2614	KdgK	-	-	Cytoplasmique
zg2615	SDR	SPase I	-	Périplasme / Sécrétée
zg2616	Perméase à sucres	-	11	Membrane cytoplasmique
zg2617	GntR-like	-	-	Cytoplasmique
zg2618	AlyA2	Spase II	-	Membrane externe
zg2619	PKD/CBM	SPase II	-	Membrane externe
zg2620	SusD-like	SPase II	-	Membrane externe
zg2621	SusC-like	SPase I	-	Membrane externe
zg2622	SDR	-	-	Cytoplasmique
zg2623	KdgF	-	-	Cytoplasmique
zg2624	AlyA3	SPase I	-	Périplasme / Sécrétée
zg4130	AlyA4	Non clivé	-	Membrane cytoplasmique
zg4131	AlyA5	SPase II	-	Membrane externe
zg4132	AlyA6	SPase I	-	Périplasme / Sécrétée
zg3445	UglA1	SPase II	-	Membrane externe
zg3452	PL15	Non clivé	-	Membrane cytoplasmique
zg3804	UglA2	SPase II	-	Membrane externe
zg4327	AlyA7	SPase I	-	Périplasme / Sécrétée

**Tableau 6-3 : Prédiction de peptides signaux et d'hélices transmembranaires dans les séquences SPase I : séquence signal clivée par la Signal Peptidase I; SPase II : séquence signal clivée par la SPase II.**

La structure modulaire des protéines a ensuite été analysée. Dans le cas de protéines modulaires, plusieurs constructions ont été envisagées afin de cloner chaque module séparément ainsi que la séquence entière. En effet, la présence de plusieurs modules séparés par des régions flexibles peut nuire à l'expression des protéines recombinantes et surtout à la cristallogénèse. De plus, l'étude de chacun des modules séparément peut donner des informations sur leur contribution à l'activité de la protéine entière. Un exemple d'analyse de modularité est donné par la suite pour la protéine AlyA1 (zg1182). La recherche de domaines conservés dans la séquence protéique d'AlyA1 (446 acides aminés) prédit deux modules (Figure 6-2).



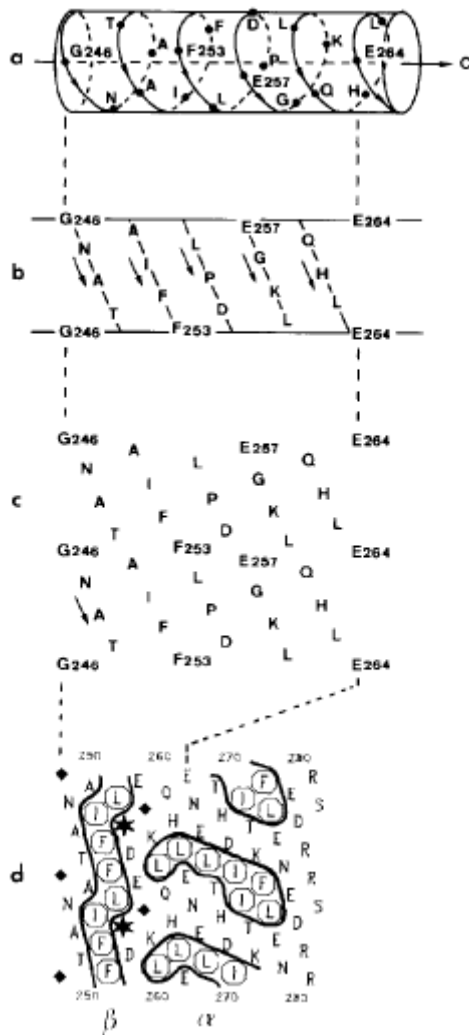
**Figure 6-2 : Résultat de l'analyse de la séquence protéique d'AlyA1 sur le serveur CDD (NCBI)**

Le premier, de la position 59 à 164, a une homologie pour la super-famille FA58C, qui contient notamment des domaines de liaison aux sucres. Le second domaine détecté, de la position 211 à 445, appartient à la super-famille Alginate lyase 2. Une recherche de type BLASTP contre la Protein Data Bank (PDB) a ensuite été effectuée pour les deux régions séparément. Pour chaque région, les cibles donnant le meilleur score ont été recherchées sur la base de données CAZy [<http://www.cazy.org/>; (Cantarel *et al.*, 2009)] afin d'assigner une famille potentielle à chacun des modules (Tableau 6-4).

Région dans AlyA1	Meilleur hit du Blastp (code PDB)	Annotation	Famille CAZy
0-164	2JD9	Chain A, Structure Of A Pectin Binding Carbohydrate Binding Module from <i>Yersinia enterocolitica</i>	CBM32
211-446	2Z42	Chain A, Crystal Structure Of Family 7 Alginate Lyase A1-Ii' From <i>Sphingomonas</i> Sp. A1	PL7

**Tableau 6-4 : Résultats du Blastp pour les deux modules d'AlyA1**

Grâce à cette analyse, la structure modulaire d'AlyA1 a pu être mise en évidence. Afin de définir les bornes des régions à cloner, une délimitation fine des limites des modules a été effectuée par la méthode des amas hydrophobes (Hydrophobic Cluster Analysis, HCA) (Gaboriaud *et al.*, 1987). Le principe de cette méthode est de prédire des éléments de structure secondaire à partir de la séquence primaire d'une protéine. La séquence primaire est écrite sur



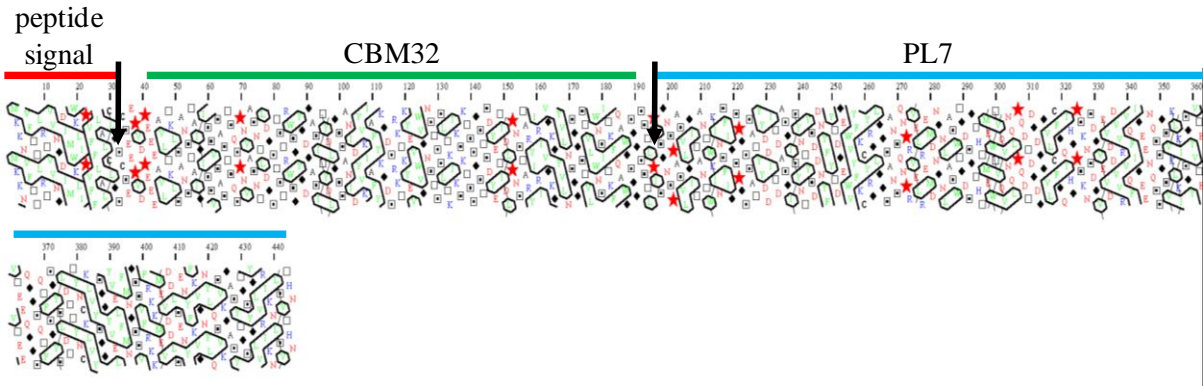
**Figure 6-3 : Principe de la méthode HCA (d'après Gaboriaud et al., 1987)**

une hélice  $\alpha$  classique (3,6 acides aminés par tour) le long d'un cylindre (Figure 6-3.a). Le cylindre est découpé selon son axe et déroulé (Figure 6-3.b). A la suite de ce traitement, certains acides aminés consécutifs dans la séquence se retrouvent artificiellement éloignés. Pour pallier à ce problème, une deuxième représentation est recopiée à l'identique sous la première (Figure 6-3.c). Les acides aminés sont alors symbolisés selon leurs propriétés. Les glycines ( $\blacklozenge$ ) et prolines ( $\blackstar$ ), qui provoquent des changements de direction dans la structure, ainsi que les thréonines ( $\square$ ) et sérines ( $\blacksquare$ ) qui peuvent présenter une certaine

amphipathie, sont représentées par des symboles pour les repérer facilement. De plus, les acides aminés hydrophobes (I, L, F, W, M, Y, V) sont représentés en vert entourés d'un liseré noir, les acides ou polaires (D, E, N, Q) en rouge et les basiques (K, R, H) en bleu. Dans cette représentation, les groupes d'acides aminés hydrophobes sont facilement repérables par la fusion

des liserés noirs (Figure 6-3.d). Ces amas hydrophobes permettent de prédire avec une bonne probabilité des éléments de structure secondaire. Les amas verticaux suggèrent la présence de brins  $\beta$ , les amas horizontaux celle d'hélices  $\alpha$ . Une région non-repliée, typique de la liaison entre deux modules structuraux, se traduit dans le diagramme HCA par une zone dépourvue d'amas hydrophobes et riches en acides aminés polaires (en rouge ou bleu) ou de petite taille (marqués d'un symbole). Le diagramme HCA obtenu pour la protéine AlyA1 est représenté

sur la Figure 6-4. Pour être adéquate, la délimitation d'un module doit se faire dans une région non structurée, présentant peu d'amas hydrophobes.



**Figure 6-4 : Analyse modulaire de l'alginate lyase AlyA1.**  
Les limites choisies pour les modules à cloner sont indiquées par les flèches noires.

Cette analyse de modularité, appliquée aux 21 séquences de gènes à cloner, a permis de définir 35 modules qui ont été traités en parallèle dans l'approche à moyen débit (Tableau 6-1).

### III-2 Clonage à moyen débit

L'approche de clonage et d'expression à moyen débit s'est basée sur un protocole développé au laboratoire (Groisillier *et al.*, 2010), permettant de traiter un grand nombre de cibles en parallèle. Les résultats des différentes étapes du clonage et des tests d'expression à moyen débit sont récapitulés dans le Tableau 6-5.

L'amplification par PCR des fragments géniques à partir de l'ADN génomique de *Zobellia galactanivorans* a abouti pour 31 des 35 cibles de départ, soit un taux de réussite de 89%. Deux vecteurs d'expression différents ont été utilisés pour tester l'expression et la solubilité des protéines recombinantes chez *E. coli*. Le premier vecteur (pGEX-4-T1) permet la fusion de la protéine d'intérêt avec une glutathion-S-transférase en N-terminal, ce qui facilite par la suite la purification sur colonne d'affinité au glutathion. Le second vecteur (pFO4) permet l'étiquetage de la protéine avec 6 histidines N-terminales (His<sub>6</sub>-tag), pour une purification sur colonne d'affinité au nickel. L'utilisation de deux vecteurs de clonage en parallèle avait pour but de multiplier les chances de succès. Un même produit d'amplification digéré par les

endonucléases adéquates a été utilisé pour des ligatures dans pFO4 et pGEX-4-T1. Après transformation dans la souche de stockage DH5 $\alpha$ , les colonies ont été criblées par PCR pour la présence de l'insert. A ce stade, 22 gènes ont été clonés avec succès dans les deux vecteurs, un gène uniquement dans pGEX-4-T1 (zg1182, n°20) et 1 gène uniquement dans pFO4 (zg2612, n°35) (Tableau 6-5). Au total, 77% des cibles amplifiées par PCR ont été clonées dans l'un ou l'autre des vecteurs. De manière surprenante, aucune des cibles utilisant la stratégie de clonage BamHI / XhoI (7 cibles sur 35) n'a pu être obtenue dans les vecteurs de clonage, ce qui diminue assez fortement le taux de réussite global. Après plusieurs essais infructueux, une explication plausible serait un défaut de coupure des produits PCR par l'enzyme XhoI utilisée. Ces 7 cibles correspondent à deux alginate lyases, AlyA3 et AlyA6, qui constituent des enzymes particulièrement intéressantes dans l'étude du métabolisme de l'alginate chez *Z. galactanivorans*. Il pourrait donc être intéressant de tenter une stratégie de clonage alternative avec une autre enzyme de restriction.



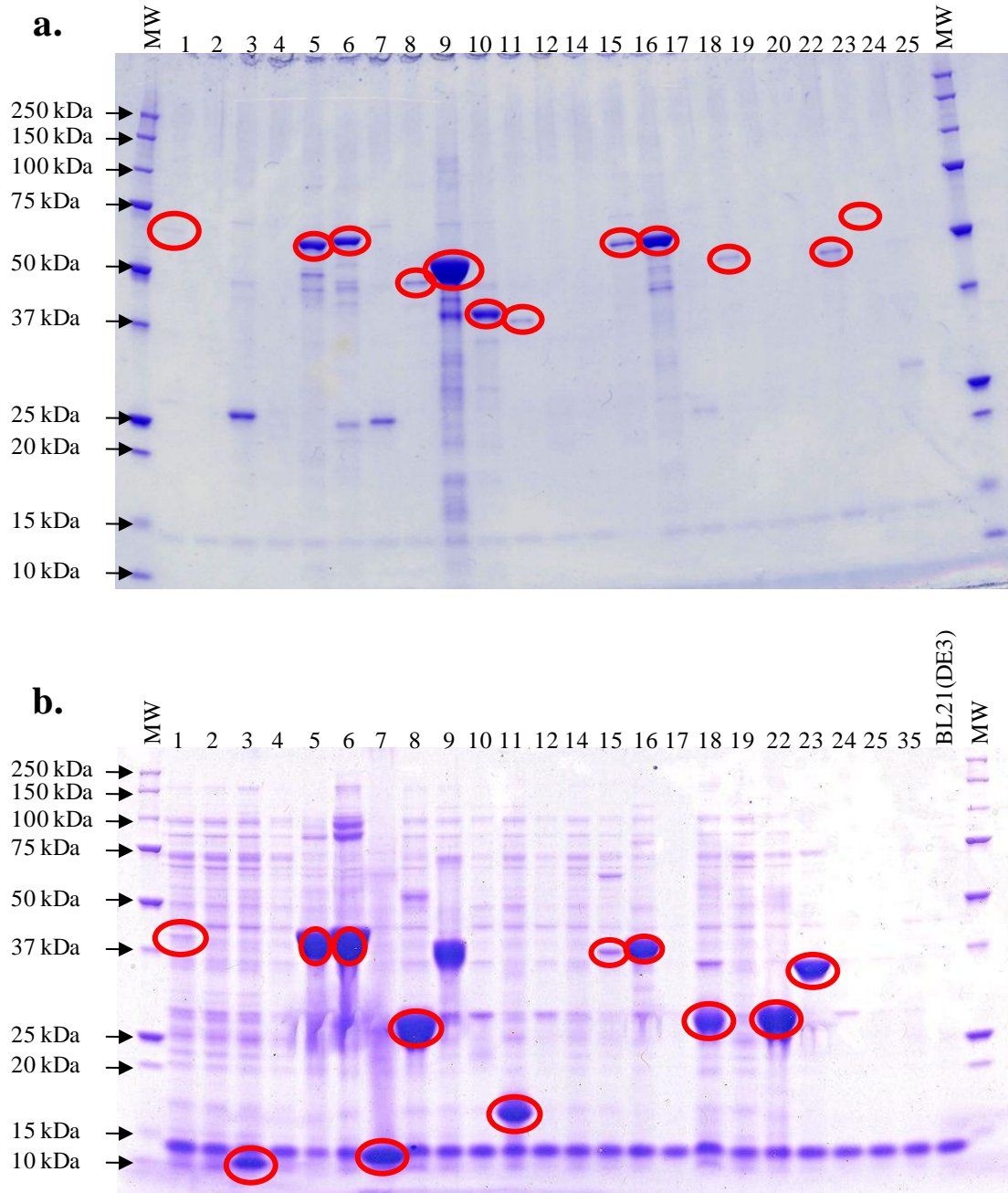
N°	Cible	Gène	Stratégie	PCR	Vecteur pGEX-4-T1		Vecteur pFO4	
					Ligation et transformation en souche DH5α	Expression en souche BL21	Ligation et transformation en souche DH5α	Expression en souche BL21(DE3)
1	AlyA7	zg4327	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
2	AlyA7 domain 1	zg4327	BamHI/EcoRI	+	+	-	+	soluble
3	AlyA7 domain 2	zg4327	BamHI/EcoRI	+	+	-	+	soluble/purifié
4	PL15	zg3452	BamHI/EcoRI	+	+	-	+	soluble
5	UglA1	zg3445	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
6	UglA2	zg3804	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
7	KdgF-like	zg2623	BamHI/EcoRI	+	+	soluble	+	soluble/purifié
8	Sdr	zg2622	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
9	zg2619	zg2619	BamHI/EcoRI	+	+	soluble/purifié	+	soluble
10	zg2619 domain 1	zg2619	BamHI/EcoRI	+	+	soluble/purifié	+	-
11	zg2619 domain 2	zg2619	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
12	Zg 2615	zg2615	BamHI/EcoRI	+	+	-	+	-
13	AlyA4	zg4130	BamHI/EcoRI	-	NA	NA	-	NA
14	AlyA4 domain 1	zg4130	BamHI/EcoRI	+	+	insoluble	+	soluble
15	AlyA4 domain 2	zg4130	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
16	AlyA5	zg4131	BamHI/EcoRI	+	+	soluble/purifié	+	soluble/purifié
17	AlyA2	zg2618	BamHI/MfeI	+	+	-	+	-
18	GntR-like	zg2617	BamHI/MfeI	+	+	soluble/purifié	+	soluble/purifié
19	Pfp	zg2613	BamHI/MfeI	+	+	-	+	-
20	AlyA1	zg1182	BglII/EcoRI	+	+	-	-	NA
21	AlyA1 domain 1	zg1182	BglII/EcoRI	-	NA	NA	-	NA
22	AlyA1 domain 2	zg1182	BglII/EcoRI	+	+	soluble/purifié	+	soluble/purifié
23	KdgK1	zg2614	BglII/EcoRI	+	+	soluble/purifié	+	soluble/purifié
24	SusD-like	zg2620	BglII/MfeI	+	+	-	+	-
25	SusD domain 1	zg2620	BglII/MfeI	+	+	-	+	-
26	AlyA3	zg2624	BamHI/XhoI	-	NA	NA	-	NA
27	AlyA3 domain 1	zg2624	BamHI/XhoI	+	-	NA	-	NA
28	AlyA3 domain 2	zg2624	BamHI/XhoI	+	-	NA	-	NA
29	AlyA3 domain 3	zg2624	BamHI/XhoI	+	-	NA	-	NA
30	AlyA6	zg4132	BamHI/XhoI	+	-	NA	-	NA
31	AlyA6 domain 1	zg4132	BamHI/XhoI	+	-	NA	-	NA
32	AlyA6 domain 2	zg4132	BamHI/XhoI	-	NA	NA	-	NA
33	SusC-like	zg2621	BamHI/BglII	+	-	NA	-	NA
34	zg2616	zg2616	BamHI/BglII	+	-	NA	-	NA
35	NRAMP	zg2612	BamHI/BglII	+	-	NA	+	-
	succès	35		31	23	14	23	16

**Tableau 6-5 : Résultats des différentes étapes du clonage et de l'expression à moyen débit**

NA: non applicable car résultat négatif à l'étape précédente

Les plasmides pGEX-4-T1 et pFO4 validés ont ensuite été transformés dans les souches d'expression BL21 et BL21(DE3), respectivement. Tous les plasmides obtenus aux étapes précédentes ont pu être transformés avec succès dans les souches d'expression, montrant que cette étape n'est pas limitante. Les tests d'expression en moyen débit ont été réalisés en milieu auto-inductible ZYP5052. L'utilisation de ce milieu, qui ne nécessite aucun ajout d'inducteur pendant la culture, permet de s'affranchir d'un suivi de croissance fastidieux à mettre en œuvre pour tous les clones en parallèle (Studier 2005). D'autre part, il permet d'atteindre des densités cellulaires élevées ( $A_{600\text{ nm}} > 10$ ), ce qui est nécessaire pour la détection de l'expression recombinante sur des petits volumes de culture. L'expression et la solubilité des protéines ont été évaluées par SDS-PAGE après lyse cellulaire. L'utilisation de mini-colonnes de purification a permis de détecter plus facilement la présence de protéines recombinantes dans les extraits solubles. Les résultats d'électrophorèse SDS-PAGE sur les fractions purifiées sont donnés sur la Figure 6-5.

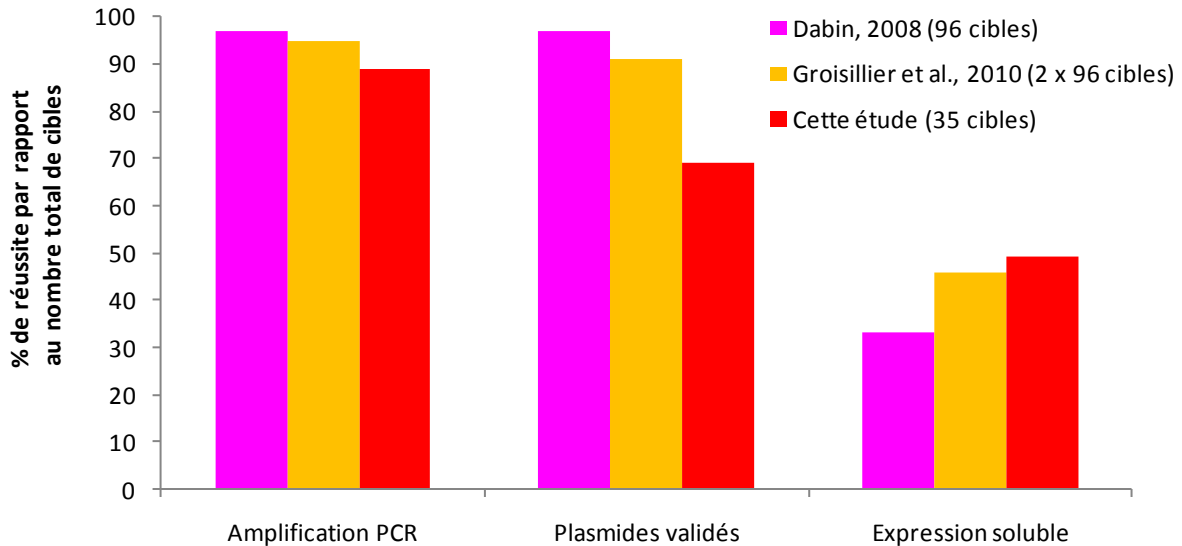
Ces tests d'expression ont révélé des taux de réussite similaires pour les deux vecteurs. 14 cibles ont pu être exprimées dans le vecteur pGEX-4-T1, et 16 cibles dans le vecteur pFO4 (Tableau 6-5). Seul le domaine 1 de la protéine AlyA4 (n°14) fusionné à la GST a été exprimé sous forme insoluble. Cependant ce même domaine cloné dans pFO4 a été exprimé de façon soluble, confirmant l'utilité de l'utilisation des deux vecteurs en parallèle pour maximiser les chances de succès. De plus, trois cibles n'ont été exprimées que dans le vecteur pFO4 (n°2, 3 et 4), et une cible uniquement dans le vecteur pGEX-4-T1 (n°10). Au total, 17 des 35 cibles initiales ont été exprimées de façon soluble dans l'un ou l'autre des vecteurs.



**Figure 6-5 : Analyse en SDS-PAGE des fractions purifiées des extraits cellulaires d'*E. coli*.** Les cercles rouges montrent la présence d'une bande à la taille attendue pour une cible donnée. a) Expression des vecteurs pGEX-4-T1 recombinants en cellule BL21. b) Expression des vecteurs pFO4 en cellules BL21(DE3).

### III-3 Discussion

Cette stratégie de clonage à moyen débit a déjà été mise en œuvre auparavant au laboratoire (Dabin 2008; Groisillier *et al.*, 2010). La comparaison des taux de réussite à trois étapes clés est donnée dans la Figure 6-6.



**Figure 6-6 : Comparaison des taux de réussite dans les trois stratégies de clonage et surexpression à moyen débit**

On constate que les taux de réussite suivent la même tendance dans les trois études. Dans le travail présenté ici, le pourcentage de plasmides validés est moindre. Ceci est à relier à l'échec de la stratégie de clonage BamHI / XhoI déjà commenté plus haut. Malgré ce handicap, le taux d'expression soluble obtenu ici est le plus élevé des trois (49%). La plupart des vecteurs recombinants validés ont donc été bien tolérés et exprimés chez *E. coli*. Ceci peut s'expliquer par les différents organismes d'origine des cibles à cloner. Dans la première étude (Dabin 2008), les 96 cibles provenaient toutes de la bactérie *Rhodospirellula baltica*, un membre du phylum Planctomycetes connu pour ses particularités comme la présence d'une structure protéique remplaçant le peptidoglycane ou de véritables compartiments cellulaires (Fuerst 1995). Les 192 cibles de la deuxième étude (Groisillier *et al.*, 2010) provenaient d'organismes très divers, eucaryotes (oursins, algues brunes et rouges, poissons, micro-algues), procaryotes (bactéries et archées) ou virus. Dans le travail présenté ici, les gènes cibles étaient tous issus de *Z. galactanivorans*, une bactérie Gram-négative classique comme *E.coli*. Les protéines recombinantes étaient donc probablement plus faciles à prendre en

charge par les cellules d'expression. Le taux d'expression soluble de 49% obtenu ici concorde avec les résultats des projets de production de protéines recombinantes à haut débit, où il a été montré que les taux de succès pour les protéines procaryotes avoisinent les 50% (Braun and LaBaer 2003; Fox *et al.*, 2008; Koehn and Hunt 2009).

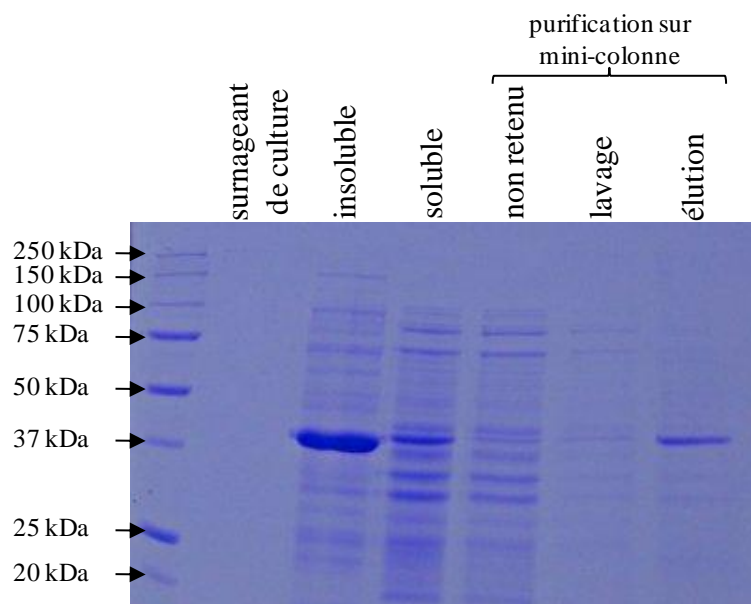
Sept cibles clonées avec succès dans un ou deux vecteur(s) n'ont pas été exprimées de manière hétérologue chez les souches d'*E. coli* utilisées. Plusieurs solutions pourraient être envisagées pour tenter de les récupérer.

Tout d'abord, d'autres souches d'expression que les cellules BL21 ou BL21(DE3) pourraient être utilisées. Par exemple, les souches C41(DE3) et C43(DE3) sont connues pour mieux tolérer la toxicité due à l'expression de protéines recombinantes, tout en atteignant des densités cellulaires élevées. Ces deux souches ont été originellement mises au point pour l'expression de protéines membranaires ou globulaires (Miroux and Walker 1996). Elles ont depuis été utilisées pour la production de tous types de protéines, avec des taux de réussite de l'ordre de 80%, alors qu'il est de 50% pour la souche BL21(DE3) (Dumon-Seignovert *et al.*, 2004). Ces souches semblent aussi mieux tolérer la transformation par les vecteurs d'expression, avec 100% de réussite contre 62% pour BL21(DE3) (Dumon-Seignovert *et al.*, 2004).

D'autre part, la méthode utilisée pour les tests d'expression à moyen débit peut sous-estimer le nombre de protéines solubles. En effet, les tests ont été réalisés sur des petits volumes de culture (3 mL), qui ne permettent pas forcément de détecter des niveaux d'expression faibles. Pour les cibles jugées négatives par cette approche, il pourrait être envisagé d'augmenter le volume de culture.

Le milieu ZYP5052 auto-inductible utilisé ici facilite le traitement d'un grand nombre de cibles en parallèle. Cependant, d'autres milieux avec une induction "classique" pourraient être utilisés pour augmenter la synthèse des protéines faiblement exprimées. A titre d'exemple, la protéine recombinante AlyA7 clonée dans pFO4 (Tableau 6-5, n°1) a un niveau d'expression faible en milieu ZYP5052, comme dénoté par la bande de faible intensité obtenue sur le gel SDS-PAGE (Figure 6-5). Le même clone a été cultivé en milieu Graffinity, où l'expression est induite par l'ajout simultané de lactose et d'IPTG (Roussel 2009). Les protéines ont ensuite été extraites à partir de 1 mL de culture, purifiées sur mini-colonne d'affinité au nickel et déposées sur gel SDS-PAGE (Figure 6-7). Alors que le volume de culture Graffinity utilisé est plus faible que celui du milieu ZYP5052, on constate que la bande obtenue pour la fraction purifiée est bien plus intense, dénotant une meilleure surexpression de la protéine

recombinante dans ces conditions (à comparer avec la première piste du gel de la Figure 6-5 B). L'utilisation du milieu Graffinity pourrait donc permettre de récupérer individuellement certaines cibles jugées négatives, et d'augmenter les niveaux de production pour celles faiblement exprimées.



**Figure 6-7 : Analyse en SDS-PAGE des extraits cellulaires d'*E. coli* exprimant la protéine AlyA7 en milieu Graffinity**

#### IV- Bilan et perspectives

L'approche de clonage et de surexpression à moyen débit menée sur 35 cibles potentiellement impliquées dans la dégradation de l'alginate a abouti à l'expression soluble de 17 protéines recombinantes. L'intérêt de traiter un grand nombre de cibles en parallèle est de maximiser les chances de succès. Cela impose ensuite de faire des choix sur les cibles à produire à plus grande échelle en vue d'une caractérisation biochimique et structurale. Dans l'étude présentée ici, les 17 protéines recombinantes sont :

- des alginate lyases de trois familles CAZy différentes (Cantarel *et al.*, 2009)
  - Famille PL6 : AlyA4 module 1 (PL6) et module 2 (fonction inconnue)
  - Famille PL7 : AlyA1 module 2 et AlyA5 entière
  - Famille PL14 : AlyA7 entière, domaine 1 et domaine 2

- des enzymes susceptibles d'intervenir en aval des lyases pour la dégradation de l'alginate (cf. Chapitre 4)
  - famille PL15 : zg3452 (oligo-alginate lyase ?)
  - famille GH88 : unsaturated glucuronyl hydrolases UglA1 et UglA2
  - Déshydrogénase : zg2622 (gluconate dehydrogenase ?)
  - Cétodésoxygluconate kinase : KdgK1
  
- des protéines non-enzymatiques
  - Facteur de transcription : *GntR-like*
  - zg2619           entière
    - module 1 : domaine PKD, présent sur les parties extracellulaires de protéines impliquées dans les interactions avec d'autres protéines ou des polysaccharides
    - module 2 : Carbohydrate Binding Module (CBM) putatif

L'objectif de cette étape de surexpression à moyen débit était d'obtenir des protéines recombinantes permettant de valider les fonctions prédites par l'annotation du génome et leur implication dans la dégradation de l'alginate par *Z. galactanivorans*. Les cibles prioritaires pour une production à grande échelle et une caractérisation biochimique et structurale sont donc les alginate lyases. Le Chapitre 7 traite de l'étude des deux enzymes de la famille PL7 AlyA1 et AlyA5. Par ailleurs, la protéine AlyA7 (PL14) est en cours de production et d'analyse biochimique. Les résultats devraient permettre de mieux comprendre le rôle particulier qui a été proposé pour cette alginate lyase dans le modèle de régulation (cf. Chapitre 4). L'étude du module PL6 de l'alginate lyase AlyA4 serait également très intéressante, notamment du point de vue de la biologie structurale. En effet, la seule structure connue à ce jour dans la famille PL6 est celle d'une chondroitinase B de *Pedobacter heparinus* (Huang *et al.*, 1999; Michel *et al.*, 2004). La caractérisation d'AlyA4 pourrait à ce titre éclairer sur les déterminants structuraux de la spécificité de substrat au sein de cette famille. Le facteur de transcription *GntR-like* (présent dans l'opéron alginolytique) a également été obtenu avec un fort niveau de surexpression dans l'approche à moyen débit. La caractérisation de ce régulateur potentiel permettrait de clarifier son implication dans la capacité de *Z. galactanivorans* à détecter la présence d'alginate dans le milieu (cf. Chapitre 4). Dans une première approche, il pourrait être envisagé des expériences de microcalorimétrie

afin de déterminer des effecteurs potentiels tels que des oligo-alginates. De plus, des sites de fixation sur l'ADN en amont des gènes impliqués dans la dégradation de l'alginate pourraient être recherchés, par des expériences d'empreinte à la DNase ou de retard sur gel, par exemple. Enfin, l'obtention d'un module CBM putatif de manière soluble est très encourageante. S'il s'avère réellement reconnaître spécifiquement l'alginate, ce domaine pourrait être utilisé comme sonde pour détecter la distribution de ce composé pariétal directement sur des coupes d'algues brunes, en s'inspirant de techniques utilisées pour les plantes terrestres (Hervé *et al.*, 2011). La production et la caractérisation de ce CBM putatif sont en cours au sein du laboratoire.

Enfin, la même méthode de clonage et surexpression à moyen débit pourrait être appliquée aux nouvelles cibles trouvées par l'analyse transcriptomique par puces à ADN (cf. Chapitre 5), afin de valider leur implication dans la dégradation de l'alginate.





## Chapitre 7

# Caractérisation phylogénétique, biochimique et cristallographique d'alginate lyases de la famille PL7

---

## I- Présentation

Au cours de l'annotation du génome, trois alginate lyases de la famille PL7 ont été identifiées chez *Z. galactanivorans*. Il s'agit d'AlyA1 (zg1182), AlyA2 (zg2618) et AlyA5 (zg4131) (cf. Chapitre 4). Les gènes codant AlyA2 et AlyA5 font partie de 2 opérons alginolytiques, alors qu'*alyA1* ne semble pas appartenir à un cluster en particulier (cf. Chapitre 4). L'expression de ces gènes est fortement induite en présence d'alginate dans le milieu. Les analyses de la modularité des protéines ont montré que AlyA2 et AlyA5 ne comprennent qu'un seul domaine (Tableau 7-1). En plus de son module catalytique C-terminal, AlyA1 contient un module auxiliaire N-terminal de fixation de sucres (*carbohydrate binding module*, CBM) prédit comme appartenant à la famille CBM32.

Enzyme	Masse moléculaire	pI théorique	Modularité
AlyA1	48,1 kDa	4,7	CBM32 + PL7
AlyA2	35,1 kDa	5,5	PL7
AlyA5	39,3 kDa	5,2	PL7

**Tableau 7-1 : Propriétés des trois alginate lyases PL7 de *Z. galactanivorans***

Afin de progresser dans la compréhension du système alginolytique de *Z. galactanivorans*, il est nécessaire de (i) confirmer les fonctions inférées lors de l'annotation et (ii) caractériser finement les mécanismes enzymatiques impliqués dans la dégradation du polysaccharide. Dans ce chapitre, je présente une étude phylogénétique des alginate lyases de la famille PL7 et la caractérisation biochimique de deux d'entre elles. Par ailleurs, une approche de biologie structurale a été initiée afin d'identifier les déterminants structuraux de la spécificité et du mode d'action des enzymes.

A ce jour, la famille PL7 regroupe 94 enzymes, dont une majorité d'origine bactérienne (<http://www.cazy.org/>, Cantarel *et al.*, 2009). Seulement 16 d'entre elles ont été caractérisées biochimiquement, et ce sont toutes des alginate lyases (EC 4.2.2.3 ou 4.2.2.11). Les trois structures disponibles montrent un repliement en sandwich  $\beta$  incurvé ( $\beta$  *jelly roll*). L'étude présentée dans ce chapitre permettra d'augmenter les connaissances sur cette famille d'enzymes.

## II- Matériels et Méthodes

### II-1 Substrats

Trois lots d'alginate de sodium polymères possédant différents ratios M/G (0,5; 0,9 et 2,0) ont été fournis par la compagnie Danisco dans le cadre du projet européen PolyModE (KBB7-2007-2A).

Les oligoalginates saturés ont été préparés par hydrolyse acide d'alginate de sodium selon le protocole décrit par Haug *et al.* (1974b). Après fractionnement, trois échantillons contenant des oligogulonates, oligomannuronates et oligosaccharides mixtes MG ont été obtenus. Cette méthode a été rapportée pour fournir des fragments avec des degrés de polymérisation (DP) allant de 4 à environ 30 (Shimokawa *et al.*, 1997). Leur ratio M/G a été vérifié par RMN du proton (polyG : 0,009; polyM : 6,064; polyMG : 0,419).

### II-2 Analyses de séquences et phylogénie

Les alignements de séquences multiples ont été réalisés en utilisant MAFFT avec l'algorithme L-INS-i et la matrice de substitution Blosom62 (Kato and Toh 2008) et formatés en utilisant le programme ESPript (Gouet *et al.*, 2003). Pour les analyses phylogénétiques de membres de la famille PL7, les séquences ont été sélectionnées dans la base de données CAZy (<http://www.cazy.org/>, Cantarel *et al.*, 2009) et alignées en utilisant MAFFT. L'alignement a été édité manuellement dans MEGA 4.1 (Tamura *et al.*, 2007). 180 sites ont été utilisés pour calculer un arbre phylogénétique en utilisant une méthode de maximum de vraisemblance conduite dans le programme PhyML (Guindon and Gascuel 2003) implémenté en ligne sur Phylogeny.fr (Dereeper *et al.*, 2008). Les valeurs de bootstrap ont été calculées à partir de 100 ré-échantillonnages du jeu de données.

### II-3 Clonage des gènes d'alginate lyases

Les régions codantes correspondant au module catalytique d'AlyA1 (zg1182) et aux protéines complètes matures d'AlyA2 et AlyA5 (zg2618 et zg4131, respectivement) ont été clonées dans le vecteur pFO4 comme décrit dans le Chapitre 6. Les inserts ont été séquencés avant l'expression selon le protocole suivant. Des réactions d'amplification par PCR ont été

conduites sur les plasmides recombinants en utilisant des nucléotides fluorescents (Big Dye Terminator BDT V3.1, Applied Biosystems, Foster City, CA, USA). Les mélanges réactionnels (5  $\mu$ l) étaient composés de 0,5  $\mu$ l de BDT, 0,75  $\mu$ l de tampon (400 mM Tris-HCl pH 9, 10 mM MgCl<sub>2</sub>), 2  $\mu$ l de plasmides (30 ng. $\mu$ l<sup>-1</sup>), 1  $\mu$ l d'amorce sens ou anti-sens (5  $\mu$ M, pFO4\_fw: GCAGCAGCCACCATCACCATCACC, pFO4\_rv: CCTTTCGGGCTTTGTTAGCAGCCGG) et 0,75  $\mu$ l d'eau *nuclease-free*. Le programme de PCR a consisté en une dénaturation initiale à 96°C pendant 5 min suivie de 50 cycles de 30 s à 96°C, 30 s à 55°C et 4 min à 60°C. Les produits d'amplification ont alors été purifiés pour éliminer les nucléotides fluorescents non incorporés et séquencés sur un appareil ABI Prism3100 (Applied Biosystems). Les séquences obtenues ont été comparées au génome de *Z. galactanivorans* et ont permis la validation des plasmides recombinants.

#### II-4 Procédures d'expression et de purification des protéines

Des précultures de 2 ml des cellules *E. coli* BL21(DE3) recombinantes ont été incubées une nuit à 37°C et 180 rpm en milieu LB contenant de l'ampicilline (100  $\mu$ g.ml<sup>-1</sup>). 200 ml de milieu autoinductible ZYP avec ampicilline (Studier 2005) ont étéensemencés avec 200  $\mu$ l de préculture et incubés pendant 3 jours à 20°C sous agitation. Les cellules ont été centrifugées pendant 20 min à 5000 rpm, 4°C et les culots ont été stockés à -20°C. Les cellules ont été resuspendues dans 20 ml de Tampon A (25 mM Tris-HCl pH 7,5; 200 mM NaCl; 5 mM imidazole) contenant un cocktail d'antiprotéases (une demie pastille pour 100 ml; Complete EDTA-free, Roche) et une paillette de DNase de pancréas de bœuf (Sigma) et lysées à la presse de French. Les échantillons ont ensuite été centrifugés pendant 2 h à 20000 g. Le surnageant a été injecté sur une colonne d'affinité métallique Hyper Cell PAL chargée avec 0,1 M NiSO<sub>4</sub> et équilibrée avec le Tampon A. Les protéines ont été éluées avec un gradient linéaire entre le Tampon A et le Tampon B (25 mM Tris-HCl pH 7,5; 200 mM NaCl; 1 M imidazole) en 60 ml à un débit de 1 ml.min<sup>-1</sup> et collectées en fractions de 1 ml. Les fractions montrant la présence de la protéine recombinante sur gel d'électrophorèse sodium dodecyl sulfate – polyacrylamide (SDS-PAGE) ont été rassemblées. Des aliquots de 5 ml ont été purifiés par chromatographie d'exclusion de taille (Superdex 75 16/60 pg, GE Healthcare) en utilisant le Tampon C (25 mM Tris-HCl pH 7,5; 200 mM NaCl) à un débit de 1 ml.min<sup>-1</sup>. Les fractions montrant la présence de la protéine recombinante pure ont été rassemblées. Des mesures de diffusion dynamique de la lumière (*Dynamic Light Scattering*, DLS) ont été

effectuées sur 40  $\mu\text{l}$  des solutions protéiques concentrées ( $> 1 \text{ mg}\cdot\text{ml}^{-1}$ ) dans une cuvette en quartz sur un instrument ZetaSizer Nano-S (Malverne Instruments).

## II-5 Tests enzymatiques

La concentration en protéines a été déterminée en mesurant l'absorbance à 280 nm ( $A_{280}$ ) et en utilisant la loi de Beer-Lambert :  $A = \epsilon * l * c$ . Les coefficients d'extinction molaires théoriques  $\epsilon_{280}$  ont été calculés à partir des séquences protéiques sur le serveur ProtParam (<http://www.expasy.ch/tools/protparam.html>) et sont de  $47455 \text{ M}^{-1}\cdot\text{cm}^{-1}$  pour AlyA1 et  $72310 \text{ M}^{-1}\cdot\text{cm}^{-1}$  pour AlyA5. L'activité alginate lyase a été testée en mesurant l'augmentation d'absorbance à 235 nm ( $A_{235}$ ) des produits de la réaction (uronates insaturés) pendant 5 min dans une cuvette en quartz de 1 cm contenant 0,5 ml de mélange réactionnel dans un spectrophotomètre à enceinte thermostatée. Une unité d'activité a été définie comme une augmentation d'une unité  $A_{235}$  par minute.

Les effets de la température, de NaCl et d'agents chélatants sur l'activité d'AlyA1 ont été testés dans un tampon 100 mM Tris-HCl à pH 7,5 en utilisant l'alginate de sodium (0,05%) comme substrat et 5  $\mu\text{l}$  d'enzyme purifiée à 1,2  $\mu\text{M}$ . Les agents chélatants acide éthylène diamine tétracétique (EDTA) ou acide éthylène glycol tétracétique (EGTA) ont été soit incubés à la concentration finale de 1 mM avec AlyA1 (une nuit, 4°C) soit ajoutés directement dans le mélange réactionnel. L'effet du pH sur l'activité d'AlyA1 a été testé en utilisant différents tampons à 100 mM: citrate (pH 3,0 to 6,0), acide 2-morpholinoéthanesulfonique-NaOH (MOPS; pH 6,0 à 7,9), HEPES (pH 7,0 à 8,0), Tris-HCl (pH 7,0 à 8,5) et glycine-NaOH (pH 8,5 à 10,0). Les constantes cinétiques ont été déterminées à 30°C dans 100 mM Tris-HCl pH 7,5, 200 mM NaCl avec 10 concentrations d'alginate différentes allant de 0,075 à 2,5  $\text{mg}\cdot\text{ml}^{-1}$ . Les réactions ont été réalisées en triplicats. L'algorithme de régression hyperbolique du logiciel Hyper32 (<http://homepage.ntlworld.com/john.easterby/software.html>) a été utilisé pour estimer les constantes cinétiques.

Les tests enzymatiques avec AlyA5 sur les polysaccharides d'alginate ont été réalisés comme pour AlyA1, en utilisant 5  $\mu\text{l}$  d'enzyme purifiée à 2,7  $\mu\text{M}$ . Les réactions sur les oligosaccharides saturés ou insaturés ont été conduites à 30°C dans 100 mM Tris-HCl pH 7,5, 4 mM  $\text{CaCl}_2$ , 50  $\mu\text{g}$  du substrat et 5  $\mu\text{l}$  d'enzyme à 5,6  $\mu\text{M}$ .

## II-6 Chromatographie d'exclusion de taille des produits de dégradation par AlyA5

Les échantillons d'oligosaccharides (200  $\mu$ l) ont été bouillis 10 minutes pour arrêter la réaction enzymatique et injectés sur une colonne analytique Superdex peptide. L'élution a été réalisée en  $\text{LiNO}_3$  100 mM à 20°C en utilisant une pompe isocratique Dionex Ultimate 3000 travaillant à un débit de 0,5  $\text{ml}\cdot\text{min}^{-1}$ . L'élution a été suivie simultanément par des détecteurs de l'index de réfraction (RI) et de l'absorbance  $A_{235}$ .

## II-7 Purification d'oligo-alginates insaturés

Afin d'obtenir des oligosaccharides insaturés purs, une digestion partielle d'alginate a été mise en place. 4,5  $\mu$ l d'AlyA1 purifiée à 41 mM ont été ajoutés à 80 ml d'alginate 0,5% (M/G = 0,5) dans 100 mM Tris-HCl pH 7,5, 200 mM NaCl. Après 4 h à 30°C et 60 rpm, le milieu réactionnel a été ultrafiltré sur cellule Amicon sous pression de diazote (*cut-off* membrane 5 kDa). Le filtrat a été concentré à l'évaporateur rotatif jusqu'à 5 ml et injecté sur un système de chromatographie d'exclusion de taille constitué de 3 colonnes connectées en série (Superdex 30, Pharmacia). Les oligosaccharides ont été élués dans 50 mM carbonate d'ammonium à 20°C en utilisant une pompe isocratique Gilson 306 travaillant à 1,5  $\text{ml}\cdot\text{min}^{-1}$ . Des fractions de 10 ml ont été collectées. La purification a été suivie par la mesure du RI. Les oligosaccharides séparés ont été lyophilisés pour éliminer le carbonate d'ammonium et repris dans l'eau.

## II-8 Analyse des oligosaccharides par électrophorèse

Après dégradation enzymatique, les aliquots ont été bouillis pendant 10 min avant leur analyse. Les oligosaccharides ont été analysés par électrophorèse en gel de polyacrylamide et détectés par coloration (Carbohydrate-PAGE, C-PAGE) ou grâce à un fluorophore (Fluorescence-Assisted Carbohydrate Electrophoresis, FACE) (Zablackis and Perez 1990; Jackson 1996). Pour les C-PAGE, des échantillons de 5  $\mu$ l ont été mélangés avec 15  $\mu$ l de tampon de chargement (10% saccharose et 0,01% rouge phénol). 5  $\mu$ l ont été soumis à l'électrophorèse dans un gel de concentration à 6% de polyacrylamide (wt/vol) et un gel de séparation à 28% de polyacrylamide (épaisseur 0,75 mm) dans un tampon 50 mM Tris-HCl, 1 mM EDTA (pH 8,7). Les gels ont été colorés au bleu d'alcan 0,5% (wt/vol), suivi par du nitrate d'argent 0,4% (wt/vol) (Cowman *et al.*, 1984; Zablackis and Perez 1990). La coloration

a été développée avec du carbonate de sodium 7% (wt/vol) contenant du formaldéhyde. La réaction a été stoppée par ajout d'acide acétique glacial (17,5 M). Pour les FACE, les échantillons ont été séchés sous vide et dissous dans 2 µl d'une solution d'AMAC (2-aminoacridone; 0,15 M dans acide acétique 15% v/v) et 2 µl d'une solution de cyanoborohydrure de sodium (1 M en DMSO). Le mélange a été incubé pendant 16 h à 37°C et séché sous vide (Speedvac). Les culots d'oligosaccharides marqués ont été resuspendus dans du glycérol 25%. 5 µl ont été analysés par électrophorèse à 4°C dans un gel de concentration à 6% de polyacrylamide (wt/vol) en tampon 125 mM Tris-HCl, 50 mM glycine (pH 6,8) et un gel de séparation à 27% en tampon 380 mM Tris-HCl, 50 mM glycine (pH 8,8). Les oligosaccharides ont été visualisés dans un transilluminateur sous lumière ultraviolet.

## II-9 Réduction des oligosaccharides

Les oligosaccharides ont été réduits selon un protocole adapté d'Abdel-Akher *et al.* (1951). Brièvement, 3 mg de borohydrure de sodium ont été ajoutés à 500 µl d'oligogulonates (1 mg.ml<sup>-1</sup>) et le mélange a été incubé une nuit à température ambiante. La réaction a été acidifiée par ajout d'acide acétique concentré au goutte à goutte jusqu'à l'arrêt du dégagement gazeux de dihydrogène H<sub>2</sub>. L'échantillon a alors été séché sous vide (Speedvac) et resuspendu dans une goutte d'acide acétique glacial (17,5 M) et 1 ml de méthanol. Le solvant a été évaporé sous un léger courant de diazote pendant environ 30 min. La quantité d'extrémités réductrices a été estimée par la méthode au ferricyanure (Kidby and Davidson 1973) en utilisant une gamme de calibration de glucose de 20 à 400 µg.ml<sup>-1</sup>. Avant le traitement, le test au ferricyanure a mesuré 29,4 µg.ml<sup>-1</sup> eq. glucose dans la solution d'oligogulonates. Aucune extrémité réductrice n'a pu être détectée par la même méthode après la réduction.

## II-10 Suivis des réactions enzymatiques par spectrométrie RMN

Toutes les expériences de RMN et de spectrométrie de masse ont été réalisées en collaboration avec Lena Lundqvist et Pr. Corine Sandström du Département de Chimie de la *Sveriges lantbruksuniversitet* (SLU, Uppsala, Suède). Les expériences de RMN ont été conduites sur un spectromètre Bruker DRX-400 MHz avec une sonde QNP 5 mm <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P ou sur un spectromètre Bruker Avance III 600 MHz avec une sonde inverse QXI 5 mm <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N/<sup>31</sup>P, avec gradient Z. Les déplacements chimiques ont été calibrés en



utilisant l'acétone comme référence interne ( $\delta_H$  2,225;  $\delta_C$  31,05). Les expériences RMN  $^1H$  ont été réalisées à 30°C et 70°C, avec les paramètres suivants : fenêtre spectrale, 3000 Hz; temps d'acquisition, 5,62 s; temps de relaxation, 4 s; largeur de pulse, 8,5  $\mu$ s; 128 scans. Le signal résiduel de l'eau a été supprimé par présaturation NOESY (temps de relaxation, 4 s; mixing time, 50 ms).

La dégradation par les enzymes AlyA1 et AlyA5 a été suivie directement dans les tubes RMN. Quatre substrats différents ont été testés à la concentration de 5 mg.ml<sup>-1</sup> : alginate de sodium, blocs G, blocs M et blocs alternés MG. Les substrats ont été dissous en tampon phosphate 100 mM deutéré, pH 7,6, avec 200 mM NaCl et 5 mM CaCl<sub>2</sub>. Un spectre de référence a été acquis sur les substrats avant la réaction enzymatique. Les enzymes ont ensuite été ajoutées à la concentration finale de 7  $\mu$ g.ml<sup>-1</sup> dans le tube RMN. Cette concentration est trop faible pour détecter des signaux protéiques en RMN. La dégradation a été suivie à 30°C en acquérant des spectres à intervalle de temps régulier.

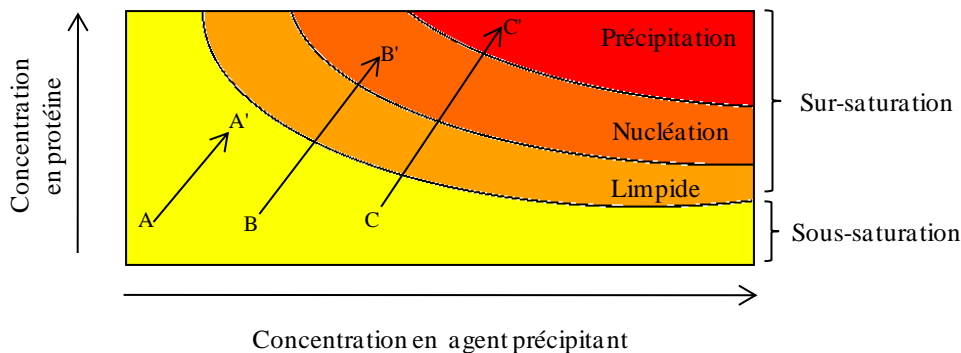
## **II-11 Analyse des produits finaux de réaction par RMN et spectrométrie de masse**

Les dégradations enzymatiques ont été réalisées à 30°C pendant 18 h sur 10 ml d'alginate de sodium à 5 mg.ml<sup>-1</sup>. Les échantillons ont ensuite été centrifugés, le surnageant lyophilisé et dissous dans de l'acétate d'ammonium 0,1 M. Les oligosaccharides ont été fractionnés par chromatographie d'exclusion de taille sur une colonne Superdex 30 16\*60 (Amersham Biosciences) sur un système ÄKTA. Les échantillons ont été élués par l'ammonium acétate 0,1 M à température ambiante, avec un débit de 0,8 ml.min<sup>-1</sup>. Les oligosaccharides ont été détectés par l'absorbance à 235 nm, et les quantités relatives des différents produits de réaction ont été estimées par l'aire des pics. Les fractions contenant les produits séparés ont été rassemblées, lyophilisées et repassées une seconde fois sur colonne d'exclusion de taille pour améliorer leur pureté. Les fractions purifiées, lyophilisées et reprises dans 1:1 méthanol-eau ont été analysées par ESI-MS sur un spectromètre de masse Bruker Esquire-LC avec un débit de 2  $\mu$ l.min<sup>-1</sup>. La gamme de masse scannée était de 100 à 1100 unités de masse atomique. Toutes les acquisitions ont été réalisées dans les modes négatif et positif et traitées dans le logiciel Bruker Daltonics esquire LC 4.5. La structure des oligosaccharides purifiés a été analysée par RMN TOCSY une et deux dimensions comme décrit précédemment.

## II-12 Cristallogenèse

Les protéines ont été concentrées en fin de purification jusqu'à leur limite de solubilité dans le tampon utilisé pour la chromatographie d'exclusion de taille (Tampon C).

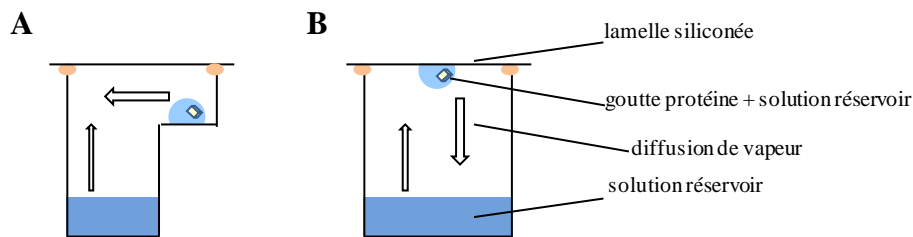
Les techniques de cristallogenèse utilisées sont basées sur la diffusion de vapeur. On permet à une goutte contenant la protéine purifiée, et une solution de cristallisation (composée d'agent précipitant, d'additifs et éventuellement d'un tampon) d'établir un équilibre avec un réservoir beaucoup plus grand qui contient la même solution de cristallisation en plus forte concentration, sans la protéine. Au départ, la goutte de protéine contient une concentration insuffisante en agent précipitant pour permettre la cristallisation. En général, le gradient de diffusion de vapeur qui s'établit aboutit à une perte d'eau pour la goutte, et donc à une concentration de la protéine et de l'agent précipitant. C'est au cours de ce lent processus que des événements de nucléation de cristaux peuvent se produire (Figure 7-1). Il faut pour cela que la protéine se trouve dans un tampon adéquat, favorisant notamment les interactions intermoléculaires.



**Figure 7-1: Diagramme de phase du principe de la croissance cristalline**

Diagramme de phase présentant les conditions nécessaires à la cristallisation par diffusion de vapeur. Cas A-A' la protéine reste soluble. Cas B-B' : Il y a nucléation et croissance cristalline. Cas C-C' : La protéine précipite.

Dans un premier temps, un grand nombre de conditions de cristallisation a été testé en parallèle par la technique de la goutte assise en utilisant des plaques 96 puits remplies par un robot HoneyBee 963 (Cartesian). Deux cribles commerciaux ont été testés, PACT et JCSG+ (Qiagen). Les meilleures conditions de cristallisation obtenues par ces cribles ont ensuite été optimisées manuellement en plaque 24 puits par la technique de la goutte suspendue (Figure 7-2). Le but est d'obtenir des monocristaux suffisamment gros et avec une bonne qualité de diffraction des rayons X.



**Figure 7-2 : Techniques de cristallogénèse par goutte assise (A) ou suspendue (B)**

Pour AlyA1 recombinante, suite à l'optimisation, les meilleurs cristaux ont été obtenus en ajoutant des oligogulonates ( $0,1 \text{ mg.ml}^{-1}$ ) à la protéine ( $11,3 \text{ mg.ml}^{-1}$ ) avant la cristallogénèse. Des gouttes composées de  $2 \mu\text{l}$  ou  $3 \mu\text{l}$  de ce mélange et  $1 \mu\text{l}$  ou  $1,5 \mu\text{l}$  (respectivement) de solution de cristallisation ( $0,2 \text{ M KSCN}$ ,  $28\% \text{ PEG-MME 2000}$ ) ont été suspendues sur un réservoir de  $500 \mu\text{l}$ . Les cristaux choisis pour les expériences de diffraction des rayons X ont été trempés dans une solution de cryo-protectant composée de la solution de cristallisation avec addition de  $10\%$  de glycérol. Ils ont été montés en boucle et congelés dans l'azote liquide.

Pour AlyA5 dimère, les meilleurs cristaux ont été obtenus en mélangeant  $2 \mu\text{l}$  de protéine purifiée ( $7,72 \text{ mg.ml}^{-1}$ ) avec  $1 \mu\text{l}$  de solution de cristallisation ( $22\% \text{ PEG 3350}$ ;  $0,2 \text{ M}$  tartrate de sodium/potassium) au-dessus d'un réservoir de  $200 \mu\text{l}$ . La solution de cryo-protectant était également composée de la solution de cristallisation additionnée de  $10\%$  de glycérol.

## II-13 Cristallographie et résolution des structures

Les données de diffraction ont été enregistrées sur la ligne Proxima 1 du Synchrotron SOLEIL. L'indexation et l'intégration des taches de diffraction ont été réalisées dans XDS (Kabsch 2010). Les données ont été mises à l'échelle en utilisant XSCALE, puis converties dans XDSCONV (Kabsch 2010). L'existence de structures pour des protéines homologues a permis d'envisager une résolution par remplacement moléculaire. Le logiciel de phasage Phaser a été utilisé (Zwart *et al.*, 2008), avec les coordonnées atomiques des alginate lyases ALY-1 de *Corynebacterium* sp. (PDB 1UAI) et A1-II' de *Sphingomonas* sp. A1 (PDB 2Z42) comme modèles. Les chaînes protéiques ont été construites automatiquement dans la carte de densité électronique en utilisant le programme ARP-WARP (Perrakis *et al.*, 1997), implémenté dans la suite logicielle CCP4i. Les solutions initiales obtenues par remplacement moléculaire ont été par la suite affinées avec le programme REFMAC5 dans CCP4i

(Murshudov *et al.*, 1997), en alternant avec des cycles de reconstruction manuelle dans COOT (Emsley and Cowtan 2004). Les molécules d'eau ont été recherchées dans ARP-WARP Solvent (Perrakis *et al.*, 1997). En fin d'affinement, le modèle a été amélioré en prenant en compte des mouvements globaux anisotropes grâce à l'application d'une matrice TLS (*Translational, Librational and Skrew tensors*) et des différences globales des positions entre les deux molécules reliées par symétrie non-cristallographique (NCS, *Non-Crystallographic Symmetry*). Ainsi, un cycle incluant des déplacements anisotropes TLS et l'application de contraintes NCS a été réalisé dans REFMAC5. Les facteurs de température (B-factor) ont été calculés par le programme Baverage implémenté dans CCP4i. Toutes les figures ont été réalisées dans le logiciel PyMOL. L'analyse des interfaces protéine-protéine a été réalisée sur le serveur PISA disponible en ligne (Krissinel and Henrick 2007; [http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)).

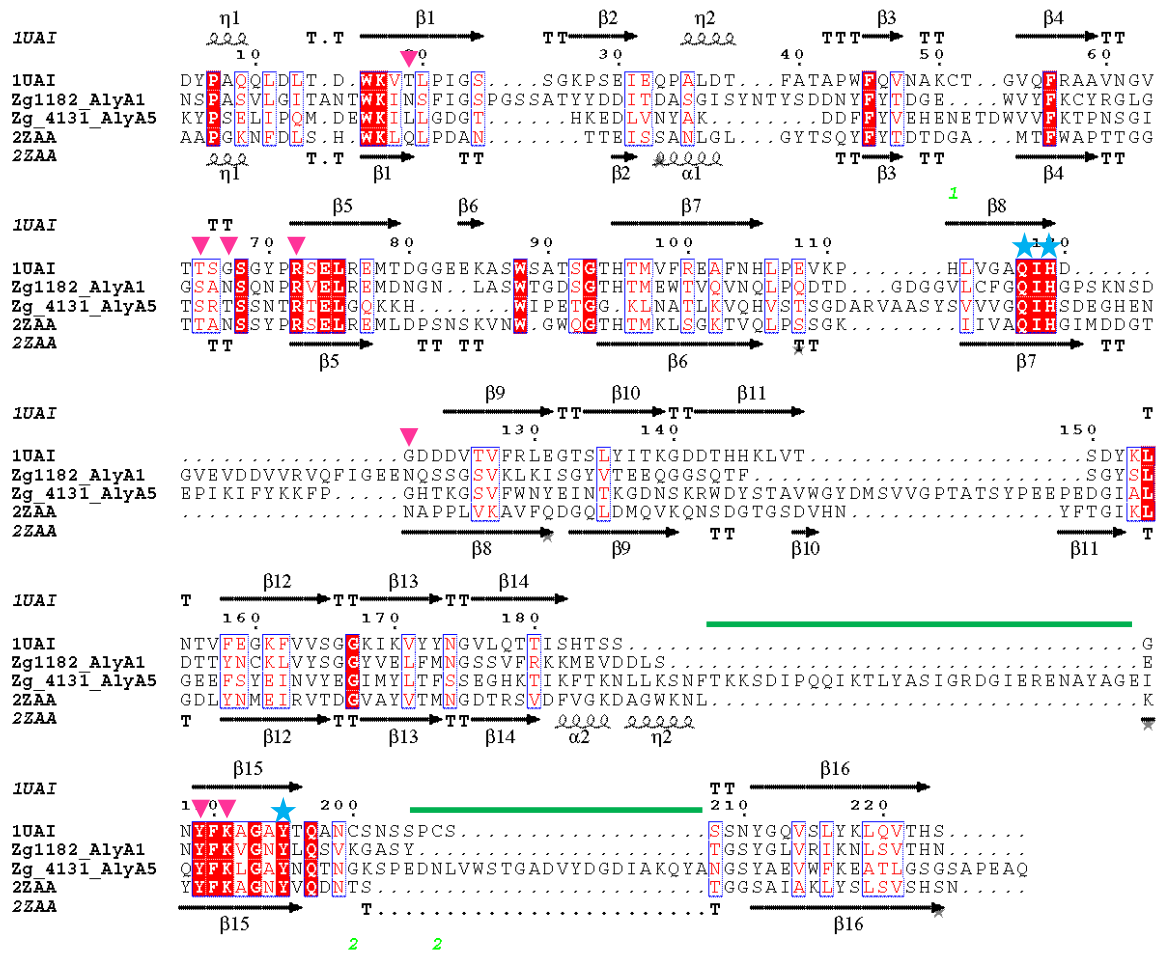
### III- Résultats et Discussion

#### III-1 Analyses de séquences et phylogénie

Les séquences d'acides aminés des trois alginate lyases PL7 AlyA1 (seulement le module catalytique), AlyA2 et AlyA5 ont été alignées avec celles d'AlyPG de *Corynebacterium* sp. ALY-1 et A1-II' de *Sphingomonas* sp. A1, deux autres membres de la famille PL7 (Matsubara *et al.*, 1998; Ogura *et al.*, 2008). Les identités de séquences sont données dans le Tableau 7-2. Les résidus catalytiques glutamine, histidine et tyrosine (**QIH** dans le feuillet  $\beta$ 7 et **YFK** dans  $\beta$ 15) sont conservés dans les 5 séquences (Figure 7-3). De manière intéressante, les résidus impliqués dans la fixation de l'alginate dans le site actif de l'enzyme A1-II' sont moins conservés. Ceci pourrait refléter des spécificités de substrats différentes. Osawa *et al.* (2005) ont précédemment proposé qu'une arginine dans le feuillet  $\beta$ 5 (RSEL**R**EM) n'est conservée que dans les lyases G-spécifiques, alors qu'un glutamate (RSEL**L**REM) est commun entre les enzymes G- et GM-spécifiques. AlyA1 et AlyA2 de *Z. galactanivorans* possèdent ces deux résidus. Cependant, l'arginine est remplacée par une glycine dans AlyA5. Cette comparaison de séquences révèle aussi qu'AlyA5 possède deux insertions de 31 et 22 acides aminés respectivement avant et après la tyrosine catalytique (Figure 7-3).

	zg1182 AlyA1	zg4131 AlyA5	zg2618 AlyA2	Q9RB42 alyPG <i>Corynebacterium</i>
Q75WP3 A1-II' <i>Sphingomonas</i>	25,8%	20,6%	28,0%	24,7%
zg1182 AlyA1		20,2%	22,4%	29,0%
zg4131 AlyA5			19,6%	15,8%
zg2618 AlyA2				23,7%

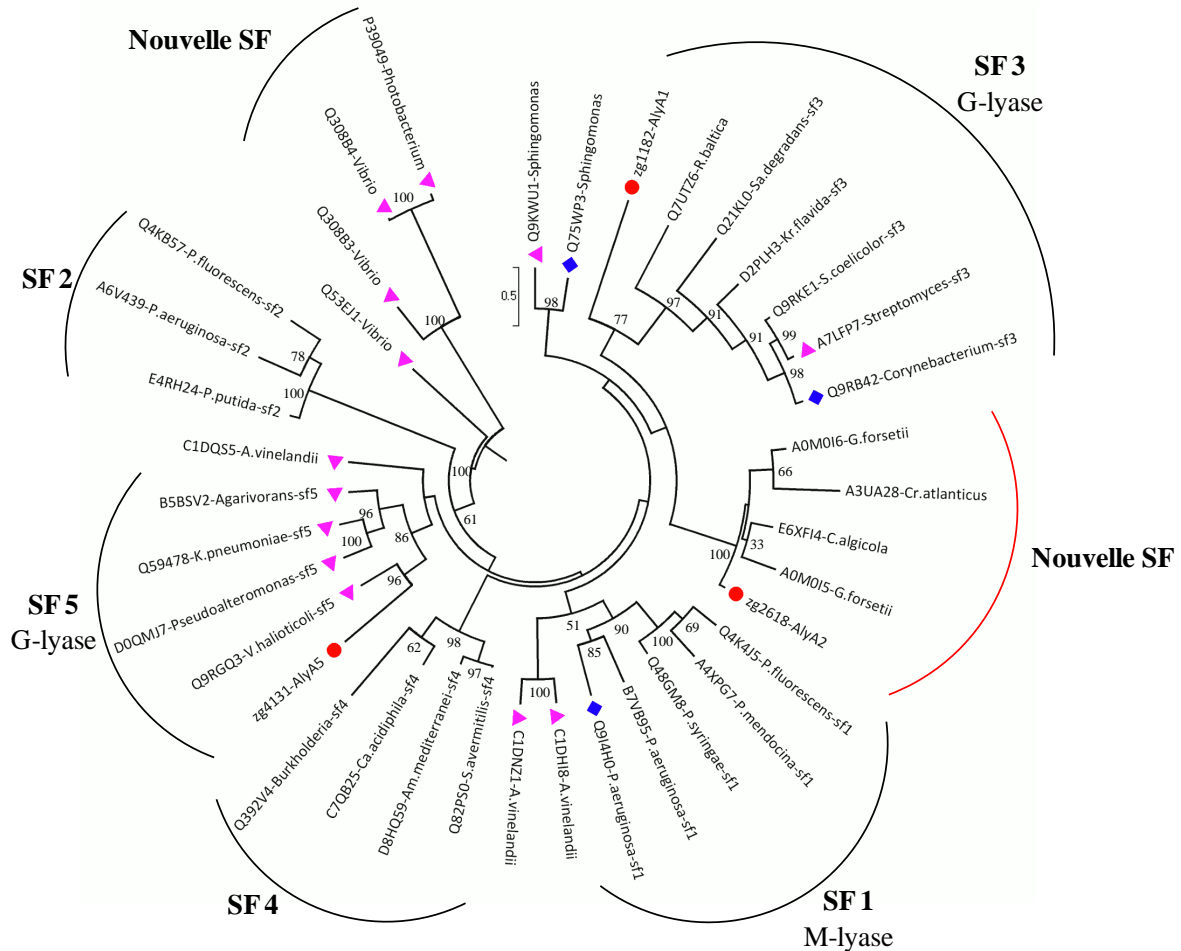
Tableau 7-2 : Pourcentages d'identité de séquences entre les PL7 alignées



**Figure 7-3 : Alignement de séquences multiples des alginate lyases PL7 de *Z. galactanivorans* avec AlyA1-II' (Q75WP3) de *Sphingomonas* sp. A1 et alyPG (Q9RB42) de *Corynebacterium* sp. ALY-1.**

Les résidus identiques sont surlignés en rouge et les résidus similaires sont écrits en rouge. Les éléments de structure secondaire pour alyPG et AlyA1-II' sont donnés respectivement au-dessus et en-dessous de l'alignement. Les résidus catalytiques sont marqués par une étoile et ceux impliqués dans la fixation du substrat par des triangles (Ogura *et al.*, 2008). Les insertions dans la séquence d'AlyA5 sont indiquées par des lignes épaisses vertes. Pour AlyA1, seule la séquence du module catalytique a été utilisée dans l'alignement.

La famille PL7 a récemment été subdivisée en sous-familles d'après des similitudes de séquences (Lombard *et al.*, 2010). Afin d'identifier les sous-familles pour AlyA1, AlyA2 et AlyA5, 39 séquences d'enzymes PL7 ont été alignées et utilisées pour calculer un arbre phylogénétique (Figure 7-4) en utilisant une méthode de maximum de vraisemblance (Guindon and Gascuel 2003).



**Figure 7-4: Arbre phylogénétique non raciné de 39 enzymes de la famille PL7.**

Cette arbre phylogénétique a été calculé par une approche de maximum de vraisemblance avec le programme PhyML (Guindon and Gascuel 2003). Les numéros d'accèsion Uniprot sont donnés. Les nombres indiquent les valeurs de bootstrap dans l'analyse. Les valeurs inférieures à 30 n'ont pas été indiquées. Les points rouges indiquent les enzymes de *Z. galactanivorans*. Les triangles roses indiquent les enzymes caractérisées biochimiquement. Les carrés bleus indiquent que la structure de la protéine a été résolue. SF: sous-famille. A: *Azotobacter*. Am: *Amycolatopsis*. C: *Cellulophaga*. Ca: *Catenulispora*. Cr: *Croceibacter*. G: *Gramella*. K: *Klebsiella*. Kr: *Kribbella*. P: *Pseudomonas*. R: *Rhodopirellula*. S: *Streptomyces*. Sa: *Saccharophagus*. V: *Vibrio*.

Cette approche distingue clairement les cinq sous-familles existant dans la famille PL7 (SF1 – SF5). Elle révèle que les trois enzymes PL7 encodées dans le génome de *Z. galactanivorans* appartiennent à différents groupes (Figure 7-4). AlyA1 est reliée de manière distante aux enzymes classées dans SF3 et pourrait donc être un nouveau membre de cette sous-famille. AlyA5 peut aussi être classée avec confiance dans SF5. AlyA2 appartient à un groupe solide comprenant 5 enzymes de Flavobacteriaceae marines. Lors de la création des sous-familles, Lombard *et al.* (2010) n'ont retenu que celles qui contenaient au moins cinq membres. Ainsi, je propose qu'AlyA2 et les 4 autres enzymes du même groupe appartiennent à une nouvelle sous-famille de PL7. Cette nouvelle sous-famille SF6 pourrait n'être conservée que chez les représentants marins des Flavobacteriaceae, et correspondre à une évolution des enzymes PL7 spécifique d'une niche écologique. Par ailleurs, trois autres enzymes semblent former une autre sous-famille qui serait à confirmer en incluant plus de séquences dans l'arbre.

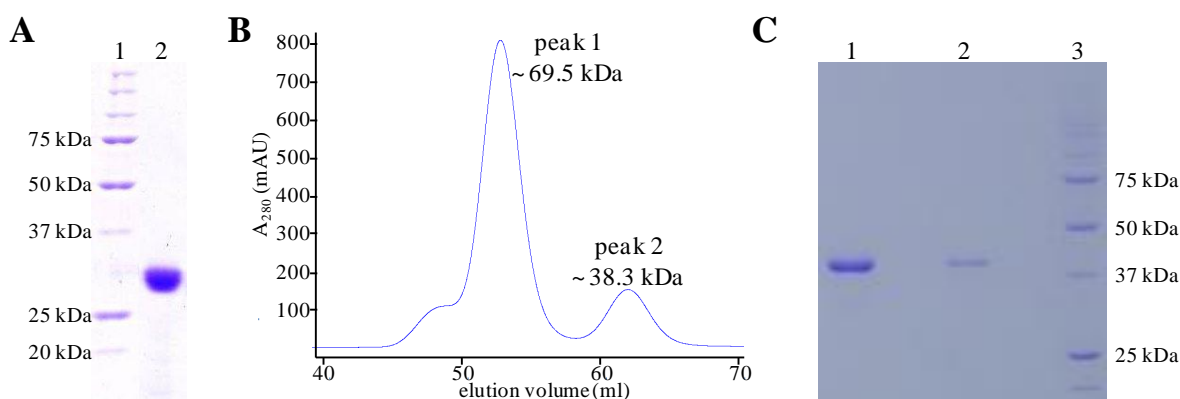
Comme les sous-familles sont corrélées avec la spécificité de substrat dans la plupart des cas, elles peuvent être utilisées pour prédire le substrat préféré de nouvelles enzymes (Lombard *et al.*, 2010). A ce jour, deux alginate lyases ont été caractérisées dans la sous-famille SF3 et quatre dans SF5 (banque de données CAZy). Elles sont toutes classées comme poly( $\alpha$ -L-gulonate) lyases (EC 4.2.2.11). Je fais donc l'hypothèse qu'AlyA1 et AlyA5 cliveraient préférentiellement les motifs G. Aucune prédiction ne peut être faite pour l'instant sur la spécificité de substrat d'AlyA2, car aucune des protéines appartenant à la nouvelle sous-famille SF6 n'a été caractérisée pour l'instant.



### III-2 Surexpression et purification des alginate lyases recombinantes

Pour caractériser les activités enzymatiques et les propriétés structurales d'AlyA1, AlyA2 et AlyA5, les séquences codantes des protéines matures (entière pour AlyA2 et AlyA5 et seulement le module catalytique pour AlyA1) ont été clonées dans le vecteur d'expression pFO4 (fusion His-tag). Les tests d'expression dans la souche *E. coli* BL21(DE3) cultivée en milieu ZYP (Studier 2005) ont été fructueux pour les protéines recombinantes AlyA1 et AlyA5, mais ont échoué pour AlyA2 (cf. Chapitre 6). Je me suis donc focalisé sur l'étude d'AlyA1 et AlyA5 dans la suite du travail.

De grandes quantités de protéines recombinantes solubles ont été obtenues et purifiées jusqu'à l'homogénéité électrophorétique par chromatographie d'affinité au nickel suivie d'une chromatographie d'exclusion de taille (Figure 7-5).



**Figure 7-5: Purification des protéines recombinantes AlyA1 et AlyA5.**

**A.** Gel SDS-PAGE de la protéine recombinante AlyA1. Les pistes 1 et 2 correspondent aux marqueurs de taille et à AlyA1, respectivement. **B.** Chromatographie d'exclusion de taille de la protéine recombinante AlyA5. La masse moléculaire a été estimée en utilisant une courbe de calibration de la colonne par différentes protéines de taille connue. **C.** Gel SDS-PAGE de la protéine recombinante AlyA5. Les pistes 1 et 2 correspondent au premier et au deuxième pic, respectivement. La piste 3 correspond aux marqueurs de taille.

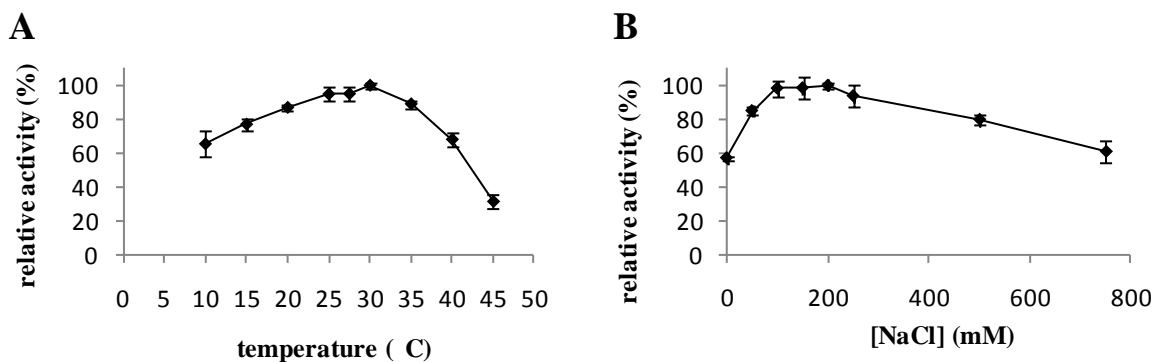
La purification en deux étapes a permis d'obtenir 30 mg d'AlyA1 recombinante à partir de 200 ml de culture (Figure 7-5 A). La masse moléculaire apparente est de 26,5 kDa (estimée par la calibration de la colonne Superdex) et 29 kDa (estimée sur SDS-PAGE). Ceci est concordant avec la masse moléculaire théorique de 27,5 kDa. Les mesures en diffusion dynamique de la lumière (DLS, *Dynamic Light Scattering*) ont indiqué un diamètre hydrodynamique de 2,77 nm. En ce qui concerne AlyA5 recombinante, deux pics ont été obtenus après chromatographie d'exclusion de taille (Figure 7-5 B). Le pic majoritaire a été

élué comme une protéine de 69,5 kDa et le pic minoritaire comme une protéine de 38,3 kDa. L'analyse en SDS-PAGE montre que ces deux populations contiennent seulement une seule et même forme en conditions dénaturantes, avec une masse moléculaire apparente de 42 kDa (Figure 7-5 C). Par comparaison avec la masse moléculaire théorique de 38,3 kDa, les protéines éluées dans le premier et le second pic correspondent environ à un dimère et un monomère d'AlyA5, respectivement. Les mesures en DLS ont indiqué un diamètre hydrodynamique 2,50 nm pour le monomère et de 3,20 nm pour le dimère. La purification d'AlyA5 recombinante a abouti à 2,2 mg de la forme monomérique et 10 mg de la forme dimérique à partir d'une culture de 200 ml. Un second passage du dimère sur colonne d'exclusion de taille a montré que cette forme est stable en solution, puisqu'un seul pic a été obtenu, correspondant à une taille de 69,5 kDa (données non présentées).

### III-3 Caractérisation biochimique d'AlyA1

#### III-3-a Effet de la température et des sels

L'activité d'AlyA1 est maximale à 30°C et 200 mM NaCl (Figure 7-6). Seulement 30% d'activité subsistent à 45°C. Ceci est bien corrélé avec une mesure de la stabilité thermique de la protéine recombinante. En utilisant le DLS, il a été montré que la dénaturation commence pour des températures supérieures à 45°C (données non présentées).



**Figure 7-6: Effet de la température et de la concentration en NaCl sur l'activité d'AlyA1**

Les expériences ont été réalisées en tampon Tris-HCl pH 7,5 en utilisant l'alginate de sodium 0,05% comme substrat et 12 nM de l'enzyme purifiée. Les valeurs représentent les moyennes  $\pm$  S.D. ( $n = 3$ ). **A**, Effet de la température. L'activité à 30°C a été prise comme 100%. **B**, Effet de la concentration en NaCl. Les réactions ont été réalisées à 30°C. L'activité avec 200 mM NaCl a été prise comme 100%.

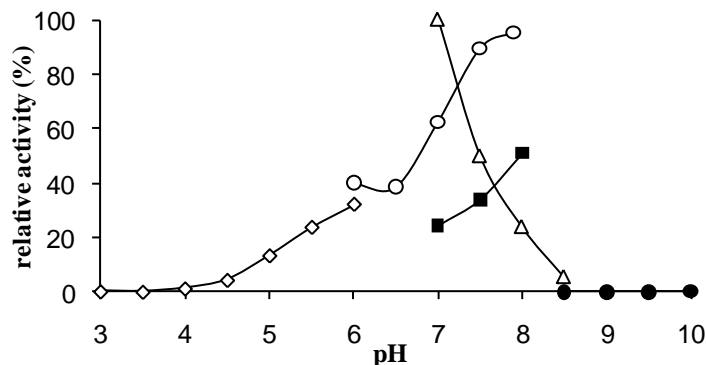
Les agents chélatants comme l'EDTA et l'EGTA inhibent l'activité d'AlyA1 (Tableau 7-3). Cette inhibition a été seulement partielle quand les agents chélatants ont été incubés avec l'enzyme avant la réaction, mais non ajoutés dans le mélange réactionnel. L'inhibition a été presque totale quand les agents chélatants ont été ajoutés directement dans la réaction. Ceci suggère que l'interaction enzyme / cation a lieu lors de la fixation du substrat, et non avant.

Traitement inhibiteur	% activité
aucun	100
EDTA 1 mM	
une nuit avec AlyA1	56,6
ajouté dans la réaction	10,7
EGTA 1mM	
une nuit avec AlyA1	42,3
ajouté dans la réaction	5,4

**Tableau 7-3 : Effet d'agents chélatants sur l'activité d'AlyA1**

### III-3-b Effet du pH

AlyA1 présente l'activité maximale à pH 7,0 en tampon Tris-HCl (Figure 7-7). Ceci est similaire au pH optimal déterminé pour d'autres alginate lyases de bactéries marines (Wong *et al.*, 2000). Le choix de la molécule tampon semble critique, comme le montre l'exemple de la baisse de 40% de l'activité en tampon MOPS à pH 7,0 comparé au Tris-HCl. La gamme d'activité est assez étroite avec 40% d'activité résiduelle à pH 6,0 et 5% à pH 8,5. L'enzyme est inactive à pH < 4,0 et pH > 8,5.



**Figure 7-7: Effet du pH sur l'activité d'AlyA1**

Les expériences ont été réalisées à 30°C en tampon à 100 mM, en utilisant l'alginate 0,05% comme substrat et 27 nM d'enzyme purifiée. Les tampons citrate de sodium (*losanges vides*), MOPS (*cercles vides*), Tris-HCl (*triangles vides*), HEPES (*carrés pleins*) et glycine-NaOH (*cercles pleins*) ont été utilisés. L'activité en tampon Tris-HCl pH 7,0 a été prise comme 100%.

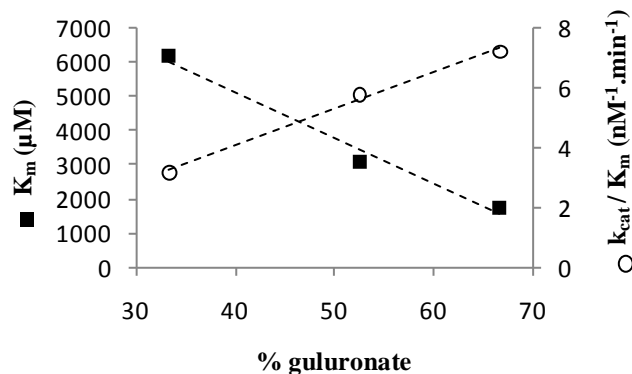
### III-3-c Détermination des constantes cinétiques d'AlyA1

Afin d'évaluer la spécificité d'AlyA1, ses paramètres cinétiques ont été déterminés en utilisant trois lots d'alginate avec différents contenus en guluronate et mannuronate. Le  $K_m$  pour l'alginate riche en G est environ 4 fois plus faible que celui pour l'alginate riche en M (Tableau 7-4). Une relation inverse a été trouvée pour le paramètre  $k_{cat} / K_m$ , qui reflète la spécificité d'une enzyme envers différents substrats (Eisenthal *et al.*, 2007). Des relations linéaires ont été obtenues en exprimant ces deux paramètres en fonction du pourcentage de motifs G dans le substrat (Figure 7-8). Par conséquent, AlyA1 est une guluronate lyase (EC 4.2.2.11). Ceci concorde avec les spécificités de substrat connues pour les deux autres alginate lyases de la sous-famille SF3 caractérisées (Figure 7-4). Cette analyse ayant été conduite avant la détermination du pH optimum, elle a été réalisée en tampon Tris-HCl pH 7,5 au lieu de 7,0.

% guluronate	$K_m$ ( $\mu\text{M}$ )	$k_{cat}$ ( $\text{s}^{-1}$ )	$k_{cat} / K_m$ ( $\text{nM}^{-1} \cdot \text{s}^{-1}$ )
66.7%	1753	12,66	7,22
52.6%	3087	17,89	5,79
33.3%	6180	19,51	3,16

L'activité alginate lyase a été mesurée par l'augmentation de  $A_{235}$  à 30°C, en tampon Tris-HCl 100 mM pH 7,5 et 200 mM NaCl, avec des concentrations en substrat de 0,075 à 2,5  $\text{mg} \cdot \text{ml}^{-1}$ . La masse molaire d'un monosaccharide constitutif de l'alginate a été utilisée pour le calcul de la concentration en substrat (194  $\text{g} \cdot \text{mol}^{-1}$ ). Le coefficient d'extinction molaire pour les produits (acides uroniques insaturés) a été considéré de  $\epsilon = 8500 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$  (Preiss and Ashwell, 1962).

**Tableau 7-4 : Constantes cinétiques d'AlyA1**

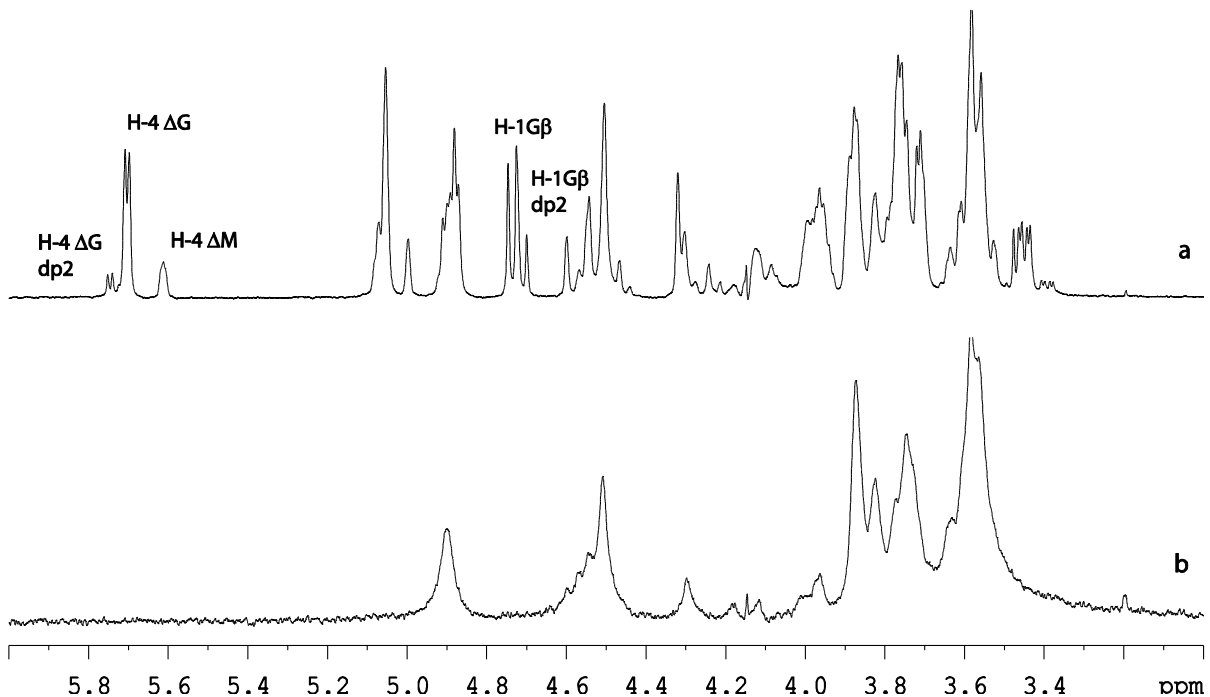


**Figure 7-8 : Relations entre les constantes cinétiques d'AlyA1 et le contenu en guluronate dans l'alginate.**

$K_m$  (carrés pleins) et  $k_{cat} / K_m$  (cercles vides) sont exprimés en fonction du pourcentage de motifs G dans le substrat.

### III-3-d Analyse de la spécificité de substrat par spectrométrie RMN

La spécificité de substrat d'AlyA1 a été déterminée en suivant la réaction enzymatique directement dans le tube RMN à 30°C. La structure du résidu à l'extrémité réductrice peut être facilement identifiée par le déplacement chimique caractéristique du signal pour le proton anomérique. Ainsi, un résidu guluronate à l'extrémité réductrice donne un doublet à 4,70-4,75 ppm avec une constante de couplage  $^3J_{\text{HH}} = 8,6$  Hz. Un résidu mannuronate donne quant à lui un pic unique à 4,75-4,80 ppm (Heyraud *et al.*, 1996). La Figure 7-9 montre que lorsque l'alginate de sodium est dégradé par AlyA1, un doublet à 4,74 ppm apparaît dans le spectre RMN. Cette observation montre que l'action d'AlyA1 produit des oligosaccharides avec un résidu guluronate à l'extrémité réductrice.



**Figure 7-9 : Suivi  $^1\text{H}$ -RMN de la dégradation d'alginate par AlyA1**

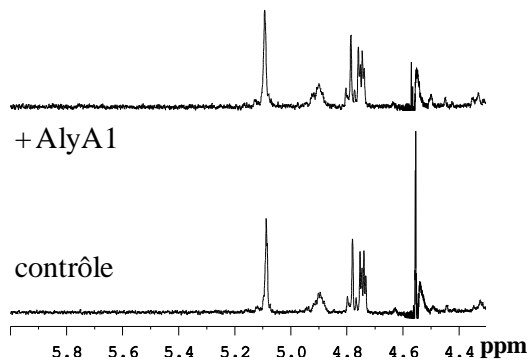
Les spectres ont été acquis à 30°C en tampon phosphate deutéré 100 mM, pH 7,6, 200 mM NaCl, 5 mM  $\text{CaCl}_2$  sur de l'alginate de sodium intact (**b**) ou après dégradation par AlyA1 (**a**).

Les déplacements chimiques des protons H4 du résidu insaturé  $\Delta$  à l'extrémité non-réductrice dépendent de la nature du résidu voisin. Un doublet H4( $\Delta$ ) à  $\delta$  5,75 ppm indique que le résidu voisin du  $\Delta$  est un guluronate. Le signal H4( $\Delta$ ) apparaît à 5,67 ppm si le résidu voisin du  $\Delta$  est un mannuronate (Heyraud *et al.*, 1996; Zhang *et al.*, 2004). Sur la Figure 7-9, on constate que l'intensité du signal H4 pour le motif  $\Delta\text{G}$  est beaucoup plus intense que pour

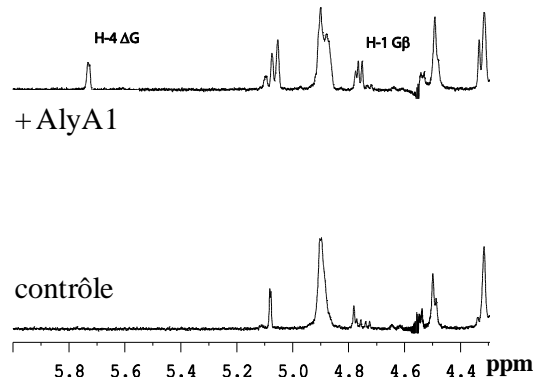
$\Delta$ M. Ceci montre que les oligosaccharides produits par AlyA1 sont majoritairement de type  $\Delta$ G(X)<sub>n</sub>G, avec X étant soit G soit M.

Ces résultats montrent que AlyA1 clive le substrat en laissant un résidu G à la nouvelle extrémité réductrice formée. Afin de déterminer si l'enzyme réalise la  $\beta$ -élimination indifféremment sur des motifs G-G et G-M, la réaction de dégradation de substrats polyG, polyM et polyMG a été suivie par RMN (Figure 7-10).

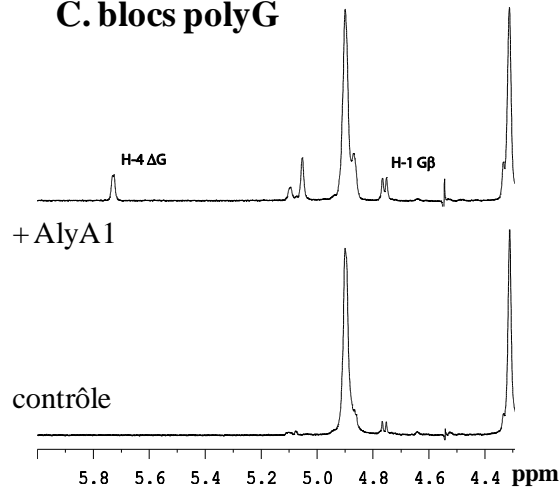
### A. blocs polyMG



### B. blocs polyM



### C. blocs polyG



**Figure 7-10 : Suivi  $^1\text{H}$ -RMN de la dégradation de blocs de motifs connus par AlyA1**

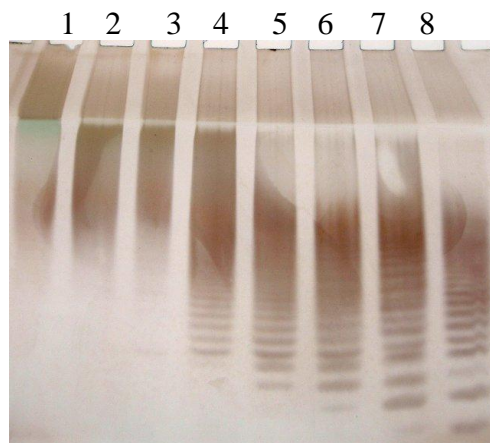
Les spectres ont été acquis avant (contrôle) et après (+AlyA1) incubation de l'enzyme avec des substrats alternés polyMG (A) ou homopolymériques polyM (B) et polyG (C).

Le spectre des blocs polyMG n'est pas modifié après incubation avec AlyA1 (Figure 7-10 A), ce qui indique que l'enzyme n'est pas active sur ce type de substrat où les résidus M

et G sont alternés. Lors de l'utilisation de blocs polyM comme substrats, aucun signal H4 pour le motif  $\Delta M$  n'apparaît dans le spectre (Figure 7-10 B). Ceci montre que l'enzyme ne peut cliver entre deux résidus mannuronate. En agrandissant le spectre pour détecter de faibles signaux, on constate l'existence d'un signal H4( $\Delta$ ) correspondant au motif  $\Delta G$ , qui provient vraisemblablement du clivage de contaminants GG dans la fraction des blocs polyM. La dégradation de blocs polyG par AlyA1 génère un signal H4( $\Delta$ ) correspondant au motif  $\Delta G$  (Figure 7-10 C). Ces résultats montrent que l'alginate lyase AlyA1 ne peut cliver qu'entre deux résidus guluronates.

### III-3-e Mode d'action

Les produits de dégradation résultant de l'action d'AlyA1 sur l'alginate ont été analysés en C-PAGE selon une cinétique. Après seulement 2 min de dégradation, l'intensité de coloration dans le gel de concentration est beaucoup plus faible qu'au temps 0 (Figure 7-11). Ceci montre que la fraction de très haut degré de polymérisation est rapidement dégradée. En suivant la réaction de 2 min à 14 h, la taille des oligosaccharides détectés devient de plus en plus petite. Par conséquent, AlyA1 agit sur l'alginate de manière endolytique. Ceci signifie que l'enzyme attaque son substrat au hasard et en milieu de chaîne, et clive si elle trouve un enchaînement de deux résidus G.

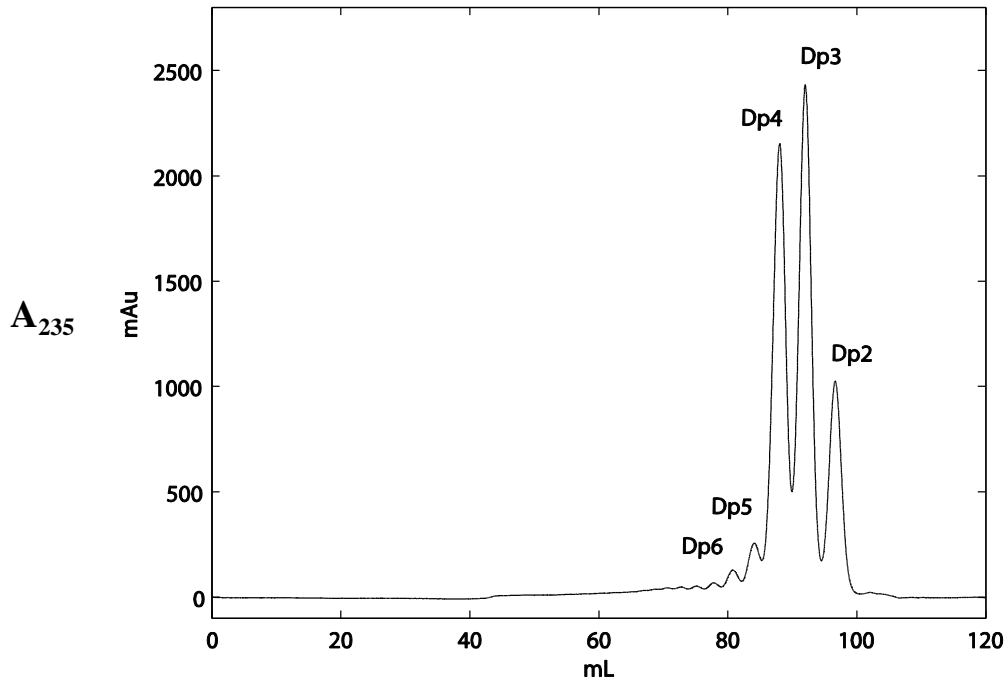


**Figure 7-11: Analyse en C-PAGE des produits de la dégradation d'alginate par AlyA1**

La réaction a été réalisée à 30°C en tampon Tris-HCl 100 mM pH 7,5 en utilisant l'alginate de sodium 0.05% comme substrat et 12 nM de l'enzyme purifiée. Des aliquots ont été prélevés avant l'ajout de l'enzyme (piste 1) et après 2 min (piste 2), 5 min (piste 3), 10 min (piste 4), 30 min (piste 5), 1 h (piste 6), 2 h (piste 7) et 14 h (piste 8).

### III-3-f Caractérisation des produits terminaux

Les oligosaccharides finaux résultant de la dégradation d'alginate par AlyA1 ont été séparés par chromatographie d'exclusion de taille. Le chromatogramme montre trois fractions majoritaires (Figure 7-12). L'analyse ESI-MS de ces fractions a mesuré des masses moléculaires de 351,0, 526,8 et 703,0 correspondant à des di-, tri- et tétrasaccharides, respectivement.



**Figure 7-12 : Chromatographie d'exclusion de taille des produits terminaux de la dégradation par AlyA1**  
dp : degré de polymérisation

Les fractions molaires des différents produits ont été estimées par l'aire des pics sur le chromatogramme. A complétion, les trisaccharides et tétrasaccharides sont les produits majoritaires de la dégradation avec une contribution totale de 41% et 36% respectivement. Les produits contiennent également 19% de disaccharides et de faibles quantités de pentamères et d'hexamères (Tableau 7-5).

Les structures et les proportions des oligosaccharides contenus dans ces fractions de taille uniforme ont été déterminées par RMN. Le spectre RMN de la fraction "dp2" a montré qu'elle contient uniquement le disaccharide  $\Delta G$ . Le spectre RMN de la fraction "dp3" présente deux signaux  $H_4(\Delta)$ , l'un à 5,75 ppm correspondant au motif à un G voisin d'un motif  $\Delta$ , et l'autre à 5,67 ppm correspondant à un M voisin d'un  $\Delta$ . Ceci montre l'existence de deux produits dans



la fraction dp3, identifiés comme  $\Delta$ GG et  $\Delta$ MG. Les proportions de  $\Delta$ GG (88%) et  $\Delta$ MG (12%) ont été déterminées en intégrant les signaux H4- $\Delta$ G et H4- $\Delta$ M, respectivement (Tableau 7-5). Le spectre RMN de la fraction "dp4" montre aussi la présence de plusieurs produits. Deux signaux H4( $\Delta$ ) ont été assignés aux motifs  $\Delta$ G- et  $\Delta$ M- d'après leur déplacement chimique caractéristique à 5,75 ppm et 5,67 ppm, respectivement. Comme la fraction ne contient que des oligosaccharides avec un G à l'extrémité réductrice, les produits possibles sont  $\Delta$ GXG et  $\Delta$ MXG. L'intégration du signal H4(D) a permis d'estimer les proportions à 84% pour  $\Delta$ GXG et 16% pour  $\Delta$ MXG. Le déplacement chimique du proton H2 d'un G à l'extrémité réductrice dépend aussi de la nature du résidu voisin, avec  $\delta=3,60$  ppm pour  $(X)_n$ GG et  $\delta=3,67$  ppm pour  $(X)_n$ MG. Ces deux signaux ont été observés dans le spectre de la fraction "dp4", montrant qu'elle contient les quatre types de tétrasaccharides  $\Delta$ GGG,  $\Delta$ GMG,  $\Delta$ MGG et  $\Delta$ MMG.

Oligosaccharides	UV	RMN	Contribution totale
<b>dp2</b>	19%	-	
$\Delta$ G		100%	19%
<b>dp3</b>	41%	-	
$\Delta$ GG		88%	36%
$\Delta$ MG		12%	5 %
<b>dp4</b>	36%	-	
$\Delta$ GXG		84%	30%
$\Delta$ MXG		16%	6%
<b>dp5</b>	2%	-	2%
<b>dp6</b>	1%	-	1%
<b>&gt; dp6</b>	1%	-	1%

**Tableau 7-5 : Proportions des différents oligosaccharides produits par AlyA1.**

Ces proportions ont été estimées en intégrant les pics du chromatogramme d'exclusion de taille (UV) et les signaux H4( $\Delta$ ) des spectres RMN.

Au total, les produits majoritaires de la dégradation d'alginate par AlyA1 sont donc  $\Delta$ GG (36%),  $\Delta$ GXG (30%) et  $\Delta$ G (19%). La forte proportion de di-, tri- et tétrasaccharides et la très

faible proportion d'oligosaccharides avec  $dp > 4$  suggère que le motif minimal reconnu soit un pentasaccharide. En effet, si un tétrasaccharide pouvait être clivé, la proportion de disaccharides devrait être plus importante que celles des oligosaccharides  $dp3$  et  $dp4$ , et on ne devrait pas retrouver de produits de type  $\Delta GGG$ .

La caractérisation biochimique d'AlyA1 montre donc qu'il s'agit d'une endo-gulonate lyase ne pouvant cliver qu'entre deux résidus guluronates. Ceci signifie que les sous-sites -1 et +1 ne peuvent accommoder que des unités G. Le motif minimal reconnu est un pentasaccharide. L'identification des produits terminaux de la dégradation permet d'inférer les spécificités de reconnaissance des différents sous-sites de l'enzyme (Figure 7-13). Ainsi, la présence de résidus M et G dans les produits finaux implique que le sous-site +2 et un éventuel sous-site +3 peuvent accommoder un résidu M, bien qu'ayant probablement une préférence pour les résidus G. De même, du côté non-réducteur, les sous-sites -3 et -2 pourraient fixer des résidus G, M et  $\Delta$ .

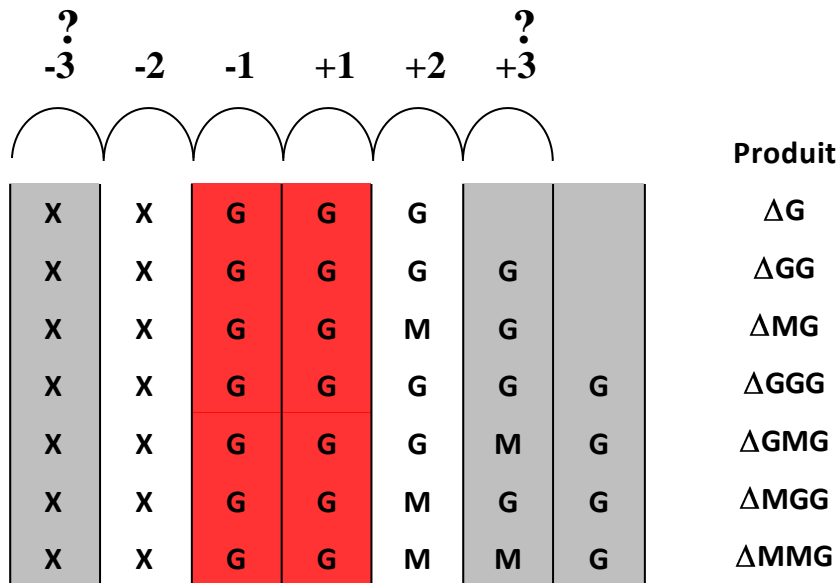


Figure 7-13 : Spécificité de reconnaissance des sous-sites d'AlyA1

### III-4 Caractérisation biochimique d'AlyA5

#### III-4-a Action d'AlyA5 sur les polysaccharides d'alginate

Dans une première tentative de caractériser l'enzyme AlyA5 recombinante, son activité a été testée sur des polysaccharides d'alginate avec différents contenus en guluronate. L'ajout d'au moins 2 mM CaCl<sub>2</sub> dans le mélange réactionnel est indispensable pour détecter une activité, montrant que AlyA5 est calcium-dépendante (données non présentées). Le monomère et le dimère d'AlyA5 dégradent l'alginate. Cependant, leur activité spécifique est 10 fois plus faible par rapport à AlyA1, indiquant que les polysaccharides ne sont peut-être pas les substrats préférés d'AlyA5 (Tableau 7-6). Dans ces conditions, l'activité spécifique est légèrement plus élevée sur l'alginate possédant le plus fort taux en guluronate.

contenu en guluronate (%)	Activité spécifique (U.mg <sup>-1</sup> )		
	AlyA1	AlyA5 monomère	AlyA5 dimère
66,7%	209,0	20,6	23,8
52,6%	189,7	13,7	17,4
33,3%	114,9	12,4	13,2

Les activités ont été mesurées à 30°C dans des mélanges composés de 100 mM Tris-HCl pH 7,5; 200 mM NaCl; 0,25 mg.ml<sup>-1</sup> alginate et 12 nM d'AlyA1 ou 27 nM d'AlyA5. CaCl<sub>2</sub> a été ajouté à la concentration finale de 4 mM pour les mesures d'activité d'AlyA5. Une unité d'activité a été définie comme une augmentation d'une unité d'A<sub>235</sub> par minute.

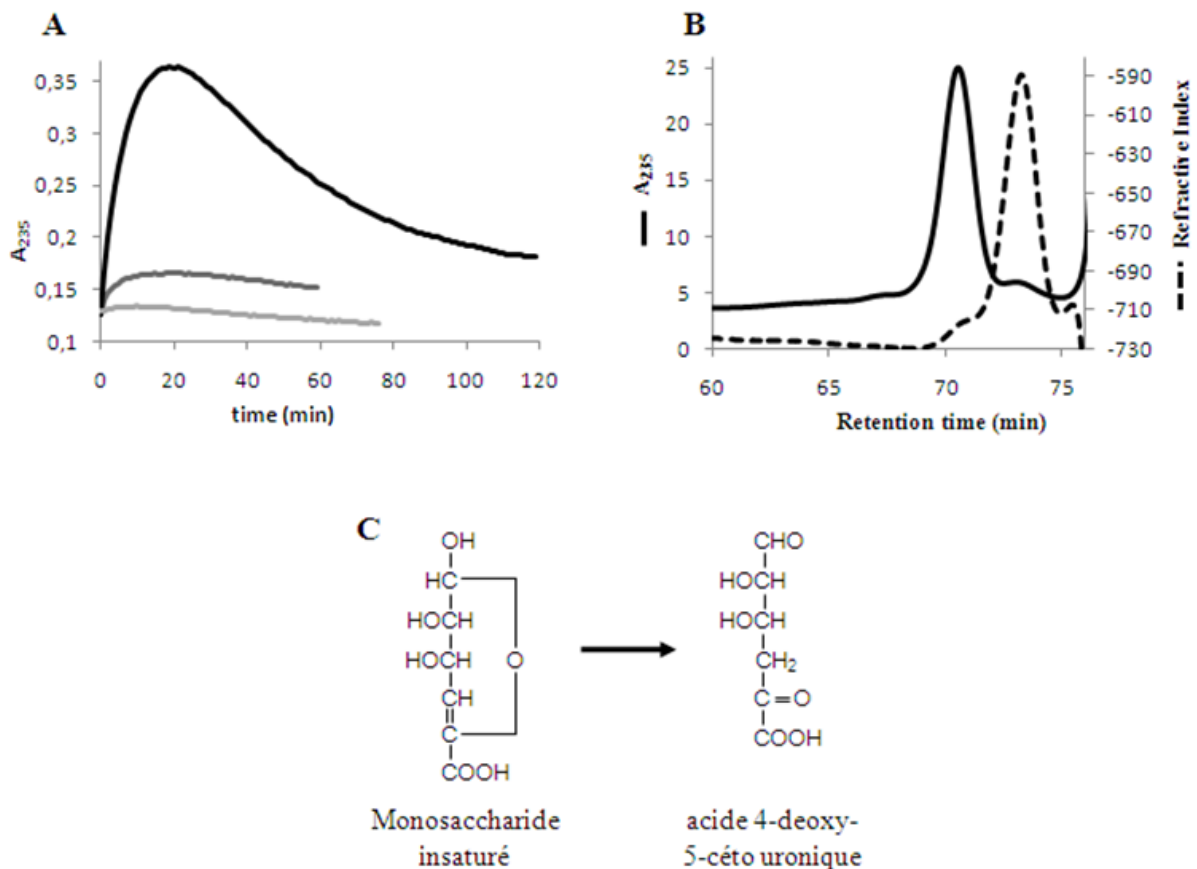
**Tableau 7-6 : Comparaison des activités spécifiques d'AlyA1 et AlyA5 sur les polysaccharides d'alginate**

A cause de son plus grand rendement de production et pour des raisons de temps, la caractérisation biochimique qui suit a été uniquement réalisée sur la forme dimérique d'AlyA5.

#### III-4-b Action d'AlyA5 sur des oligo-alginates saturés

L'activité d'AlyA5 dimère a été testée sur des oligomannuronates (oligoM), oligoguluronates (oligoG) et des blocs mixtes (oligoMG) avec des degrés de polymérisation allant de DP4 à DP30. Les suivis de A<sub>235</sub> montrent une cinétique en deux phases (Figure 7-14 A). Dans la première phase, on détecte une augmentation de l'absorbance. Ceci montre

qu'AlyA5 clive les oligosaccharides par un mécanisme de  $\beta$ -élimination. Les activités spécifiques ont été calculées dans cette première phase. AlyA5 est beaucoup plus active sur les oligoG ( $449,3 \text{ U.mg}^{-1}$ ) que les oligoMG ( $13,5 \text{ U.mg}^{-1}$ ) ou oligoMM ( $2,0 \text{ U.mg}^{-1}$ ). Cette première phase est suivie par une décroissance de  $A_{235}$ , qui suggère la nature transitoire des composés absorbant dans l'UV. Un tel profil est typique pour les lyases produisant des monosaccharides insaturés (Hoffman *et al.*, 1960; Preiss and Ashwell 1963). En effet, ces composés sont convertis de manière non-enzymatique en  $\alpha$ -cétoacides, i.e., acide 4-désoxy-5-céto-uronique (Figure 7-14C). Ceux-ci n'absorbent plus à 235 nm à cause du transfert de l'insaturation vers le groupe cétone (Preiss and Ashwell 1962a; Hashimoto *et al.*, 2000; Ochiai *et al.*, 2006). Cependant, on remarque sur la Figure 7-14 A que l'absorbance ne redescend pas jusqu'à son niveau initial, ce qui suggère que le monosaccharide insaturé ne soit pas le seul produit de la réaction.



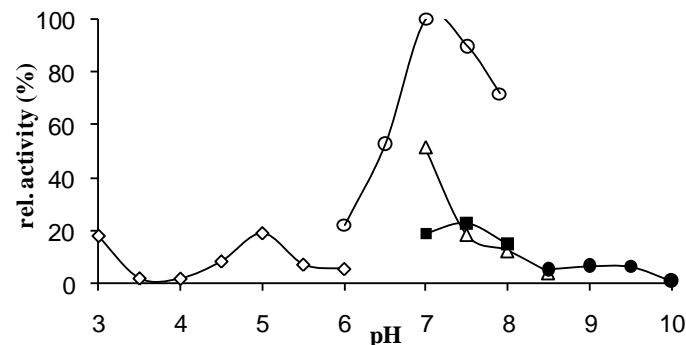
**Figure 7-14 : Action d'AlyA5 dimère sur les oligo-alginates saturés**

**A.** Cinétiques de dégradation d'oligoG (noir), oligoMG (gris foncé) et oligoM (gris clair) par AlyA5.  $A_{235}$  a été mesurée à  $30^\circ\text{C}$  dans des réactions ( $500 \mu\text{l}$ ) composées de  $100 \text{ mM}$  Tris-HCl pH 7,5,  $4 \text{ mM}$   $\text{CaCl}_2$ ,  $50 \mu\text{g}$  de substrat et  $27 \text{ nM}$  d'enzyme. **B.** Chromatographie d'exclusion de taille des produits finaux de la dégradation d'oligoG par AlyA5. L'éluion a été suivie par  $A_{235}$  (trait continu) et l'index de réfraction (trait pointillé). **C.** Conversion non-enzymatique du monosaccharide insaturé en acide 4-désoxy-5-céto-uronique n'absorbant plus à  $235 \text{ nm}$ .

Les produits finaux de la dégradation d'oligoG par AlyA5 ont été analysés par chromatographie d'exclusion de taille (Figure 7-14 B). Le signal RI présente un premier pic peu élevé (temps de rétention RT=70,5 min) suivi d'un pic majoritaire (RT=73 min). Le signal UV présente un profil inverse : le premier pic est le plus grand (RT=70,5 min) et seulement un faible signal est détecté à RT=73 min. Le composé élué dans le second pic a ainsi été identifié comme l'acide 4-désoxy-5-céto uronique (n'absorbant pas à 235 nm), résultant de la conversion du monosaccharide insaturé (Figure 7-14 C). Le premier pic, composé de matériel absorbant à 235 nm, est le disaccharide insaturé  $\Delta$ G. L'intensité de l'index de réfraction RI est proportionnelle à la masse des molécules : en quantité égale, un disaccharide devrait donner un signal environ doublé par rapport à un monosaccharide. Par conséquent, le produit majoritaire de la dégradation par AlyA5 est l'acide 4-désoxy-5-céto-uronique, avec de faibles traces de disaccharide insaturé.

Sur la base des données présentées, AlyA5 a donc pu être montrée comme agissant de manière exolytique. Ceci signifie que l'enzyme attaque son substrat en extrémité de chaîne. L'activité relativement faible sur les polysaccharides d'alginate par rapport aux oligosaccharides peut s'expliquer par le nombre plus faible d'extrémités disponibles pour l'attaque. A la fois sur les polysaccharides et les oligosaccharides, l'activité est plus forte sur le substrat ayant un contenu en guluronate plus important. Ainsi, AlyA5 est une exo-guluronate lyase. Cette spécificité de substrat correspond à ce qui a déjà été montré pour d'autres enzymes de la même sous-famille SF5.

L'effet du pH sur l'activité d'AlyA5 a été testé. Le dimère d'AlyA5 présente la plus forte activité à pH 7,0 en tampon MOPS (Figure 7-15). La gamme d'activité d'AlyA5 est encore plus limitée que celle d'AlyA1 (cf. Figure 7-7). Le choix du tampon est là aussi critique.

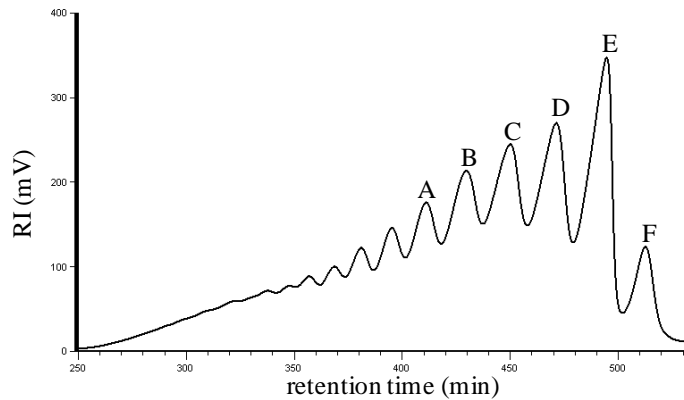


**Figure 7-15: Effet du pH sur l'activité d'AlyA5**

Les expériences ont été réalisées à 30°C en tampon à 100 mM en utilisant les oligoG (100 $\mu$ g.ml<sup>-1</sup>) comme substrat et 56 nM d'enzyme purifiée. Les tampons citrate de sodium (*losanges vides*), MOPS (*cercles vides*), Tris-HCl (*triangles vides*), HEPES (*carrés pleins*) et glycine-NaOH (*cercles pleins*) ont été utilisés. L'activité en tampon MOPS pH 7,0 a été prise comme 100%.

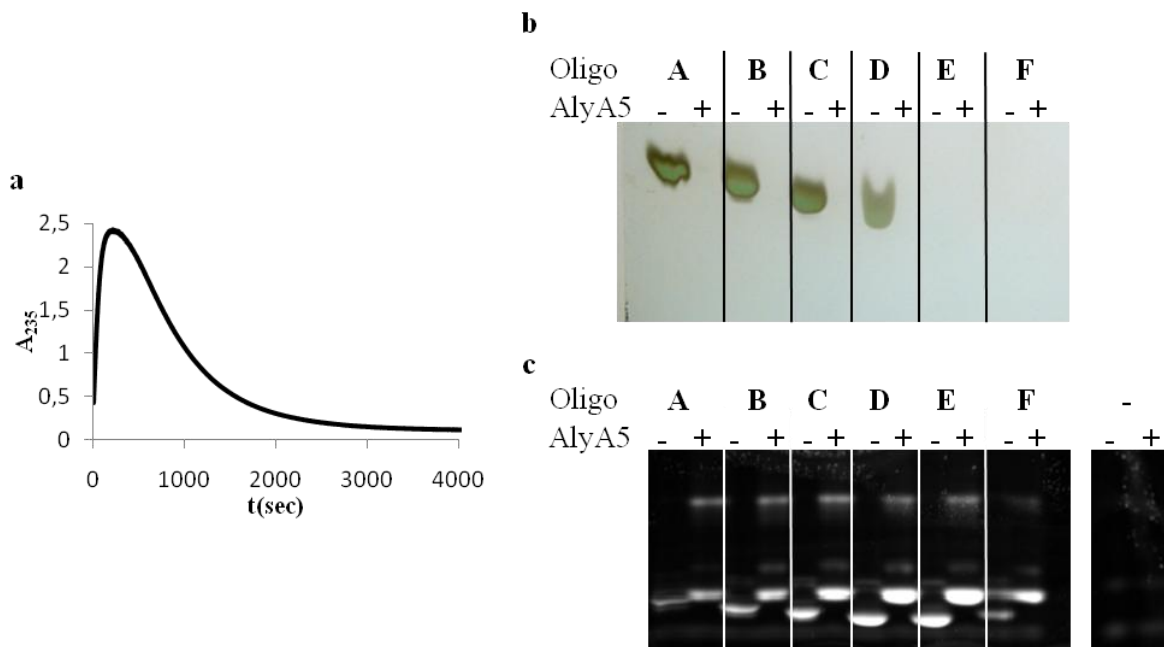
**III-4-c Action d'AlyA5 sur des oligosaccharides insaturés**

Les oligosaccharides insaturés ont été préparés et purifiés à partir d'une dégradation partielle d'alginate par l'endo-gulonate lyase AlyA1 (Figure 7-16). Six d'entre eux (oligoA à oligoF) ont ensuite été utilisés comme substrat pour AlyA5.



**Figure 7-16: Chromatographie d'exclusion de taille des oligosaccharides insaturés résultant de la dégradation partielle par AlyA1.**

A, B, C, D, E, F: oligosaccharides utilisés comme substrats pour AlyA5



**Figure 7-17: Dégradation d'oligosaccharides insaturés par AlyA5**

50 µg d'oligosaccharides insaturés ont été utilisés comme substrat pour 27 nM d'AlyA5 dimère dans 100 mM Tris-HCl pH 7,5; 200 mM NaCl; 4 mM CaCl<sub>2</sub>. **a.** L'action d'AlyA5 sur l'oligoA purifié a été suivie par A<sub>235</sub>. **b, c.** Les produits finaux de l'action d'AlyA5 sur les 6 oligosaccharides ont été analysés en C-PAGE (b) et en FACE (c).

Comme pour les oligosaccharides saturés, l'action d'AlyA5 sur des oligosaccharides insaturés provoque une première phase d'augmentation rapide de  $A_{235}$ , suivie par une baisse plus lente (Figure 7-17 a). Ceci montre qu'AlyA5 peut dégrader les oligo-alginates insaturés en monosaccharides, convertis par la suite en acide 4-désoxy-5-céto-uronique. On constate cependant que dans ce cas, l'absorbance revient assez rapidement à son niveau initial, contrairement à ce qui avait été trouvé lors de l'utilisation d'oligoalginates saturés comme substrat (Figure 7-14 A). Ce point sera rediscuté ultérieurement.

AlyA5 dégrade complètement les oligo-alginates insaturés de différentes tailles en un composé trop petit pour être coloré par la méthode du C-PAGE. (Figure 7-17 b). Les oligoE et oligoF sont également trop petits pour être détectés. L'analyse en FACE a révélé que les produits finaux de la réaction d'AlyA5 sur les six tailles d'oligo-alginates sont les mêmes (Figure 7-17 c). Deux bandes principales sont observées après dégradation par AlyA5 : une bande intense, correspondant à un composé avec une mobilité électrophorétique relativement élevée, et une bande moins intense avec une mobilité beaucoup plus faible. Ces deux composés pourraient correspondre au disaccharide insaturé et à l'acide 4-désoxy-5-céto-uronique, respectivement. La purification des produits finaux de la réaction d'AlyA5 sur les oligo-alginates est en cours. Leur caractérisation devrait permettre d'identifier les deux bandes observées en FACE.

#### III-4-d Action d'AlyA5 sur des oligo-alginates réduits

Le fait qu'AlyA5 puisse dégrader de manière exolytique à la fois des oligosaccharides saturés et insaturés permet d'obtenir des indications sur l'extrémité attaquée et le motif reconnu. En effet, seuls deux cas de figure sont possibles. AlyA5 pourrait soit attaquer un motif guluronate à partir de l'extrémité réductrice, soit attaquer du côté non-réducteur et tolérer un motif G ou insaturé. Pour commencer à élucider cette question, les extrémités réductrices des oligoG saturés ont été réduites par le borohydrure de sodium. Sur de tels oligosaccharides, l'extrémité réductrice ne devrait plus être utilisable pour un éventuel clivage. Une telle méthode a déjà été utilisée pour étudier le mécanisme d'action d'une glucanase (Nelson *et al.*, 1963). Comparé à la réaction sur des oligoG intacts, AlyA5 conserve 75% de son activité sur des oligoG réduits. La diminution de 25% peut être attribuée à la perte de matériel lors des étapes successives d'évaporation / resolubilisation incluses dans le protocole. Par conséquent, AlyA5 semble attaquer son substrat du côté non-réducteur. Le site catalytique doit donc pouvoir accommoder à la fois des motifs G ou insaturés.

### III-4-e Suivi de la dégradation par spectrométrie $^1\text{H}$ -RMN

La dégradation des différents types de blocs polyM, polyG et polyMG par AlyA5 a été suivie par RMN du proton à 30°C directement dans le tube RMN. Les spectres des blocs alternés polyMG ou homopolymériques polyM restent identiques avant et après incubation plusieurs heures avec l'enzyme, démontrant qu'AlyA5 n'est pas active sur ce type de substrat (données non présentées). En revanche, les blocs polyG sont dégradés par AlyA5, comme le montrent les changements dans les spectres RMN durant l'incubation avec l'enzyme (Figure 7-18). Ces résultats démontrent qu'AlyA5 n'est pas active sur les motifs MM, MG et GM et qu'elle peut cliver la liaison entre deux G ou entre un G et un  $\Delta$ .

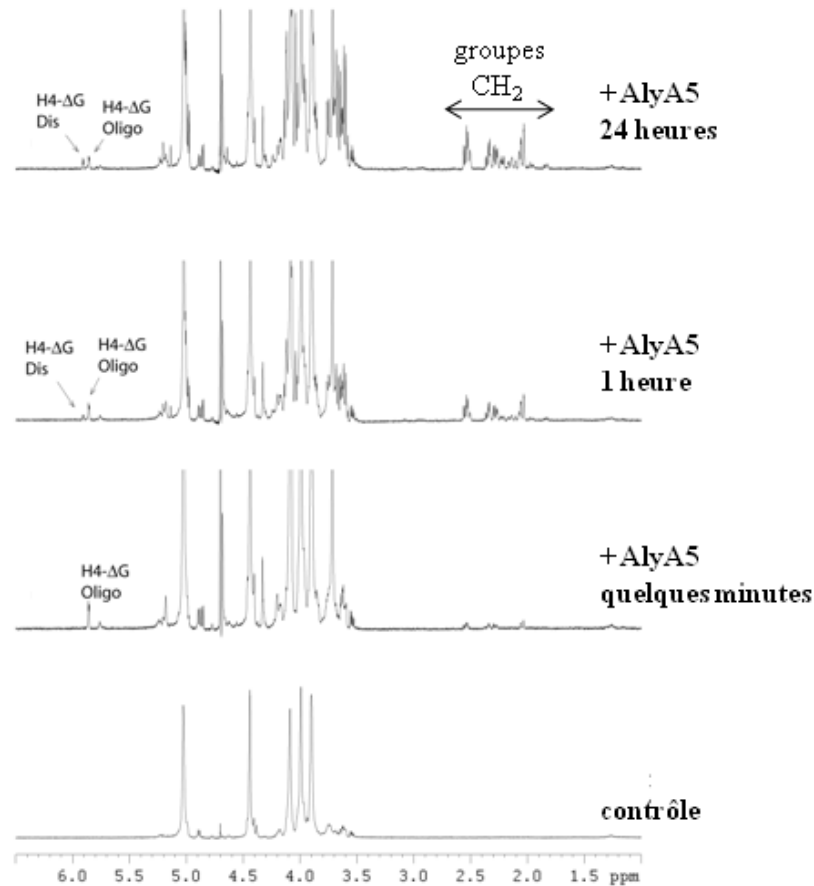


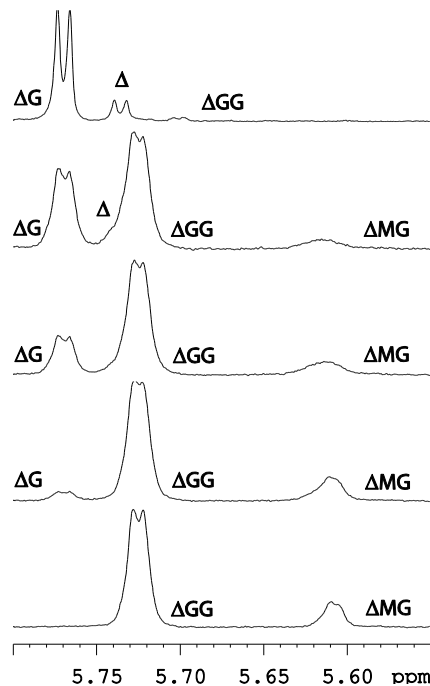
Figure 7-18 : Suivi en  $^1\text{H}$ -RMN de l'action d'AlyA5 sur les blocs polyG

Quelques minutes après le début de la réaction, on observe un pic  $\text{H4}(\Delta)$  à 5,75 ppm caractéristique des oligosaccharides  $\Delta\text{G}(\text{X})_n$  plus longs que  $\text{dp}2$ . Au cours de l'incubation, ce signal diminue, et un autre signal apparaît correspondant au disaccharide  $\Delta\text{G}$ . Cependant, l'intensité de ce signal reste relativement faible, même après 24 heures de réaction. Ce résultat



confirme que le disaccharide  $\Delta G$  est un produit minoritaire de la dégradation par AlyA5. Comme il a été décrit précédemment, AlyA5 agit de manière exolytique, et le produit majoritaire de la réaction est le monomère. Ce composé est spontanément converti en acide 4-désoxy-5-céto-uronique. De manière concordante, le suivi de la réaction par RMN montre l'apparition de signaux correspondants à deux groupes  $CH_2$  différents. Ces signaux ont des déplacements chimiques typiques pour les groupes  $CH_2$  proches d'un groupe cétone ou hémiacétal (Kuorelahti *et al.*, 2006). L'intensité de ces signaux augmente au cours de la réaction. Ainsi, l'analyse RMN confirme qu'un des produits majoritaires de la dégradation par AlyA5 est l'acide 4-désoxy-5-céto-uronique.

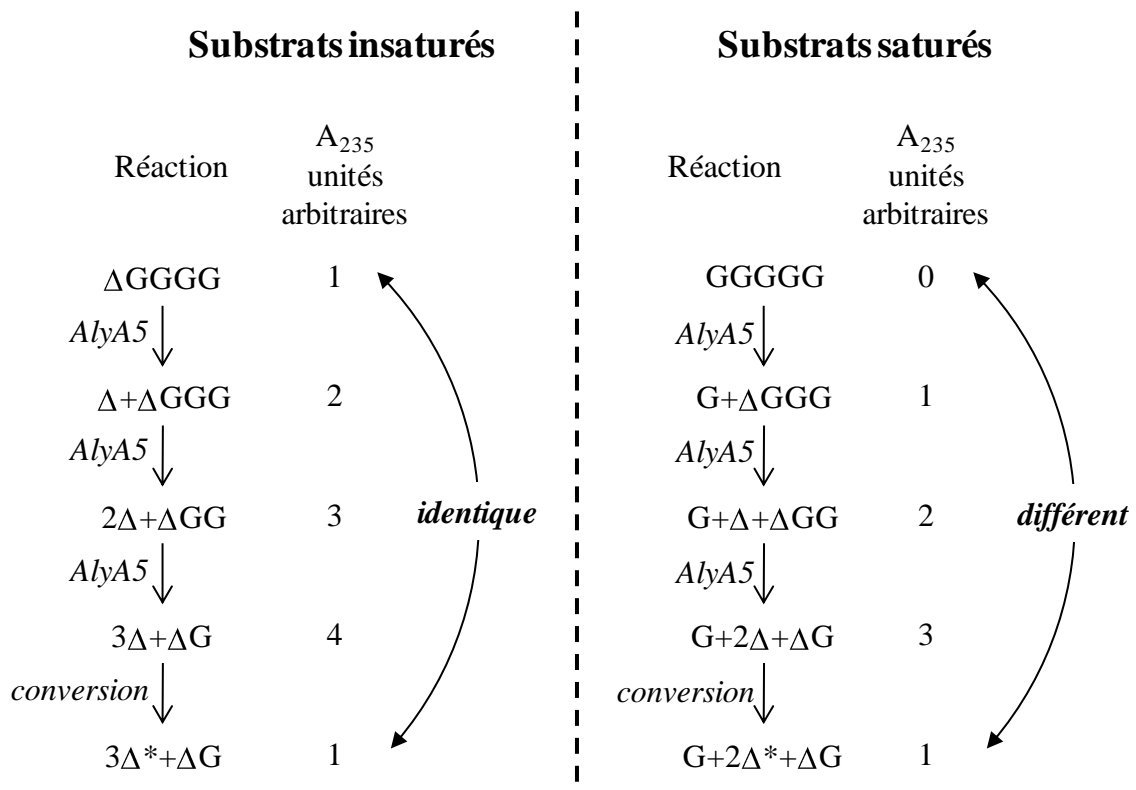
L'action d'AlyA5 sur l'oligosaccharide purifié  $\Delta GG$  a été suivie par RMN. On constate l'apparition de signaux correspondant au disaccharide  $\Delta G$  et au monosaccharide  $\Delta$  avant conversion (Figure 7-19). La quantité de  $\Delta G$  est beaucoup plus importante que celle de  $\Delta$ , indiquant comme précédemment que le monosaccharide insaturé n'est pas stable en solution. Ceci est confirmé par l'apparition de signaux dans la région 1,5-2,6 ppm, dûs à sa conversion en acide 4-désoxy-5-céto-uronique. Par ailleurs, l'oligosaccharide  $\Delta GG$  utilisé comme substrat contenait une faible contamination par  $\Delta MG$ , comme indiqué par le signal pour  $H_4(\Delta)$  à 5,61 ppm. Ce signal disparaît au cours de la réaction, montrant qu'AlyA5 est également capable de cliver la liaison glycosidique  $\Delta$ -M.



**Figure 7-19 : Suivi en  $^1H$ -RMN de l'action d'AlyA5 sur l'oligosaccharide  $\Delta GG$**

Spectre du bas : avant ajout de l'enzyme. De bas en haut : Suivi au cours de la réaction. Seule la région des protons  $H_4(D)$  à l'extrémité non-réductrice est montrée.

La caractérisation biochimique d'AlyA5 a donc mis en évidence qu'il s'agit d'une exogulonate lyase agissant du côté non réducteur du substrat. L'enzyme clive les liaisons G-G,  $\Delta$ -G et  $\Delta$ -M. Ainsi, le sous-site +1 de l'enzyme pourrait fixer un résidu G ou M, alors que le sous-site -1 accommoderait à la fois les résidus G et  $\Delta$ . La détection de  $\Delta$ G en fin de dégradation (Figure 7-14 B et Figure 7-19) montre que le disaccharide n'est pas clivé par l'enzyme. Ceci suggère que le motif minimal de reconnaissance d'AlyA5 est le trisaccharide. Ce résultat explique les différences observées dans les suivis de l'absorbance lors des réactions d'AlyA5 sur des oligoalginates insaturés (retour de  $A_{235}$  au niveau initial; cf Figure 7-17 A) ou saturés (baisse de  $A_{235}$ , mais pas de retour au niveau initial; cf. Figure 7-14 A). En effet, les substrats insaturés possèdent une absorbance intrinsèque. En fin de dégradation par AlyA5, on obtient autant de disaccharides insaturés que de molécules de substrats au départ, et donc l'absorbance est la même en début et en fin de réaction (Figure 7-20). Au contraire, les substrats saturés n'ont aucune absorbance au départ, et la présence de disaccharide insaturé comme produit terminal explique que l'absorbance soit plus forte en fin de réaction qu'au début.



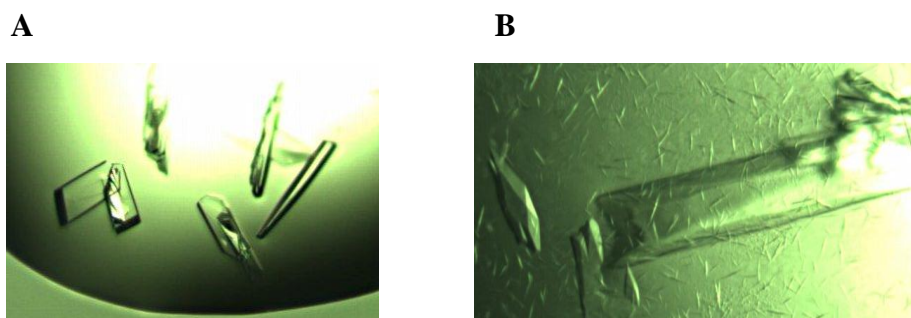
**Figure 7-20 : Schéma d'interprétation de l'activité d'AlyA5 sur des oligogulonates saturés et insaturés**

Les différentes étapes de la dégradation d'un pentasaccharide saturé ou insaturé sont représentées. La présence d'un résidu insaturé est comptée comme 1 unité arbitraire d'absorbance. Pour simplifier, la conversion non enzymatique du monosaccharide insaturé  $\Delta$  en forme acide 4-désoxy-5-céto-uronique ( $\Delta^*$ ) est considérée comme ne commençant qu'en fin de réaction, ce qui n'est pas le cas dans la réalité.

### III-5 Etude structurale des deux alginate lyases AlyA1 et AlyA5

#### III-5-a Cristallogenèse

Les cribles de cristallisation effectués grâce au robot ont mis en évidence un certain nombre de conditions prometteuses. Pour AlyA1, la gamme de conditions optimales obtenues à l'issue du crible va de 15% à 25% de PEG 1500 ou 3350, avec un pH variant autour de 7. Pour AlyA5, une seule condition des deux cribles testés a donné des cristaux encourageants, avec 20% PEG 3350 et 0,2 M  $\text{Na}_2\text{SO}_4$ . L'optimisation manuelle a consisté à faire varier un ou plusieurs paramètres, tels que la concentration en agent précipitant, en sels, l'ajout d'additif pour ralentir la nucléation (ex. isopropanol), le volume de la goutte ou du réservoir, etc. Pour AlyA1, les premières conditions testées ont donné de grandes plaquettes collées les une aux autres. Il a alors été envisagé d'ajouter des oligoG dans la solution. Un nouveau crible a été réalisé au robot. L'ajout d'oligoG a complètement changé la forme des cristaux et les conditions de cristallisation, cette fois-ci dans une gamme de 28 à 33% de PEG-MME 2000 et en présence de thiocyanate de potassium. Après optimisation manuelle, des monocristaux isolés et de volume plus important ont été obtenus dans la condition 0,2 M KSCN, 28% PEG-MME (Figure 7-21 A). Pour le dimère d'AlyA5, la meilleure trouvée condition (22% PEG 3350; 0,2 M NA/ tartrate) permet la formation de gros cristaux en un à deux mois (Figure 7-21 B).

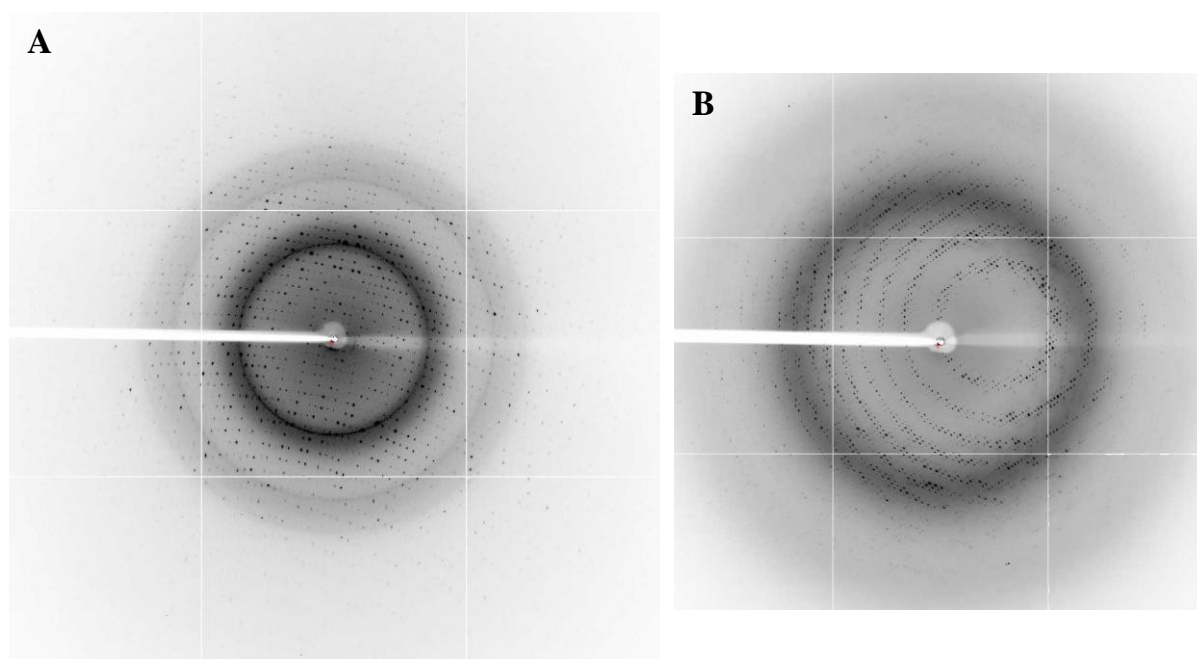


**Figure 7-21 : Cristaux des deux alginate lyases recombinantes**

**A.** Cristaux de la protéine AlyA1 recombinante (module catalytique). **B.** Cristaux de la forme dimérique d'AlyA5 recombinante.

### III-5-b Cristallographie aux rayons X

Les cristaux obtenus ont été utilisés pour des expériences de diffraction des rayons X afin de déterminer la structure des deux protéines AlyA1 et AlyA5. Deux exemples de clichés de diffraction obtenus sont donnés sur la Figure 7-22. Les paramètres de la collecte sont donnés dans le Tableau 7-7. Les cristaux d'AlyA1 diffractent jusqu'à 1,43 Å et appartiennent au groupe d'espace triclinique P1. Un jeu de données à haute résolution (1,05 Å) a également été collecté. Le cristal d'AlyA5 qui a été analysé diffracte jusqu'à 1,75 Å et appartient au groupe d'espace orthorhombique P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>.



**Figure 7-22 : Clichés de diffraction obtenus pour les cristaux d'AlyA1 (A) et AlyA5 (B)**

Les deux protéines AlyA1 et AlyA5 cristallisent chacune avec deux molécules dans l'unité asymétrique. Selon le calcul du volume spécifique  $V_M$  d'après Matthews (1968), le cristal d'AlyA1 contient 41% de solvant, et celui d'AlyA5 68%. Les structures des deux protéines ont été résolues par remplacement moléculaire, en utilisant les coordonnées atomiques des alginate lyases ALY-1 de *Corynebacterium* sp. (PDB 1UAI) et A1-II' de *Sphingomonas* sp. A1 (PDB 2Z42) comme modèle. Les statistiques de l'affinement des structures sont données dans le Tableau 7-7.

	AlyA1	AlyA5 dimère
<b>Données de la collecte</b>		
Ligne de lumière	PROXIMA 1	
Longueur d'onde	0,98 Å	
Groupe d'espace	P1	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Paramètres de maille	a=43,86 Å ; b=50,39 Å ; c=55,52 Å $\alpha=69,14^\circ$ ; $\beta=90,02^\circ$ ; $\gamma=84,92^\circ$	a=93,42 Å ; b=93,91 Å ; c=130,15 Å $\alpha=\beta=\gamma=90,00^\circ$
Gamme de résolution (Å)	43,66-1,43 (1,47-1,43)	41,95-1,75 (1,80-1,75)
Nombre d'observations	358902	967134
Réflexions uniques	78881	114168
Complétude (%)	96,1 (94,2)	99,4 (93,3)
Moyenne I/ $\sigma$ (I)	14,16 (7,79)	16,53 (2,30)
R <sub>merge</sub> (%)	7,8 (21,9)	9,6 (84,4)
Redondance	4,55	8,47
<b>Données de l'affinement</b>		
Gamme de résolution	40,66-1,43 (1,47-1,43)	41,95-1,80 (1,85-1,80)
Réflexions uniques	74933 (5454)	101057 (7343)
Nombre de réflexions R <sub>free</sub>	3944 (287)	5319 (386)
R / R <sub>free</sub> (%)	14,6/20,1 (11,0/19,4)	15,4/17,8 (21,1/24,6)
RMSD longueurs de liaison	0,024 Å	0,039 Å
RMSD angles de liaison	2,00°	3,02°
Facteur B moyen (Å <sup>2</sup> )	12,74	29,16
Facteur B Molécule A (Å <sup>2</sup> )	10,39	21,74
Facteur B Molécule B (Å <sup>2</sup> )	10,32	21,31
Facteur B Solvant (Å <sup>2</sup> )	25,00	22,67
Facteur B Ligands (Å <sup>2</sup> )	-	31,56

**Tableau 7-7 : Données de la collecte et statistiques de l'affinement pour AlyA1 et AlyA5.**

Les valeurs entre parenthèses concernent la dernière couche de résolution.

La protéine AlyA1 a pu être modélisée quasi-entièrement. Le modèle final contient 248 résidus : seuls les quatre premiers acides aminés de l'extrémité N-terminale n'ont pu être construits dans la densité électronique. Malgré la présence d'oligogulonates dans le tampon de cristallisation, aucun sucre n'a pu être identifié dans la densité électronique autour du site actif. L'enzyme étant active, on peut penser que ces substrats aient été clivés. Les oligogulonates ou leurs produits de dégradation auraient donc agi comme additifs, favorisant la cristallisation de l'enzyme.

Les extrémités d'AlyA5 sont mal définies dans la densité. Les 22 premiers acides aminés en N-ter et les 3 derniers acides aminés en C-ter ne sont pas visibles dans la carte de densité électronique 2Fo-Fc. Le modèle final comporte 312 résidus. On retrouve dans le site actif de l'enzyme une molécule de tartrate provenant du tampon de cristallisation. Une molécule de glycérol provenant du cryoprotectant est également présente dans la structure.

### III-5-c Les structures d'AlyA1 et AlyA5 de *Z. galactanivorans*

Comme attendu pour des enzymes de la famille PL7, les deux alginate lyases AlyA1 et AlyA5 adoptent un repliement en sandwich  $\beta$  incurvé (Figure 7-23 A et B, page suivante). Celui-ci est globalement similaire au repliement de l'enzyme A1-II' de *Sphingomonas* sp. A1 (Figure 7-23 C), obtenu en complexe avec un oligosaccharide (Ogura *et al.*, 2008). La gorge de fixation du substrat est formée par la courbure des feuillets  $\beta$  le long de l'enzyme et renferme le site catalytique. La position des acides aminés catalytiques (Q127, H129, Y226 pour AlyA1 et Q131, H133, Y283 pour AlyA5) est conservée avec A1-II' (Figure 7-23, résidus matérialisés en rouge). Malgré la similitude du repliement, on constate d'importantes différences entre les structures d'AlyA1 et AlyA5 entre elles et avec celle de l'alginate lyase A1-II'. Dans la structure d'A1-II', deux boucles viennent recouvrir la gorge catalytique (Figure 7-23 C, boucles en bleu). Dans la structure obtenue pour l'alginate lyase AlyA1 de *Z. galactanivorans*, ces deux boucles sont en conformation ouverte (Figure 7-23 A, boucles en rouge). Cette différence est encore plus évidente si on observe la surface des enzymes en vue de dessus. Dans l'enzyme de *Sphingomonas* sp. A1, les deux boucles en conformation fermée viennent fermer la gorge, créant ainsi un tunnel dans lequel le substrat se fixe (Figure 7-25 A). Pour AlyA1, les deux boucles en position ouverte laissent la gorge libre (Figure 7-25 B). La superposition de l'oligosaccharide modélisé dans A1-II' sur la structure d'AlyA1 native montre que la topologie en sillon peut accueillir le substrat. Cette topologie est compatible avec le mode d'action endolytique déterminée pour AlyA1. Chez *Sphingomonas* sp. A1, il a été montré que des liaisons hydrogène entre deux asparagines N141 et N199 assurent l'interaction des deux boucles au-dessus de la gorge pour former un tunnel. La fermeture irréversible du tunnel par la mutation des deux asparagines en cystéines et la création d'un pont disulfure diminue fortement l'activité de l'enzyme (Ogura *et al.*, 2008). Ceci montre que la flexibilité des boucles est déterminante pour permettre la fixation du substrat dans le site actif et la libération des produits de la réaction. Chez AlyA1, il est possible que les deux boucles soient également flexibles. Elles auraient été piégées en conformation ouverte dans le cristal et pourraient se refermer lors de la fixation du substrat. Les deux boucles contiennent effectivement des arginines candidates (N75 ou N78 dans la boucle 1, N134 dans la boucle 2) qui pourraient former des liaisons hydrogène comme dans l'enzyme A1-II'. Le facteur B des résidus (facteur de température) dans la structure donne des indications sur la fluctuation des atomes autour de leur position moyenne. Dans les chaînes A et B d'AlyA1, trois régions présentent un facteur B relativement plus élevé que le reste de la protéine (Figure 7-24). De

manière intéressante, deux de ces régions (résidus 72-78 et 133-139) correspondent aux deux boucles qui pourraient se refermer au-dessus du site actif pour former le tunnel. Ce résultat montre donc que les deux boucles trouvées en conformation ouverte dans la structure d'AlyA1 sont flexibles. Il vient étayer l'hypothèse de leur fermeture lors de la fixation du substrat.

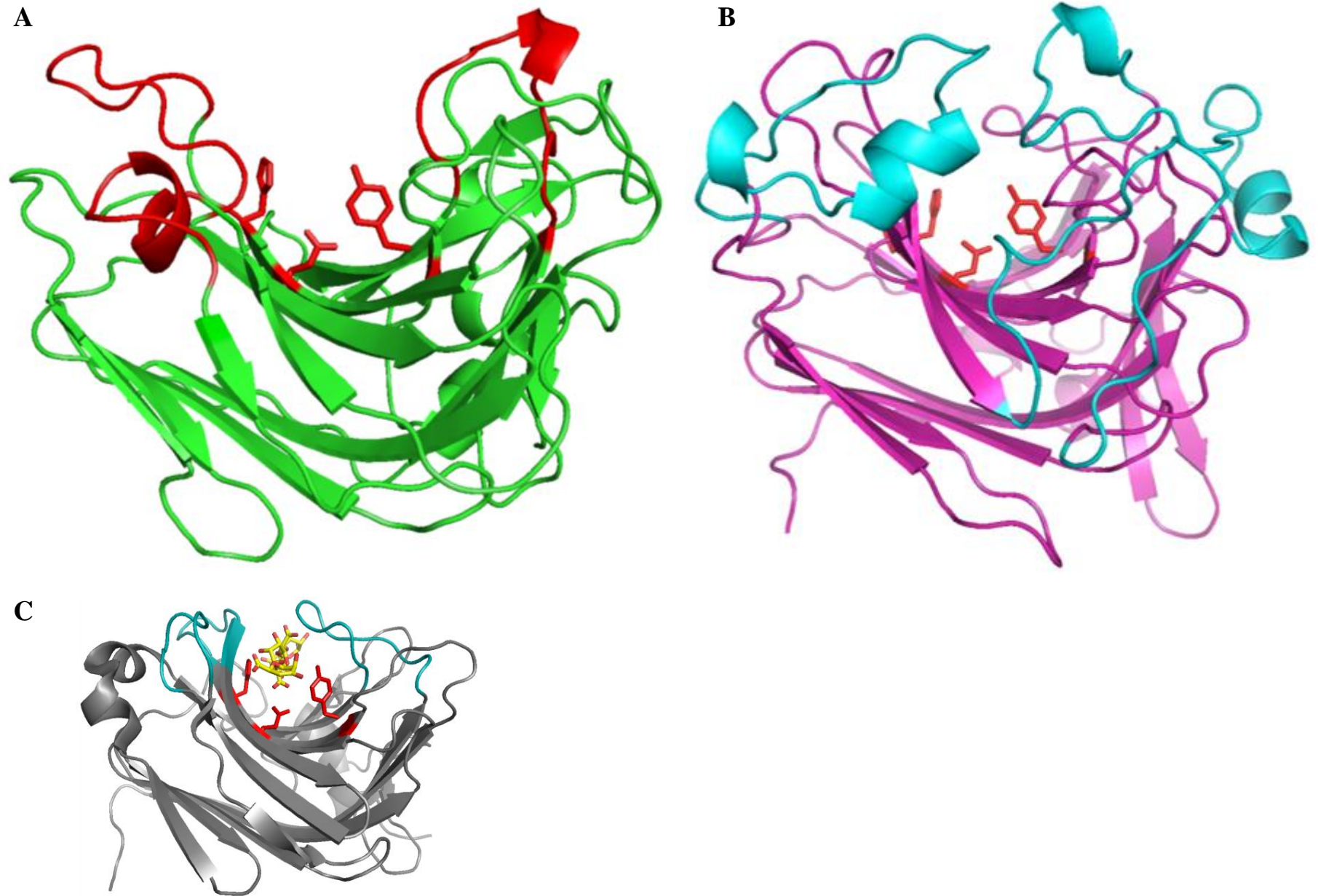
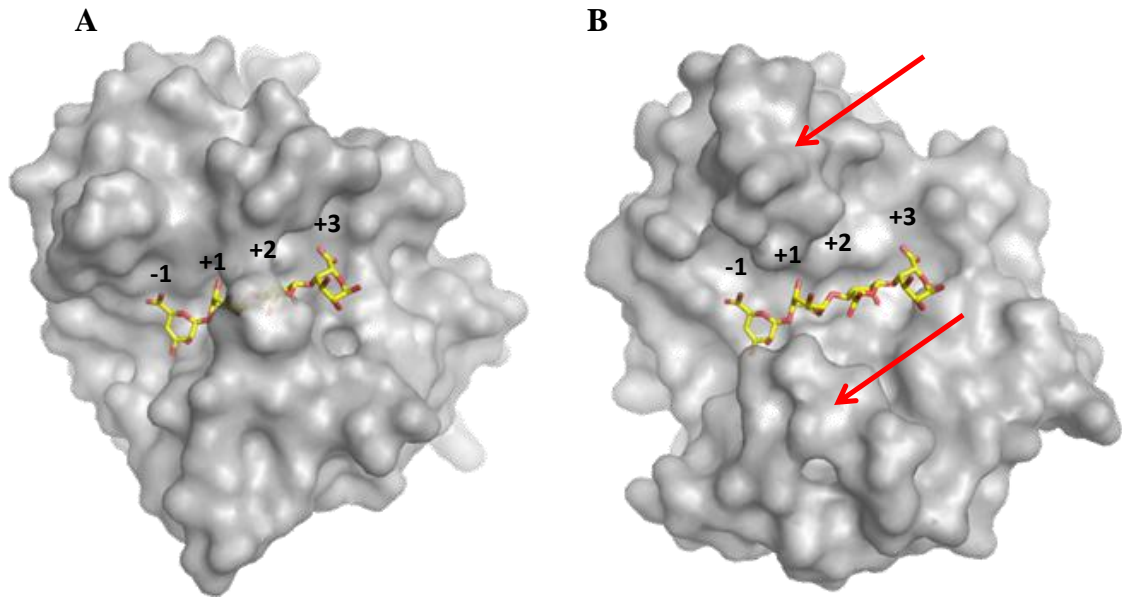


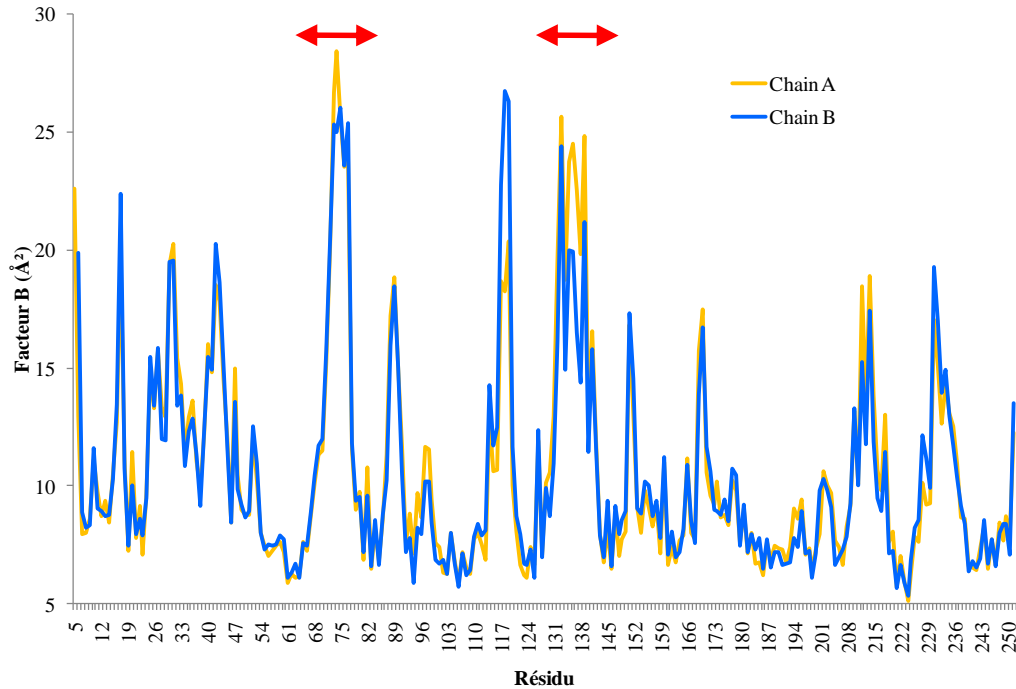
Figure 7-23 : Comparaison des structures d'AlyA1 (A) et AlyA5(B) avec celle d'A1-II' de *Sphingomonas* sp.A1 (C) obtenue en complexe avec un oligosaccharide.





**Figure 7-25 : Comparaison des surfaces des enzymes A1-II' et AlyA1.**

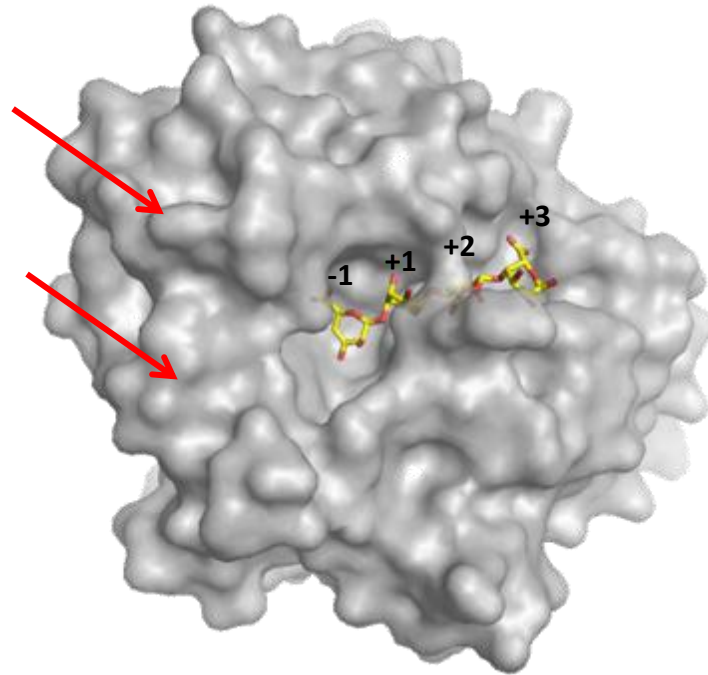
**A.** Modélisation de la surface de A1-II' en complexe avec un tétrasaccharide  $\Delta$ GMG (PDB 2ZAA, Ogura *et al.*, 2008). La position des sous-sites de l'enzyme est indiquée. **B.** Modélisation de la surface de AlyA1 de *Z.galactanivorans*. L'oligosaccharide provenant de la structure d'A1-II' a été superposé. Les flèches rouges indiquent les deux boucles en conformation ouverte.



**Figure 7-24 : Evolution du facteur B des résidus le long de la séquence d'AlyA1.**

Les facteurs B sont représentés pour la chaîne A (jaune) et B (bleu). Les flèches rouges indiquent les deux régions flexibles correspondant aux boucles qui pourraient se refermer sur le site catalytique.

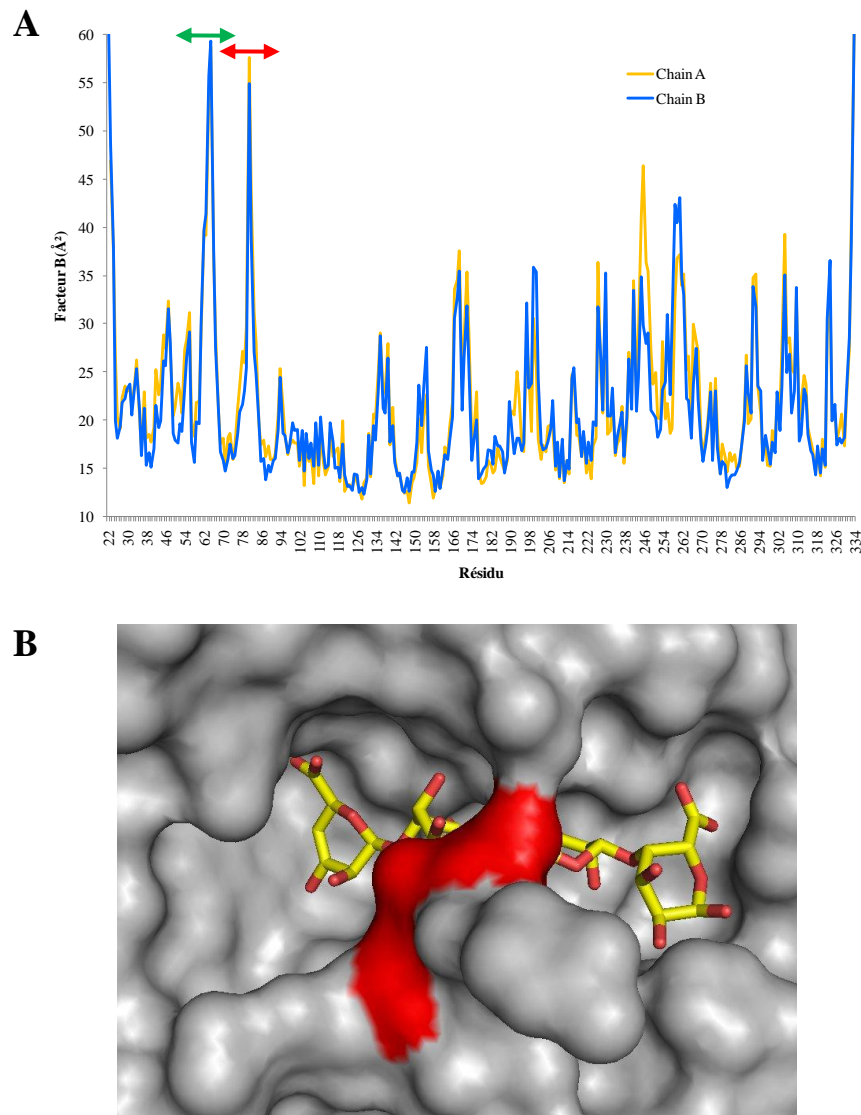
La structure d'AlyA5 de *Z. galactanivorans* est la première structure pour une exo-alginate lyase dans la famille PL7. Elle présente une différence majeure par rapport à celles d'AlyA1 et d'A1-II'. En effet, deux longues boucles additionnelles semblent venir fermer la gorge d'un côté (Figure 7-23 B, boucles en bleu). Ces boucles correspondent aux insertions spécifiques à AlyA5 déjà mentionnées dans le paragraphe III-1 de ce Chapitre (Figure 7-3). Dans la visualisation de la surface de l'enzyme, on constate que ces boucles obstruent complètement le sillon formé par le sandwich  $\beta$  (Figure 7-26). L'émergence de ces deux boucles chez AlyA5 donne à l'enzyme une topologie en forme de poche. Ceci explique le mode d'action exolytique démontré lors de la caractérisation biochimique de cette alginate lyase. En effet, la superposition dans AlyA5 du tétrasaccharide provenant de la structure d'A1-II' en complexe montre clairement que cette topologie empêche la fixation d'un sucre au-delà du sous-site -1.



**Figure 7-26 : Surface de l'enzyme AlyA5 de *Z. galactanivorans*.**

Modélisation de la surface de AlyA5. Le tétrasaccharide  $\Delta$ GMG provenant de la structure d'A1-II' a été superposé (PDB 2ZAA, Ogura *et al.*, 2008). Les flèches rouges indiquent les deux insertions venant fermer la gorge pour former une poche.

L'analyse du facteur B le long de la séquence montre deux pics majeurs (Figure 7-27 A). L'un d'eux correspond à une région au dos de l'enzyme (flèche verte). L'autre (résidus 81-83, flèche rouge) correspond à la boucle qui recouvre le site actif de l'enzyme et confère la topologie en tunnel (Figure 7-27 B). Cette boucle semble donc flexible. Elle peut probablement s'ouvrir pour permettre la fixation du substrat dans la poche d'AlyA5.



**Figure 7-27 : Étude de la flexibilité de la boucle couvrant le site actif d'AlyA5**

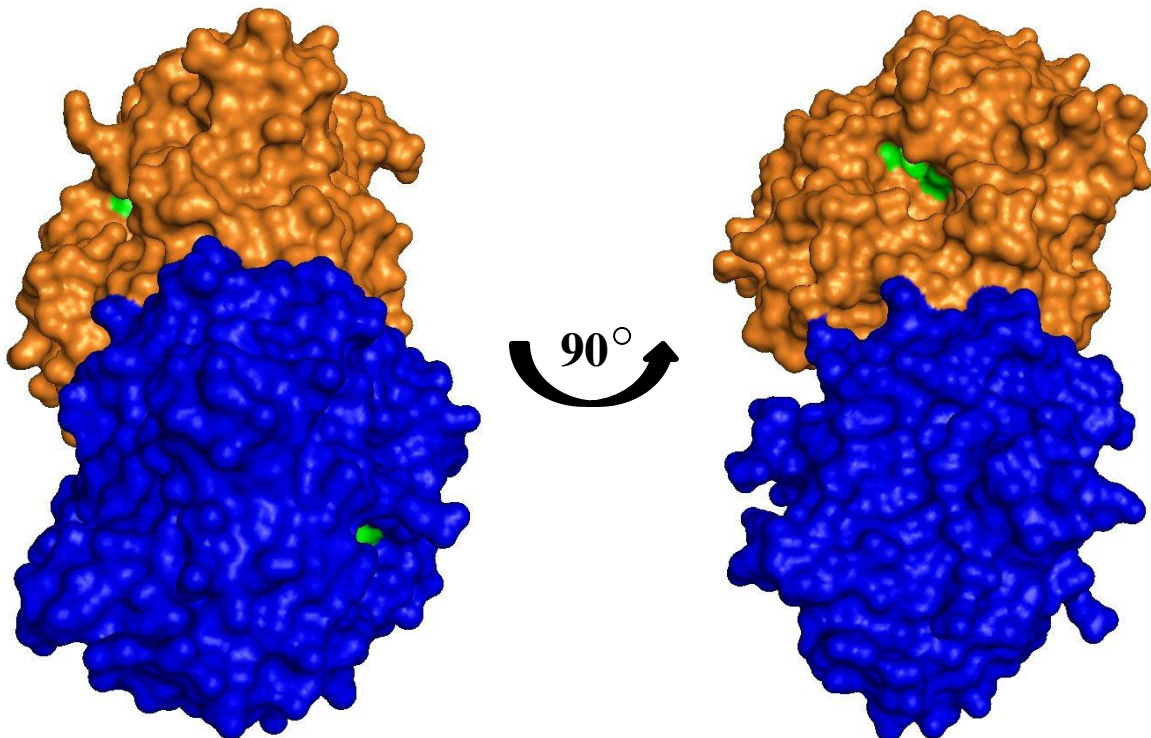
**A.** Evolution du facteur B le long de la séquence d'AlyA5. Les facteurs B sont représentés pour la chaîne A (jaune) et B (bleu). Les flèches rouges et vertes indiquent les deux pics mentionnés. **B.** Visualisation de la boucle flexible au-dessus du site actif de l'enzyme (en rouge), correspondant au pic de facteur B représenté par la flèche rouge dans le panel A.

Par ailleurs, AlyA5 existe sous deux formes en solution : un monomère et un dimère. La forme dimérique ayant été utilisée pour préparer les cristaux, une analyse de l'interface protéine-protéine au sein de la structure obtenue a été réalisée sur le serveur PISA. Il apparaît que l'interface entre les deux molécules cristallisées dans l'unité asymétrique est trop faible pour être stable en solution (Tableau 7-8). Ce dimère est donc artificiel et n'est pas la forme purifiée. En revanche, l'étude de l'interface entre deux chaînes appartenant à des molécules symétriques adjacentes montre que l'interface est deux fois plus importante. Cette forme correspond donc plus probablement au dimère formé en solution, où les molécules interagissent via une zone d'interface avec un cœur non-polaire. La visualisation de ce dimère montre que sous cette forme, les poches catalytiques des deux monomères pointent dans des directions opposées (Figure 7-28).

	Chaîne A et Chaîne B dans l'unité asymétrique	Chaîne A dans l'unité asymétrique Chaîne B dans la maille voisine
Aire d'interface totale	750 Å <sup>2</sup>	1526 Å <sup>2</sup>
% de la surface totale	3%	6%
Nombres de résidus impliqués	25	49
Aire d'interface non-polaire	602 Å <sup>2</sup>	970 Å <sup>2</sup>

**Tableau 7-8 : Analyse des interfaces protéine-protéine au sein du cristal d'AlyA5**

Les calculs ont été réalisés sur le serveur PISA ([http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html))



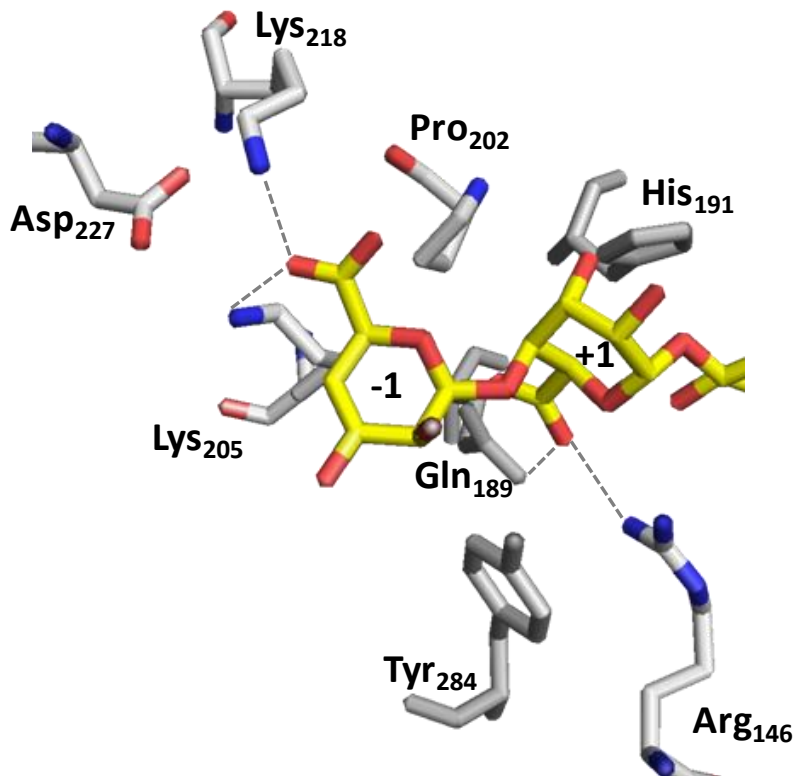
**Figure 7-28 : Visualisation du dimère d'AlyA5 le plus probable**

Les surfaces moléculaires des deux chaînes sont colorées en orange (chaîne A) et bleu (chaîne B). La position des sites catalytiques est matérialisée en vert. La vue de droite est obtenue après rotation de 90° autour d'un axe vertical.

### III-5-d Discussion sur la reconnaissance du substrat chez AlyA1 et AlyA5

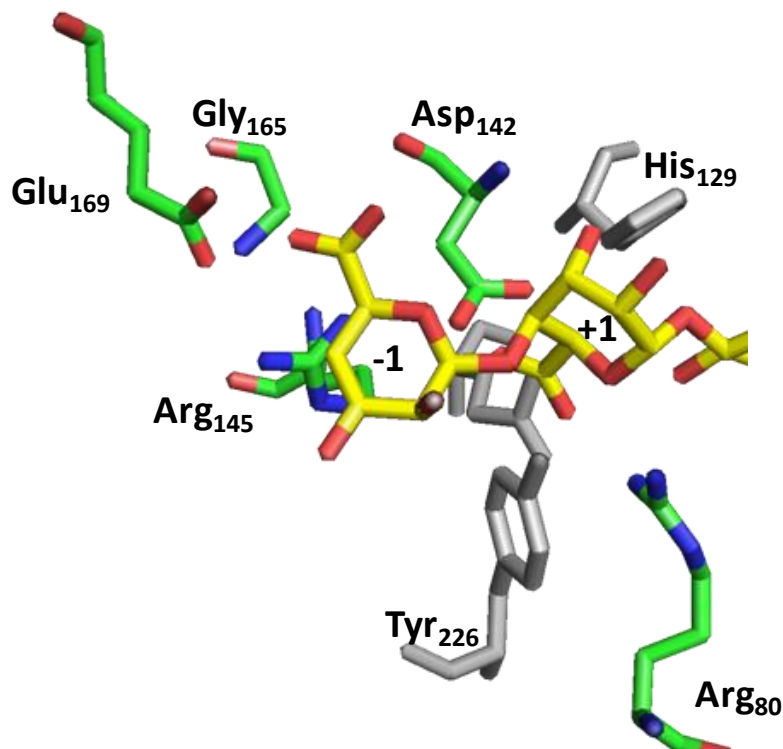
Afin d'identifier le mode de fixation du substrat dans le site catalytique des enzymes, le tétrasaccharide  $\Delta$ GGG provenant de la structure d'A1-II' de *Sphingomonas* sp. A1 en complexe a été superposé aux structures natives obtenues pour AlyA1 et AlyA5. La comparaison des sous-sites de fixation -1 et +1 donnent des indications intéressantes sur d'éventuelles différences dans le mode de reconnaissance du substrat.

Dans l'alginate lyase A1-II' de *Sphingomonas* sp. A1, les sucres en -1 et +1 sont stabilisés par des acides aminés chargés positivement (Figure 7-29). En plus de sa neutralisation par la glutamine catalytique Q189, le groupement carboxylate du sucre en +1 peut former des liaisons électrostatiques avec l'arginine R146. D'autre part, le groupement carboxylate du résidu en -1 est reconnu par les deux lysines K205 et K218.



**Figure 7-29 : Mode de reconnaissance du substrat chez A1-II' de *Sphingomonas* sp.A1**  
D'après Ogura *et al.* (PDB 2ZAA, 2008). Les résidus H191N et Y284F mutés pour obtenir le complexe ont été remplacés par les acides aminés de l'enzyme sauvage.

La superposition du tétrasaccharide dans la structure native d'AlyA1 de *Z. galactanivorans* montre que la position des résidus catalytiques Q127, H129, Y226 est bien conservée avec la structure d'A1-II' (Figure 7-30). De même, l'arginine R80 se positionne de manière similaire à R146 dans A1-II'. Cependant, cette superposition du substrat met en évidence une différence majeure. En effet, on constate la présence d'un acide aspartique en position centrale (D142) à la place de la proline dans A1-II' (P202). Dans cette visualisation, l'acide aspartique chargé négativement viendrait d'une part en conflit stérique et d'autre part en répulsion électrostatique avec le groupement carboxylate du sucre dans le sous-site +1. Cette observation montre que le sucre ne peut pas être reconnu de la même façon dans A1-II' de *Sphingomonas* sp. A1 que dans AlyA1 de *Z. galactanivorans*.



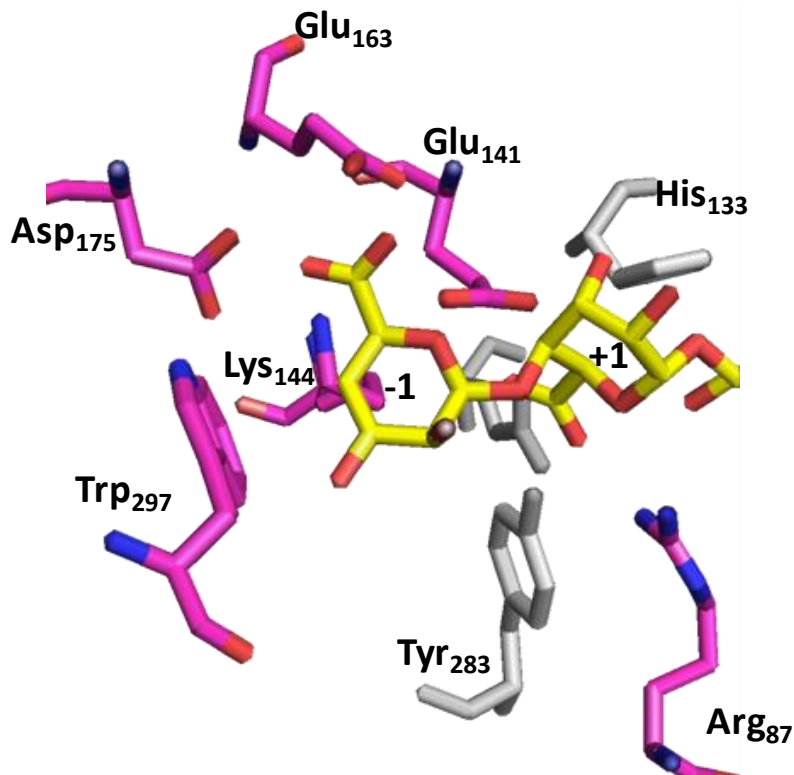
**Figure 7-30 : Modélisation du substrat dans le site catalytique d'AlyA1**

Le tétrasaccharide  $\Delta$ GGG provenant de la structure d'A1-II' de *Sphingomonas* sp.A1 a été superposé à la structure de l'enzyme AlyA1 native.

Pour AlyA5, on constate également que la position des acides aminés catalytiques Q131, H133 et Y283 est conservée, ainsi que celle de l'arginine R87 (Figure 7-31). Cependant, on peut remarquer deux particularités. D'une part, on constate la présence d'un tryptophane W297 juste en amont du sous-site -1. Ce tryptophane provient d'une des deux



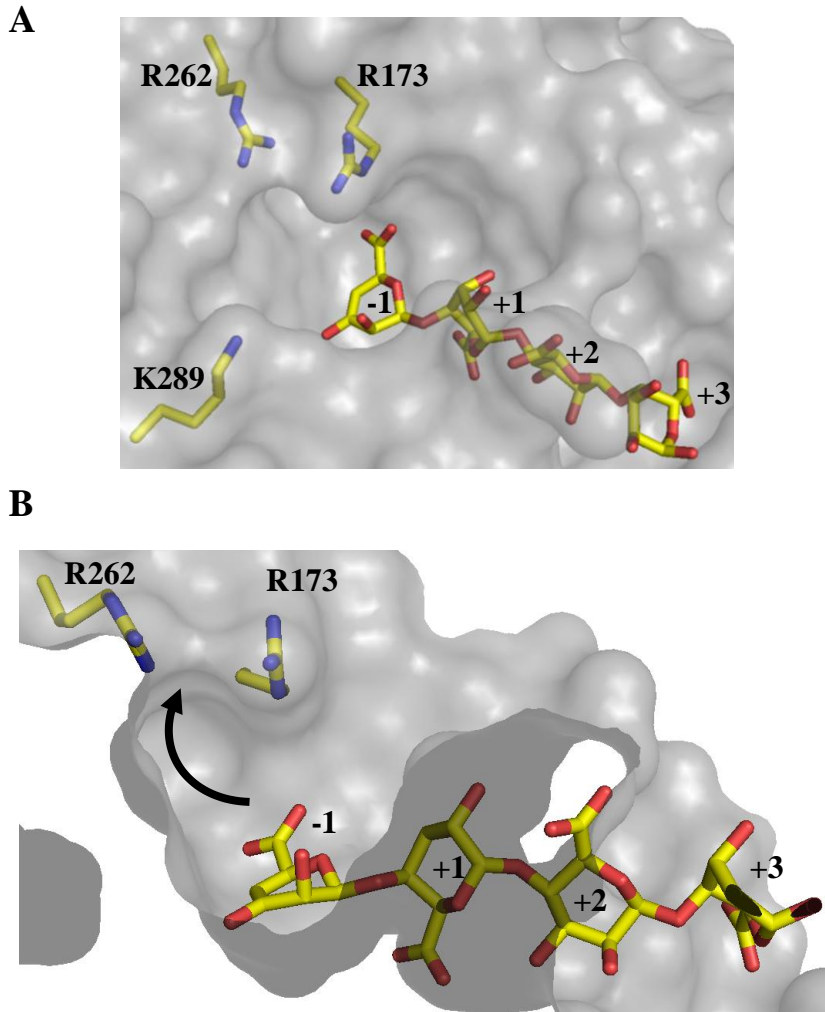
boucles additionnelles déjà mentionnées. Il vient littéralement obstruer la gorge de l'enzyme et empêche la fixation d'un sucre au-delà du sous-site -1, expliquant le mode d'action exolytique pour cette alginate lyase. D'autre part, trois résidus chargés négativement entourent les sous-sites -1 et +1 : les deux acides glutamiques E141 et 163 et l'acide aspartique D175. Ces trois acides aminés entrent en conflit stérique et répulsion électrostatique avec le sucre modélisé à partir de la structure de l'enzyme A1-II' de *Sphingomonas* sp. A1. Comme pour AlyA1, ceci suggère que la fixation du substrat dans le site actif d'AlyA5 est différente de chez A1-II'.



**Figure 7-31 : Modélisation du substrat dans le site catalytique d'AlyA5**

Le tétrasaccharide  $\Delta$ GGG provenant de la structure d'A1-II' de *Sphingomonas* sp.A1 a été superposé à la structure de l'enzyme AlyA5 native.

Par ailleurs, la vue de la surface d'AlyA5 montre qu'il existe une cavité juste au-dessus du sous-site -1. On peut remarquer que trois acides aminés chargés positivement (R173, R262, K289) sont positionnés à la sortie de cette cavité. De manière spéculative, on peut supposer que ces résidus attirent le sucre libéré lors du clivage de la liaison glycosidique. Ce mécanisme permettrait d'extraire rapidement le produit de la réaction du site catalytique et de libérer le sous-site -1 pour un clivage ultérieur.



**Figure 7-32 : Mode spéculatif d'extraction des produits de réaction dans AlyA5**

Vues de dessus (A) et en coupe (B) montrant les trois résidus chargés positivement positionnés au-dessus du sous-site -1. Le tétrasaccharide  $\Delta$ GGG provenant de la structure d'A1-II' de *Sphingomonas* sp.A1 a été superposé à la structure de l'enzyme AlyA5 native.

Contrairement à A1-II', l'activité d'AlyA1 et d'AlyA5 est dépendante du calcium. Ceci pourrait expliquer les différences dans le mode de fixation du substrat. En effet, une hypothèse envisageable serait la réorganisation des acides aminés chargés négativement autour d'un ion calcium lors de l'interaction avec le substrat. Cette réorganisation permettrait à la fois d'écranter les charges négatives et d'éviter l'encombrement stérique des sous-sites -1 et +1. Plus généralement, les différences dans la reconnaissance du substrat pourraient être liées au type d'alginate dégradé. En effet, l'enzyme A1-II' provient d'une bactérie du sol, *Sphingomonas* sp. A1. Les autres structures d'alginate lyases de la famille PL7 proviennent de *Corynebacterium* sp. et de *Pseudomonas aeruginosa* (Yamasaki *et al.*, 2004; Osawa *et al.*, 2005). Dans leur environnement naturel, ces espèces dégradent de l'alginate bactérien. Du fait de la présence de groupement acétyls, l'alginate bactérien a beaucoup moins tendance à



gélifier en présence de cations divalents. Au contraire, les enzymes d'origine marine AlyA1 et AlyA5 ont pour substrat naturel de l'alginate d'algues brunes. Celui-ci se retrouve principalement à l'état substrat semi-solide, formant une matrice compacte et organisée autour de cations divalents tels que le calcium dans le contexte de la paroi. L'obtention de ces deux premières structures d'alginate lyases de la famille PL7 d'origine marine suggère donc l'existence d'adaptations du mode de reconnaissance du substrat à la niche écologique.

#### IV- Conclusions

L'objet de ce chapitre était l'étude et la caractérisation des alginate lyases de la famille PL7 chez *Z. galactanivorans*. L'analyse phylogénétique a mis en évidence que les trois PL7 codées dans le génome de *Z. galactanivorans* appartiennent à des sous-familles différentes. AlyA1 et AlyA5 appartiennent respectivement aux sous-familles SF3 et SF5. L'alginate lyase AlyA2 fait partie d'une nouvelle sous-famille, qui ne contient pour l'instant que des enzymes d'origine marine. Sur ces trois enzymes, deux ont été exprimées de manière hétérologue dans *E. coli* et caractérisées. Bien qu'appartenant à la même famille CAZy, elles présentent des modes d'action différents. AlyA1 est une endo-guluronate lyase clivant les liaisons glycosidiques entre deux résidus G. AlyA5 est une exo-guluronate lyase attaquant le substrat du côté non réducteur et pouvant cliver les liaisons G-G et  $\Delta$ -G. L'étude cristallographique d'AlyA1 et AlyA5 a permis de résoudre pour la première fois les structures d'alginate lyases PL7 d'origine marine. Les déterminants structuraux du mode d'action ont été mis en évidence. AlyA1 adopte une topologie en gorge ouverte, compatible avec le mode d'action endolytique. Au contraire, deux insertions dans la séquence d'AlyA5 lui confèrent une topologie en poche qui explique son mode d'action exolytique. L'étude des deux structures suggère par ailleurs que le mode de reconnaissance du substrat est différent dans les enzymes d'origine marine ou terrestre. Pour déterminer plus précisément les modalités de fixation du substrat dans les sites catalytiques d'AlyA1 et AlyA5, il serait intéressant d'obtenir la structure des enzymes en complexe avec des oligosaccharides. Pour cela, l'inactivation des deux enzymes par mutation des acides aminés catalytiques doit être effectuée, en vue de la formation de cristaux en présence d'oligo-alginates.

# Conclusions et Perspectives générales

---

L'objectif de ce travail de thèse était d'identifier et de déchiffrer le système alginolytique chez la flavobactérie marine *Zobellia galactanivorans*. Au début de ce projet en janvier 2009, très peu de données étaient disponibles sur l'utilisation des polysaccharides d'algues brunes par cette espèce. Contrairement aux voies d'utilisation des polysaccharides d'algues rouges (agars, carraghénanes, porphyrane), celles de la dégradation des composés d'algues brunes n'avaient jamais été étudiées expérimentalement. Il était connu que *Z. galactanivorans* pouvait utiliser l'alginate comme seule source de carbone. Par ailleurs, l'annotation experte du génome fraîchement séquencé avait révélé des gènes codant des protéines potentiellement impliquées dans la dégradation de l'alginate. Ce projet est donc le premier à s'intéresser à cette voie métabolique chez *Z. galactanivorans*. Pour étudier cette question, j'ai conduit un travail pluri-disciplinaire, mêlant des approches de transcriptomique, de génomique comparative, de biochimie et de biologie structurale. L'intégration des différents résultats a permis de fournir les bases d'un modèle du système de perception et d'utilisation de l'alginate chez *Z. galactanivorans*.

## I- Principaux résultats

Au cours de ce travail, j'ai pu **identifier les acteurs du système alginolytique** de *Z. galactanivorans*. Pour cela, j'ai étudié l'expression des gènes prédits comme potentiellement impliqués dans la dégradation de l'alginate. Cette approche transcriptomique ciblée m'a permis de caractériser deux nouveaux opérons dédiés à cette voie catabolique. L'un d'entre eux constitue un véritable **Locus d'Utilisation de l'Alginate**, avec une organisation similaire à d'autres PUL (*Polysaccharide Utilization System*). Il code notamment des enzymes de dégradation, un régulateur transcriptionnel, un système SusCD-like, etc. L'expression des gènes est fortement induite en présence d'alginate dans le milieu, ce qui démontre que *Z. galactanivorans* est capable de sentir la nature de la source de carbone à sa disposition et d'adapter son métabolisme en fonction. Par ailleurs, j'ai découvert que le PUL dédié à l'alginate est conservé chez d'autres flavobactéries marines. De manière particulièrement intéressante, ce PUL a été partiellement transmis à des *Bacteroidetes* du microbiome intestinal et à des bactéries marines d'autres phyla (*Alpha-* et *Gammaproteobacteria*). L'existence de tels **transferts latéraux de gènes** apporte une vision nouvelle sur l'acquisition par les bactéries de capacités métaboliques au cours de l'évolution, en fonction de la biomasse à leur disposition. Elle soulève par ailleurs la question de l'émergence des premiers *Bacteroidetes* en

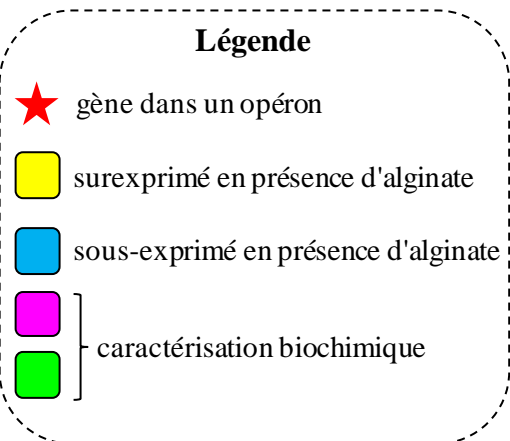
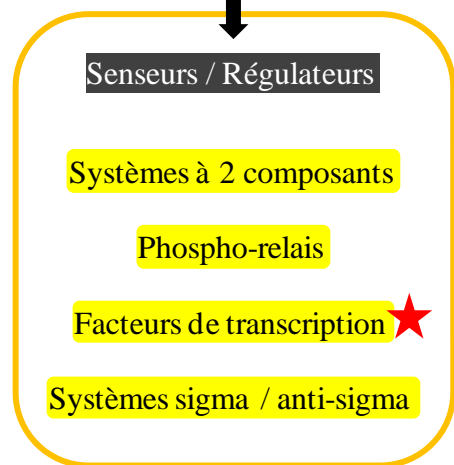
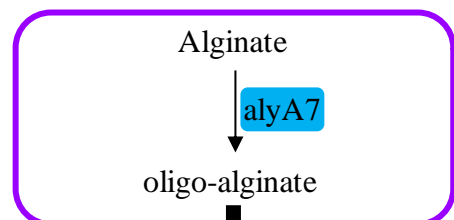
tant que symbiontes intestinaux des animaux. La vie étant apparue dans l'océan, on peut spéculer que la colonisation du tractus intestinal a commencé par l'installation de *Bacteroidetes* associés à des végétaux marins, ingérés par un protiste herbivore ancestral lors de phénomènes de broutage.

L'approche transcriptomique globale à l'échelle du génome m'a permis d'identifier de nouvelles **voies potentielles de perception et de transduction du signal** liées à la présence de l'alginate. Les résultats suggèrent également que lors de la dégradation d'alginate, *Z. galactanivorans* puisse se préparer à **résister aux réactions de défense** induites chez les algues attaquées. Cette découverte est pour l'instant préliminaire et reste à confirmer, mais elle pourrait être le signe d'une co-évolution étroite entre les algues et les bactéries avec lesquelles elles interagissent. Par ailleurs, les études transcriptomiques globales ont été menées en parallèle sur des cultures en présence de polysaccharides d'algues brunes et rouges. Les données indiquent qu'il pourrait exister deux "programmes" transcriptionnels différents, déclenchés en fonction du type de biomasse présent.

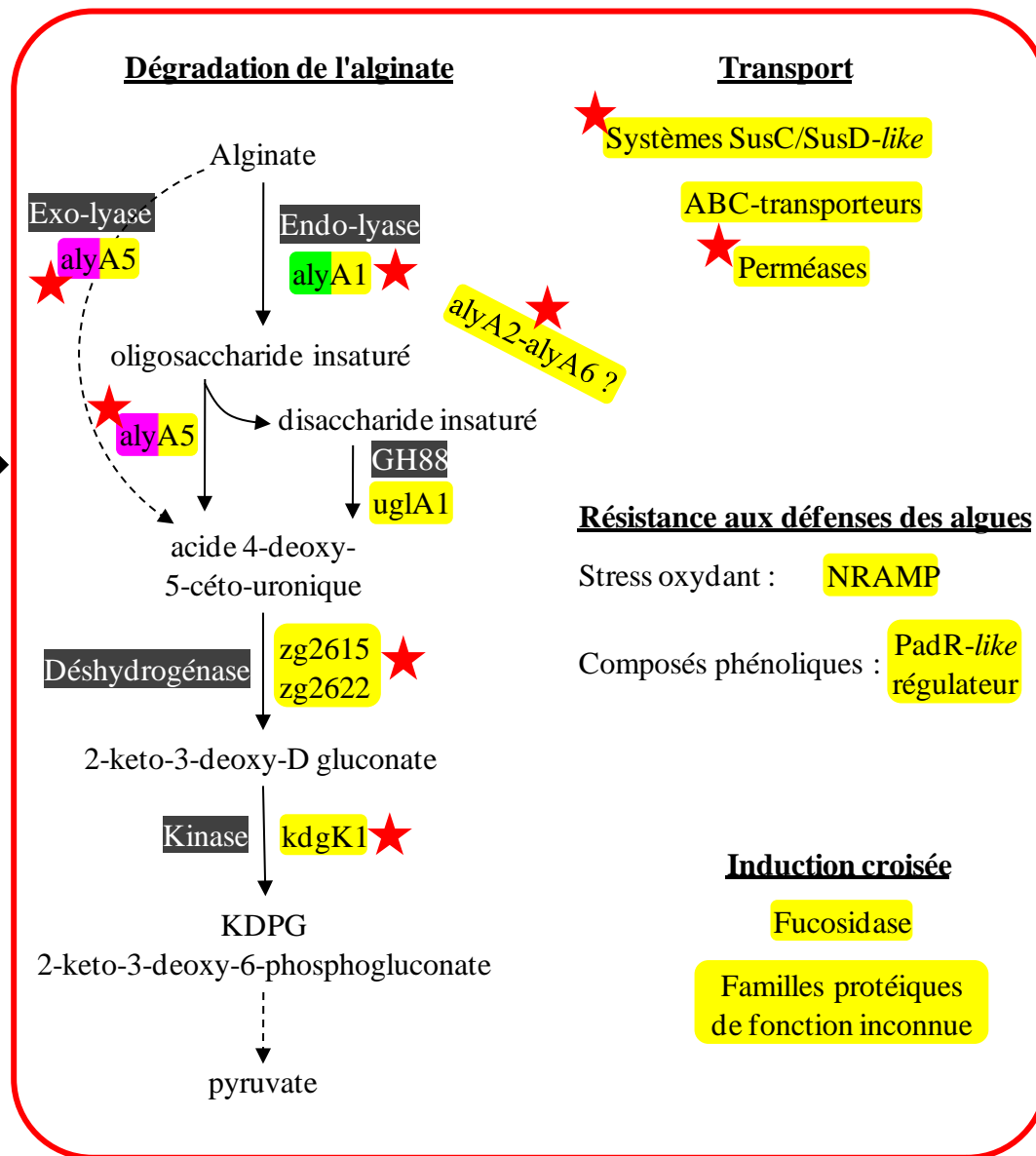
Au cours de ce travail, je me suis également intéressé dans le détail **aux alginate lyases de la famille PL7** présentes chez *Z. galactanivorans*. L'étude phylogénétique montre que l'une d'entre elles, AlyA2, appartient à une nouvelle sous-famille qui ne contient pour l'instant que des représentants d'origine marine. J'ai caractérisé biochimiquement deux nouvelles alginate lyases, AlyA1 et AlyA5. Bien qu'appartenant à la même famille CAZy, ces deux enzymes ont des modes d'action très différents. Les deux enzymes sont des **guluronate lyases**, mais **AlyA1 est endolytique**, alors qu'**AlyA5 est exolytique**. L'obtention des structures cristallographiques de ces deux enzymes a mis en évidence les déterminants expliquant ces différences. Les structures suggèrent de plus l'évolution de modes de reconnaissance contrastés entre les enzymes dont le substrat naturel est l'alginate d'algues et celles agissant sur l'alginate bactérien.

L'intégration des résultats provenant des différentes approches m'a permis de construire les bases d'un **modèle de dégradation de l'alginate** chez *Z. galactanivorans* (Figure 8-1).

### 1) Perception



### 2) Réponse alginolytique



**Figure 8-1 : Modèle hypothétique final du système alginolytique de *Z. galactanivorans***  
Ce modèle intègre les résultats issus des différentes approches conduites au cours de ce travail de thèse.

Ce modèle met en évidence les différentes phases impliquées dans la dégradation de l'alginate. Dans un premier temps, les bactéries doivent percevoir la présence du polysaccharide dans le milieu. Différents systèmes senseurs ont été identifiés comme répondant à l'alginate. Cette étape de perception nécessite la présence d'une molécule effectrice. Un effecteur potentiel pourrait être un oligo-alginate produit par l'alginate lyase constitutive AlyA7. Suite à la perception, des mécanismes de transduction du signal aboutissent au déclenchement de la réponse alginolytique proprement dite. Un grand nombre de gènes impliqués dans la voie de dégradation de l'alginate sont surexprimés. De manière intéressante, certains d'entre eux sont regroupés au sein de loci d'utilisation de l'alginate, constituant de vrais opérons. La caractérisation biochimique de deux alginate lyases impliquées dans la voie de dégradation a montré qu'elles peuvent agir de manière complémentaire. AlyA1 dégrade l'alginate et produit des oligosaccharides substrats d'AlyA5. La fonction précise des autres alginate lyases (AlyA2-AlyA6) reste à déterminer. Par ailleurs, la réponse alginolytique comporte l'expression de systèmes de transport qui permettraient l'internalisation des produits de dégradation dans le périplasma, puis dans le cytoplasme. Enfin, il semble que le déclenchement de la réponse alginolytique soit couplé à (i) une augmentation des capacités de résistance aux défenses des algues et (ii) à l'induction croisées de voies cataboliques d'autres composés présents chez les algues brunes.

## II- Perspectives

### II-1 Du pain sur la planche...

Le travail réalisé au cours de ma thèse ouvre de nombreuses perspectives à plus ou moins long terme. L'analyse des résultats provenant de l'approche transcriptomique à l'échelle du génome devra être finalisée. Notamment, la dissection des données obtenues en présence des polysaccharides d'algues rouges ou de laminarine pourrait mettre en évidence de nouvelles voies métaboliques et des systèmes de dégradation complets, comme cela a été le cas pour le système alginolytique. Les gènes impliqués seront de bons candidats pour un futur projet de clonage et d'expression à moyen débit, en vue de caractériser de nouvelles activités enzymatiques. L'analyse de l'expression des sondes couvrant les régions intergéniques aidera également à prédire l'existence de nouveaux opérons dans le génome. Par exemple, ces résultats sont déjà utilisés par Aurore Labourel, qui étudie au cours de sa thèse dans notre équipe le métabolisme de la laminarine chez *Z. galactanivorans*. Lors des expériences

réalisées au cours de ma thèse, les bactéries ont été cultivées en présence de polysaccharides purifiés. Afin de se rapprocher des conditions naturelles, il serait particulièrement intéressant de tester l'effet d'extraits de polysaccharides totaux d'algues rouges ou brunes, contenant un mélange de composés pariétaux.

La stratégie de clonage et d'expression hétérologue à moyen débit que j'ai menée a abouti à l'obtention de 17 protéines recombinantes solubles. J'ai pu en caractériser deux d'entre elles au cours de ma thèse, AlyA1 et AlyA5. La production et l'étude des autres protéines pourront être envisagées, afin de compléter la compréhension du modèle proposé ci-dessus. Par exemple, des premiers tests enzymatiques ont déjà été réalisés sur la protéine recombinante AlyA7. Ils confirment qu'il s'agit bien d'une alginate lyase.

Enfin, l'étude structurale d'AlyA1 et AlyA5 a suggéré l'existence d'un mode original de fixation du substrat. Pour confirmer cette hypothèse et élucider le type d'interactions, il serait crucial de résoudre la structure des enzymes en complexe avec des oligosaccharides. Dans ce but, la production de protéines recombinantes inactives mutées sur les acides aminés catalytiques pourra être entreprise en vue de l'obtention de cristaux en présence de substrats.

## **II-2 Le potentiel biotechnologique de *Z. galactanivorans***

Les résultats acquis au cours de cette thèse permettent de mettre en avant le potentiel biotechnologique de *Z. galactanivorans*.

On peut envisager l'utilisation des alginate lyases recombinantes AlyA1 et AlyA5 issues de cette bactérie marine comme des outils moléculaires avec diverses applications. Grâce à leur caractérisation biochimique détaillée, on connaît aujourd'hui précisément leur mode d'action et les motifs qu'elles clivent dans une séquence d'alginate. En identifiant les produits de dégradation obtenus à partir d'alginate de séquence inconnue, on peut donc déduire des informations sur l'agencement des résidus mannuronates et guluronates dans le polysaccharide de départ (Ostgaard 1993; Heyraud *et al.*, 1996). Une telle utilisation des alginate lyases décrites dans cette thèse en tant qu'outils de "séquençage" est d'ailleurs l'objet du programme européen PolyModE (*Polysaccharide Modifying Enzymes*, FP7) auquel notre équipe prend part. Dans le cadre de ce projet, les nouvelles alginate lyases seront employées pour élucider la structure de molécules d'alginate après l'action de

mannuronane C5-épimérase, qui convertissent les résidus M en G. Les alginate lyases pourront également être employées pour la production d'oligosaccharides de motifs et de taille connus, en vue de tester d'éventuels effets biologiques sur des plate-forme de criblage (F. Le Sourd, communication personnelle).

Les *Carbohydrate Binding Modules* (CBM) retrouvés dans certaines des protéines de l'opéron alginolytique sont également des cibles ayant une application biotechnologique intéressante. En effet, de tels CBM sont utilisés en biologie végétale pour étudier l'architecture de la paroi des plantes (McCartney *et al.*, 2006). Les CBM sont employés comme des sondes spécifiques pour visualiser la localisation des polysaccharides directement sur des coupes des végétaux, par des techniques d'immunohistochimie. Des CBM dirigés contre des polysaccharides d'algues pourraient donc être utilisés de la même façon. Cette approche est actuellement mise au point par Cécile Hervé, chercheur dans notre équipe. Le CBM putatif obtenu sous forme recombinante soluble lors de mon travail (zg2619, module 2) a déjà été produit à grande échelle et purifié. Le clonage du fragment génomique codant le CBM associé à l'alginate lyase AlyA1 a échoué lors de l'approche à moyen débit. Il pourrait être envisagé de retenter un clonage spécifiquement pour ce fragment, afin d'obtenir l'expression de la protéine recombinante. La caractérisation de la spécificité de reconnaissance de ces deux CBM permettra ensuite de les utiliser comme sonde pour étudier l'architecture de la paroi chez les algues brunes, qui est à ce jour moins bien connue que chez les plantes terrestres.

Enfin, on peut évoquer une application en termes de valorisation de la biomasse algale. Une étape-clé de l'utilisation de cette biomasse est la conversion des polysaccharides pariétaux en sucres simples, un processus nommé saccharification. Ces monosaccharides peuvent ensuite être utilisés comme substrats par des levures réalisant la fermentation alcoolique. Ce procédé a récemment été utilisé pour produire de l'éthanol à partir de l'algue brune *Saccharina japonica*, en présence d'une bactérie dégradant l'alginate (*Erwinia tasmaniensis*) et de la levure *Candida lusitaniae* (Lee and Lee 2011). J'ai pu montrer lors de mon travail que *Z. galactanivorans* possède une voie de dégradation complète de l'alginate, aboutissant à la formation de monosaccharides. Par ailleurs, *Z. galactanivorans* étant capable d'utiliser la laminarine comme seule source de carbone, on peut penser qu'elle dégrade également ce composé en glucose. Par conséquent, il semble possible d'employer *Z. galactanivorans* comme agent de saccharification de la biomasse des algues brunes, en vue de la production d'éthanol. On



pense bien sûr à l'utilisation de ce processus pour obtenir du biofuel, mais pourquoi ne pas tenter de créer une nouvelle gamme de bières d'algues ?

### II-3 *Z. galactanivorans*, un nouvel organisme dégradateur modèle

Le séquençage du génome de *Z. galactanivorans* a révélé son formidable arsenal enzymatique de dégradation des polysaccharides marins. Un nombre croissant d'études, dont celle-ci, rapporte la caractérisation biochimique de quelques unes de ses glycosidases. L'organisation des gènes permet également de prédire un grand nombre de PUL, potentiellement impliqués dans l'utilisation de polysaccharides d'algues rouges ou brunes. La caractérisation expérimentale du premier d'entre eux, le PUL dédié à l'alginate, ouvre la voie à des études similaires pour d'autres substrats. D'autre part, la bactérie semble bien équipée pour résister aux défenses des algues. *Z. galactanivorans* est donc bien placée pour devenir un nouvel organisme modèle dans l'étude des interactions algues-bactéries, en particulier en ce qui concerne la colonisation des surfaces algales et la dégradation de la biomasse.

Durant ma thèse, j'espère avoir contribué à l'essor de ce nouvel "*Erwinia de la mer*" en mettant au point et en développant de nouvelles techniques dédiées à l'étude de l'expression des gènes. Ainsi, les méthodes de RT-qPCR et de Northern Blot non-radioactif sont aujourd'hui disponibles pour *Z. galactanivorans*. La validation des gènes de normalisation pour la RT-qPCR fournit un protocole fiable et rapidement mis en œuvre de mesure du taux de transcription de gènes d'intérêt sélectionnés. Pour des études plus globales à l'échelle du génome, la puce pangénomique mise au point au cours de ma thèse est un outil de choix. Elle permettra notamment de suivre les voies métaboliques sur- ou sous-exprimées dans diverses conditions. Cette approche où tous les gènes sont considérés sans *a priori* pourra faciliter l'identification de cibles impliquées dans une voie donnée, alors même que l'annotation n'a pas pu prédire leur fonction.

Une des techniques majeures pour l'analyse fonctionnelle de voies métaboliques est l'étude de mutants pour un ou plusieurs des gènes mis en jeu. Dans ce cadre, un projet visant à développer la mutagenèse chez *Z. galactanivorans* a été initié dans notre équipe par Lionel Cladière. Les premiers mutants aléatoires ont été obtenus récemment, et leur caractérisation phénotypique et génétique est en cours (L. Cladière et D. Goudier, communication personnelle). Suite à ces premiers essais encourageants, il a été décidé d'adapter un protocole de mutagenèse dirigée, qui permettra de cibler le gène à invalider. Pour tester ce protocole,

trois gènes du système alginolytique de *Z. galactanivorans* ont été sélectionnés comme candidats pour la mutation : zg1182 (*alyAI*) zg2617 (codant le régulateur transcriptionnel GntR-like) et zg2620 (*susD-like*). Si de tels mutants sont obtenus, leur étude fournira de précieux éléments sur le rôle fonctionnel de ces gènes dans la perception et l'utilisation de l'alginate chez *Z. galactanivorans*.

Les perspectives ouvertes à l'issue de mon travail de thèse sont donc nombreuses et variées. L'essor des techniques dans l'ère de la post-génomique permet aujourd'hui d'envisager des études intégratives telles que celle que j'ai conduite, voire de les compléter par d'autres approches à développer. La confrontation des différents résultats permet de bâtir une vue d'ensemble d'une voie métabolique et d'identifier de nouvelles cibles à étudier. Bien sûr, les questions restent nombreuses et l'étude de *Z. galactanivorans* en tant qu'organisme modèle de la dégradation de végétaux marins n'en est qu'à son début. Qu'on se le dise !



# Annexe 1

## Gènes surexprimés en présence d'alginate

---

**Annexe 1 :**  
Gènes surexprimés en présence d'alginate par rapport au glucose  
Critères : FC > 2 et q-value < 0,05

Gene ID	Fold-change	q-value	Annotation
zobellia_18	2.68	0.02	Putative membrane protein
zobellia_19	3.44	0.01	Conserved hypothetical protein
zobellia_28	2.75	0.02	Two-component system - Response regulator
zobellia_43	2.34	0.02	tRNA guanosine-2'-O-methyltransferase
zobellia_54	2.54	0.05	Carboxy-terminal processing peptidase, family S41
zobellia_75	3.24	0.01	Cardiolipin synthetase
zobellia_87	2.09	0.02	RNA polymerase ECF-type sigma factor
zobellia_89	2.00	0.02	Conserved hypothetical protein
zobellia_93	5.73	0.02	Conserved hypothetical periplasmic protein
zobellia_106	2.40	0.01	SAM-dependent methyltransferase
zobellia_115	3.23	0.04	Bacterial surface antigen
zobellia_143	2.60	0.02	Putative membrane lipoprotein
zobellia_176	2.06	0.05	Conserved hypothetical periplasmic protein
zobellia_177	2.73	0.03	Cytochrome c-containing protein
zobellia_223	2.14	0.01	ABC-2 type exporter, permease component
zobellia_269	2.44	0.02	One-component system sensor protein
zobellia_324	2.52	0.04	Creatinase family protein
zobellia_332	2.39	0.04	Mucin-desulfating sulfatase, family S1-11
zobellia_358	2.76	0.04	anti-sigma factor
zobellia_390	3.66	0.02	Putative protein
zobellia_515	4.76	0.02	Major facilitator family transporter
zobellia_538	2.22	0.01	RDD family membrane protein
zobellia_631	3.21	0.04	Thrombospondin type 3 repeats protein
zobellia_632	11.24	0.01	Conserved hypothetical protein
zobellia_633	11.54	0.01	Thrombospondin type 3 repeats protein
zobellia_634	10.66	0.02	Conserved hypothetical protein
zobellia_675	3.10	0.04	Conserved hypothetical protein
zobellia_685	2.32	0.04	Two-component system - Sensor histidine kinase
zobellia_686	7.15	0.02	Two-component system - Response regulator, receiver domain
zobellia_701	2.02	0.03	TetR-type transcriptional regulator
zobellia_704	4.59	0.04	Conserved PDZ-containing protein
zobellia_734	2.16	0.00	Conserved hypothetical periplasmic protein
zobellia_737	2.80	0.01	Ankyrin repeats protein
zobellia_754	3.35	0.01	Ribosomal RNA small subunit methyltransferase G
zobellia_780	2.15	0.02	Zinc-transporting P-type ATPase
zobellia_831	2.43	0.02	hypothetical protein
zobellia_890	3.27	0.02	Conserved hypothetical periplasmic protein
zobellia_904	2.33	0.03	Protein TolR
zobellia_920	4.60	0.03	AraC-type transcriptional regulator
zobellia_932	2.13	0.03	OmpA-like periplasmic protein
zobellia_934	2.46	0.03	Serine peptidase, family S8
zobellia_951	2.21	0.03	Conserved hypothetical lipoprotein
zobellia_1132	3.59	0.04	Conserved hypothetical protein.
zobellia_1133	4.38	0.04	Conserved cytochrome c -containing protein
zobellia_1144	2.15	0.01	Receptor-like protein kinase
zobellia_1164	3.31	0.04	Metal transport-related protein
zobellia_1182	3.83	0.01	Alginate lyase, family PL7
zobellia_1186	2.01	0.02	Lipid-A-disaccharide synthase, family GT19
zobellia_1214	2.23	0.04	Formylmethionine deformylase
zobellia_1292	2.11	0.02	Conserved hypothetical protein
zobellia_1313	3.65	0.03	Saccharopine dehydrogenase
zobellia_1360	2.75	0.01	Conserved hypothetical membrane protein
zobellia_1381	2.37	0.02	Conserved hypothetical membrane protein
zobellia_1428	2.22	0.02	Two-component system - Sensor histidine kinase
zobellia_1445	3.36	0.01	5-Formyltetrahydrofolate cyclo-ligase
zobellia_1475	3.13	0.03	Conserved hypothetical lipoprotein
zobellia_1479	2.48	0.04	Conserved hypothetical membrane protein
zobellia_1481	2.95	0.03	Hypothetical protein
zobellia_1483	2.19	0.02	Conserved hypothetical membrane protein
zobellia_1515	2.36	0.03	CYTH family protein
zobellia_1595	2.59	0.01	PAS fold sensor protein
zobellia_1619	6.02	0.04	Conserved hypothetical lipoprotein
zobellia_1672	2.32	0.01	Conserved hypothetical protein
zobellia_1681	2.56	0.03	Carboxy-terminal processing peptidase, family S41
zobellia_1688	2.32	0.02	Conserved hypothetical protein
zobellia_1704	2.95	0.03	Conserved hypothetical protein
zobellia_1705	2.20	0.01	Putative protein
zobellia_1706	4.28	0.01	Conserved hypothetical protein
zobellia_1707	2.64	0.05	Conserved hypothetical periplasmic protein
zobellia_1708	2.94	0.01	Conserved hypothetical periplasmic protein
zobellia_1715	2.14	0.04	Conserved hypothetical membrane protein
zobellia_1768	11.24	0.00	PadR-type transcriptional regulator
zobellia_1769	8.90	0.05	Conserved hypothetical membrane protein

Annexe 1 : Gènes surexprimés en présence d'alginate

zobellia_1770	2.66	0.04	Conserved hypothetical membrane protein
zobellia_1773	2.24	0.04	Gliding motility membrane lipoprotein SprE
zobellia_1806	4.01	0.02	Gliding motility membrane lipoprotein
zobellia_1834	9.92	0.04	Conserved hypothetical protein
zobellia_1851	6.21	0.02	Conserved hypothetical lipoprotein
zobellia_1869	2.40	0.04	Conserved hypothetical periplasmic protein
zobellia_1873	2.89	0.03	Conserved hypothetical periplasmic protein
zobellia_1878	4.04	0.01	Putative exported protein
zobellia_1885	2.46	0.04	Putative protein
zobellia_1888	2.50	0.03	Deoxycytidylate deaminase
zobellia_1909	2.67	0.03	Conserved hypothetical periplasmic protein
zobellia_1930	2.58	0.02	Conserved hypothetical lipoprotein
zobellia_1956	2.54	0.03	Two-component system - Sensor histidine kinase
zobellia_1990	7.54	0.00	Conserved hypothetical lipoprotein
zobellia_2022	4.03	0.03	Conserved hypothetical periplasmic protein
zobellia_2023	5.48	0.02	Conserved hypothetical protein
zobellia_2024	3.82	0.03	OmpA-like outer membrane protein
zobellia_2029	3.96	0.01	Conserved hypothetical protein
zobellia_2034	2.26	0.04	Conserved hypothetical protein
zobellia_2099	2.01	0.00	Conserved hypothetical protein
zobellia_2132	2.47	0.03	Putative protein
zobellia_2184	2.10	0.04	Conserved protein containing type I restriction enzyme R N-terminal domain
zobellia_2224	2.16	0.04	LacI-type transcriptional regulator
zobellia_2226	2.00	0.02	Conserved hypothetical protein
zobellia_2269	2.86	0.02	BlaI-type transcriptional repressor
zobellia_2283	5.19	0.04	conserved hypothetical protein
zobellia_2284	3.43	0.02	Conserved hypothetical membrane protein
zobellia_2356	2.38	0.01	Sodium/hydrogen exchanger
zobellia_2401	3.78	0.02	Conserved hypothetical protein
zobellia_2454	2.07	0.01	CBS domain protein
zobellia_2456	2.99	0.02	Conserved hypothetical protein
zobellia_2491	5.50	0.01	Conserved hypothetical protein
zobellia_2495	6.29	0.05	Rhomboid-like protease, family S54
zobellia_2568	3.01	0.02	Putative membrane lipoprotein
zobellia_2607	3.36	0.01	Conserved hypothetical periplasmic protein
zobellia_2612	4.51	0.05	Manganese/fer transporter
zobellia_2613	5.76	0.02	Diphosphate-fructose-6-phosphate 1-phosphotransferase
zobellia_2614	3.64	0.01	2-dehydro-3-deoxygluconokinase
zobellia_2615	3.10	0.01	Acetoin(diacetyl) reductase
zobellia_2616	5.89	0.01	Sugar permease
zobellia_2618	19.13	0.02	Alginate lyase, family PL7
zobellia_2619	8.53	0.00	Conserved hypothetical membrane protein
zobellia_2620	9.53	0.01	SusD/RagB family lipoprotein
zobellia_2621	16.69	0.01	TonB-dependent Receptor
zobellia_2622	3.06	0.03	Short-chain dehydrogenase/reductase
zobellia_2623	5.45	0.01	Pectin degradation protein KdgF
zobellia_2624	11.39	0.03	Alginate lyase, family PL17
zobellia_2626	2.74	0.01	Xylose isomerase-like TIM barrel protein
zobellia_2627	3.79	0.01	Conserved hypothetical protein
zobellia_2628	4.37	0.02	Conserved hypothetical protein
zobellia_2634	2.53	0.01	Conserved hypothetical protein
zobellia_2636	5.81	0.04	Putative membrane protein
zobellia_2637	2.15	0.01	Gluconokinase
zobellia_2638	6.26	0.03	5-dehydro-2-deoxygluconokinase
zobellia_2673	2.30	0.05	SusD/RagB family lipoprotein
zobellia_2716	8.70	0.02	Conserved hypothetical membrane protein
zobellia_2747	7.50	0.02	Conserved hypothetical lipoprotein
zobellia_2761	6.00	0.04	Conserved hypothetical periplasmic protein
zobellia_2762	5.73	0.01	Conserved hypothetical periplasmic protein
zobellia_2795	2.04	0.01	Rieske [2Fe-2S] iron-sulfur protein
zobellia_2796	2.96	0.03	Thioredoxin family protein
zobellia_2797	3.69	0.04	Thiamine biosynthesis protein apbE
zobellia_2805	2.26	0.01	Conserved hypothetical protein
zobellia_2816	4.08	0.01	Protein CrcB
zobellia_286	4.54	0.01	Conserved hypothetical protein
zobellia_2882	4.85	0.01	Conserved hypothetical protein
zobellia_2884	2.47	0.03	Putative protein
zobellia_2885	2.34	0.02	Putative protein
zobellia_2887	3.02	0.01	Putative protein
zobellia_2888	5.21	0.01	Conserved hypothetical protein
zobellia_2895	5.05	0.03	Conserved hypothetical protein
zobellia_2983	2.82	0.01	Histidine-containing phosphotransfer protein
zobellia_2984	2.94	0.02	Two-component system - Response regulator
zobellia_2988	2.56	0.02	Conserved hypothetical membrane protein

## Annexe 1 : Gènes surexprimés en présence d'alginate

zobellia_2995	2.77	0.04	DSBA thioredoxin family protein
zobellia_2996	2.13	0.02	Conserved hypothetical membrane protein
zobellia_3000	2.34	0.02	Conserved hypothetical protein
zobellia_3017	2.22	0.03	Conserved hypothetical periplasmic protein
zobellia_3032	4.94	0.02	Hypothetical periplasmic protein
zobellia_3033	6.26	0.02	Hypothetical periplasmic protein
zobellia_3158	2.18	0.02	High-affinity gluconate transporter
zobellia_3187	3.24	0.02	Integron-related Integrase
zobellia_3239	8.50	0.01	Conserved hypothetical protein
zobellia_3243	2.07	0.04	TonB-dependent Receptor
zobellia_3255	2.05	0.03	Conserved hypothetical lipoprotein
zobellia_3344	2.23	0.04	TonB-dependent Receptor
zobellia_3350	2.50	0.04	Sodium/solute symporter
zobellia_3353	3.93	0.04	Serine peptidase, family S12
zobellia_3363	3.27	0.02	Conserved hypothetical membrane protein
zobellia_3399	2.62	0.02	Putative protein
zobellia_3446	2.00	0.02	Carbohydrate esterase, family CE10
zobellia_3464	2.11	0.04	Alpha-L-fucosidase, family GH29
zobellia_3583	2.99	0.04	Two-component system - Response regulator, receiver domain
zobellia_3610	2.36	0.03	Carbon starvation protein A
zobellia_3612	2.23	0.05	Possible FAD dependent oxidoreductase
zobellia_3626	2.33	0.04	Conserved hypothetical periplasmic protein
zobellia_3647	4.11	0.02	Hypothetical protein
zobellia_3648	3.03	0.01	Conserved hypothetical protein
zobellia_3649	2.94	0.02	OmpA-like periplasmic protein
zobellia_3650	6.67	0.04	Conserved hypothetical protein
zobellia_3651	2.95	0.01	Conserved hypothetical membrane protein
zobellia_3652	2.94	0.04	OmpA-like periplasmic protein
zobellia_3727	2.21	0.04	Conserved hypothetical protein
zobellia_3751	2.19	0.04	Hypothetical membrane protein
zobellia_3770	2.13	0.04	Pectate lyase, family PL1
zobellia_3789	2.22	0.02	Conserved hypothetical protein
zobellia_3922	2.88	0.03	Plasmid stabilisation system protein
zobellia_3971	2.34	0.00	Deoxyribodipyrimidine photo-lyase
zobellia_3973	2.16	0.04	Conserved hypothetical protein
zobellia_3988	3.12	0.03	SusD/RagB family lipoprotein
zobellia_4003	9.05	0.01	Two-component system - Sensor histidine kinase
zobellia_4009	2.33	0.01	AraC-type transcriptional regulator
zobellia_4025	3.05	0.02	1,4-dihydroxy-2-naphthoate octaprenyltransferase
zobellia_4057	2.88	0.01	Conserved hypothetical membrane protein
zobellia_4071	3.61	0.00	RNA polymerase ECF-type sigma factor
zobellia_4102	2.03	0.05	Conserved hypothetical protein
zobellia_4130	6.60	0.03	Alginate lyase, family PL6
zobellia_4131	10.07	0.00	Alginate lyase, family PL7
zobellia_4132	7.82	0.00	Alginate lyase, family PL6
zobellia_4133	2.31	0.02	Two-component system - Sensor histidine kinase
zobellia_4138	2.14	0.01	Conserved hypothetical membrane protein
zobellia_4151	3.69	0.02	Histidine-containing phosphotransfer protein
zobellia_4152	4.08	0.03	Response regulator receiver
zobellia_4156	3.02	0.04	Conserved hypothetical periplasmic protein
zobellia_4195	3.57	0.01	Conserved hypothetical protein
zobellia_4196	4.06	0.03	GCN5-related N-acetyltransferase
zobellia_4225	2.26	0.02	Conserved hypothetical membrane protein
zobellia_4228	2.05	0.01	Intracellular peptidase, family C56
zobellia_4267	2.22	0.02	Beta-agarase C, family GH16
zobellia_4295	2.52	0.04	Two-component system - Sensor histidine kinase
zobellia_4340	3.29	0.01	Conserved hypothetical periplasmic protein
zobellia_4341	8.53	0.04	ABC-2 type exporter, permease component
zobellia_4342	2.19	0.05	ABC-2 type exporter, permease component
zobellia_4350	2.08	0.04	Conserved hypothetical protein
zobellia_4388	2.20	0.04	Conserved hypothetical protein
zobellia_4389	2.06	0.05	Aromatic-ring hydroxylase, FAD-dependent oxidoreductase family
zobellia_4403	2.29	0.02	Proton/sodium-glutamate symporter
zobellia_4407	2.23	0.02	Conserved hypothetical periplasmic protein
zobellia_4436	3.55	0.02	AsnC/Lrp-type transcriptional regulator
zobellia_4437	3.22	0.02	Conserved hypothetical protein
zobellia_4461	3.80	0.02	Serine peptidase, family S12
zobellia_4466	2.03	0.03	Glyoxalase superfamily protein
zobellia_4472	2.38	0.04	Conserved hypothetical membrane protein
zobellia_4575	2.28	0.01	Conserved hypothetical protein
zobellia_4576	4.67	0.02	Conserved hypothetical membrane protein
zobellia_4584	3.47	0.04	Genetic Element Protein A
zobellia_4585	2.75	0.01	Hypothetical membrane protein
zobellia_4588	4.08	0.01	Putative protein

---

Annexe 1 : Gènes surexprimés en présence d'alginate

---

zobellia_4600	2.21	0.02	Acyl-homoserine lactone acylase, family S45
zobellia_4664	4.91	0.05	AraC-type transcriptional regulator
zobellia_4729	3.13	0.01	Conserved hypothetical lipoprotein
zobellia_4761	2.11	0.01	Conserved hypothetical protein
zobellia_4765	2.87	0.02	Conserved hypothetical periplasmic protein





## Annexe 2

# Gènes sous-exprimés en présence d'alginate

---

**Annexe 2 :**  
Gènes sous-exprimés en présence d'alginate par rapport au glucose  
Critères : FC < 0,5 et q-value < 0,05

Gene ID	Fold-change	q-value	Annotation
zobellia_0	0.47	0.02	Chromosomal replication initiator protein, DnaA
zobellia_15	0.24	0.03	Isocitrate dehydrogenase NADP-dependent, monomeric type
zobellia_20	0.40	0.03	Metallopeptidase, family M42
zobellia_22	0.37	0.02	Histidyl-tRNA synthetase
zobellia_26	0.36	0.05	Ribosomal protein S18
zobellia_34	0.30	0.01	Phosphoserine phosphatase / D-3-phosphoglycerate dehydrogenase
zobellia_40	0.20	0.04	RNA helicase CsdA
zobellia_46	0.42	0.02	Adenylosuccinate lyase
zobellia_66	0.41	0.04	Glucose-6-phosphate 1-dehydrogenase
zobellia_67	0.30	0.01	6-Phosphogluconate dehydrogenase, decarboxylating
zobellia_85	0.23	0.00	Phosphoglucomutase
zobellia_91	0.33	0.01	Ribosomal RNA small subunit methyltransferase B
zobellia_102	0.50	0.03	Major facilitator family transporter
zobellia_103	0.45	0.04	DEAD/DEAH RNA helicase
zobellia_108	0.36	0.01	Conserved hypothetical lipoprotein
zobellia_109	0.36	0.01	Di-heme cytochrome c peroxidase
zobellia_110	0.14	0.05	Conserved hypothetical periplasmic protein
zobellia_122	0.49	0.04	5-Methyltetrahydropteroyltriglutamate-homocysteine methyltransferase
zobellia_124	0.40	0.03	RNA helicase RhIE
zobellia_153	0.45	0.01	Two-component system - Response regulator
zobellia_154	0.25	0.01	Galactokinase
zobellia_155	0.47	0.03	C-alpha-formylglycine-generating protein
zobellia_171	0.31	0.04	Transaldolase
zobellia_173	0.38	0.03	Xylulose kinase
zobellia_202	0.49	0.03	Conserved hypothetical protein
zobellia_245	0.18	0.01	Ribonucleoside-diphosphate reductase, small subunit
zobellia_279	0.46	0.05	Short-chain dehydrogenase/reductase
zobellia_301	0.50	0.04	Conserved hypothetical membrane protein
zobellia_305	0.39	0.04	Sulfite reductase [NADPH] flavoprotein alpha-component
zobellia_306	0.39	0.05	Peroxiredoxin
zobellia_369	0.11	0.04	Alpha-amylase, family GH13
zobellia_370	0.17	0.01	Starch utilization lipoprotein
zobellia_371	0.11	0.01	Starch utilization lipoprotein
zobellia_372	0.06	0.05	TonB-dependent Receptor
zobellia_385	0.24	0.03	Short-chain dehydrogenase/reductase
zobellia_403	0.40	0.04	Conserved hypothetical protein
zobellia_467	0.27	0.01	TPR repeats protein
zobellia_468	0.20	0.01	DNA gyrase, subunit A
zobellia_472	0.14	0.02	Aspartate kinase I / Homoserine dehydrogenase I
zobellia_477	0.37	0.02	RmlD substrate binding domain family protein
zobellia_478	0.43	0.03	Pterin-4-alpha-carbinolamine dehydratase
zobellia_505	0.32	0.01	Bifunctional protein pyrR
zobellia_506	0.40	0.02	Ribosomal protein S1
zobellia_523	0.19	0.01	Dihydrolipoyl dehydrogenase, E3 component of branched-chain alpha-keto acid dehydrogenase complex
zobellia_530	0.20	0.01	Oligo-1,6-glucosidase, family GH13
zobellia_531	0.46	0.03	NADH dehydrogenase
zobellia_554	0.45	0.03	Inner membrane protein oxaA
zobellia_555	0.33	0.03	CTP synthase
zobellia_568	0.39	0.01	GMP synthase [glutamine-hydrolyzing]
zobellia_601	0.44	0.01	Recombination protein RecR
zobellia_607	0.19	0.02	Chaperone protein htpG
zobellia_615	0.03	0.02	Conserved hypothetical periplasmic protein
zobellia_628	0.13	0.03	Mannose-6-phosphate isomerase, type I
zobellia_638	0.20	0.05	OmpA-like outer membrane protein
zobellia_640	0.15	0.01	Hypothetical membrane protein
zobellia_641	0.42	0.01	peptide chain release factor RF-3
zobellia_646	0.43	0.04	Ribosomal protein L10
zobellia_667	0.41	0.01	Anthranilate synthase, component I
zobellia_668	0.42	0.03	Anthranilate synthase component II (glutamine amidotransferase)
zobellia_669	0.39	0.03	Anthranilate phosphoribosyltransferase
zobellia_670	0.38	0.03	Indole-3-glycerol phosphate synthase
zobellia_672	0.37	0.02	Tryptophan synthase beta subunit
zobellia_673	0.48	0.02	Tryptophan synthase, alpha subunit
zobellia_682	0.42	0.02	PASTA domain protein
zobellia_683	0.45	0.03	D-alanine--D-alanine ligase
zobellia_690	0.43	0.01	Conserved hypothetical protein
zobellia_716	0.31	0.01	Conserved hypothetical protein
zobellia_728	0.32	0.03	Conserved hypothetical periplasmic protein
zobellia_733	0.49	0.05	315 Conserved hypothetical periplasmic protein
zobellia_738	0.42	0.01	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase
zobellia_739	0.48	0.03	Methionyl aminopeptidase, family M24
zobellia_742	0.39	0.01	Conserved hypothetical protein
zobellia_743	0.22	0.01	Na(+)-translocating NADH-quinone reductase subunit F

Annexe 2 : Gènes sous-exprimés en présence d'alginate

zobellia_745	0.31	0.02	Na(+)-translocating NADH-quinone reductase subunit D
zobellia_746	0.23	0.04	Na(+)-translocating NADH-quinone reductase subunit C
zobellia_747	0.19	0.01	Na(+)-translocating NADH-quinone reductase subunit B
zobellia_748	0.28	0.01	Na(+)-translocating NADH-quinone reductase subunit A
zobellia_752	0.40	0.02	l-pyrroline-5-carboxylate dehydrogenase
zobellia_764	0.43	0.02	Fructose-bisphosphate aldolase, class II
zobellia_765	0.41	0.04	Acetyl-coenzyme A carboxylase, carboxyl transferase subunit beta
zobellia_768	0.28	0.02	Polyribonucleotide nucleotidyltransferase
zobellia_789	0.40	0.04	Conserved hypothetical protein
zobellia_809	0.22	0.01	Gliding motility membrane lipoprotein
zobellia_813	0.49	0.03	DNA primase
zobellia_818	0.25	0.04	S-adenosylmethionine:tRNA ribosyltransferase-isomerase
zobellia_819	0.15	0.02	3-Phosphoshikimate 1-carboxyvinyltransferase
zobellia_826	0.25	0.00	3-Deoxy-7-phosphoheptulonate synthase / Chorismate mutase
zobellia_827	0.27	0.02	Prephenate dehydrogenase
zobellia_828	0.25	0.01	Aminotransferase
zobellia_829	0.42	0.05	Prephenate dehydratase
zobellia_856	0.36	0.01	S-adenosylmethionine synthetase
zobellia_879	0.38	0.04	Calcium-dependent ATPase
zobellia_908	0.27	0.05	Butyryl-CoA dehydrogenase
zobellia_953	0.23	0.03	Lysyl-tRNA synthetase
zobellia_960	0.31	0.01	Conserved hypothetical periplasmic protein
zobellia_963	0.41	0.01	Glutamate-1-semialdehyde 2,1-aminomutase.
zobellia_1053	0.30	0.05	GntR-type transcriptional regulator
zobellia_1106	0.41	0.02	Exopolysaccharide production protein
zobellia_1108	0.14	0.05	Tyrosine-protein kinase
zobellia_1114	0.21	0.04	Glycosyltransferase, family GT4
zobellia_1156	0.41	0.02	MiaB-like tRNA modifying protein
zobellia_1157	0.48	0.05	Sigma-54-dependent transcriptional regulator
zobellia_1162	0.19	0.04	60 kDa chaperonin
zobellia_1171	0.20	0.00	Bifunctional purine biosynthesis protein PurH
zobellia_1184	0.50	0.03	Stationary-phase survival protein surE
zobellia_1193	0.29	0.02	Peptidyl-prolyl isomerase family protein
zobellia_1197	0.47	0.03	TPR repeats protein
zobellia_1199	0.48	0.03	Type III pantothenate kinase
zobellia_1204	0.42	0.03	Purine nucleoside phosphorylase
zobellia_1227	0.15	0.04	GCN5-related N-acetyltransferase
zobellia_1228	0.43	0.03	Argininosuccinate synthase
zobellia_1230	0.24	0.02	Pyrroline-5-carboxylate reductase
zobellia_1231	0.45	0.05	Acetylornithine aminotransferase
zobellia_1233	0.26	0.04	Glutamate 5-kinase
zobellia_1234	0.32	0.04	N-acetylornithine carbamoyltransferase
zobellia_1236	0.40	0.05	Acetylornithine deacetylase, family M20A
zobellia_1237	0.22	0.04	Argininosuccinate lyase
zobellia_1276	0.45	0.01	Crp/Fnr-type transcriptional regulator
zobellia_1301	0.28	0.05	Hypothetical protein
zobellia_1303	0.47	0.01	Tyrosyl-tRNA synthetase
zobellia_1325	0.39	0.02	Conserved hypothetical protein
zobellia_1336	0.36	0.05	Transcription elongation factor GreA
zobellia_1344	0.14	0.00	Conserved hypothetical periplasmic protein
zobellia_1345	0.07	0.02	Conserved hypothetical periplasmic protein
zobellia_1361	0.42	0.04	Metallopeptidase, family M16
zobellia_1377	0.44	0.04	Conserved hypothetical periplasmic protein
zobellia_1382	0.50	0.04	3-oxoacyl-[acyl-carrier-protein] synthase
zobellia_1405	0.16	0.03	ATP synthase, F1 sector beta subunit
zobellia_1437	0.26	0.01	TonB-dependent Receptor
zobellia_1438	0.33	0.01	Glucose-6-phosphate isomerase
zobellia_1449	0.19	0.01	Isoleucyl-tRNA synthetase
zobellia_1463	0.35	0.04	Cytochrome c
zobellia_1467	0.48	0.03	Glycerol-3-phosphate dehydrogenase [NAD(P)+]
zobellia_1532	0.19	0.02	Conserved hypothetical protein
zobellia_1533	0.27	0.04	Conserved hypothetical protein
zobellia_1585	0.26	0.02	Amidophosphoribosyltransferase
zobellia_1586	0.13	0.01	Carbohydrate kinase
zobellia_1587	0.47	0.04	Putative protein
zobellia_1593	0.29	0.01	Conserved hypothetical protein
zobellia_1613	0.31	0.02	Riboflavin synthase alpha chain
zobellia_1677	0.46	0.02	Metal-dependent hydrolase
zobellia_1678	0.25	0.04	Serine endopeptidase, family S8
zobellia_1679	0.33	0.05	Aminoamidase, family M1
zobellia_1698	0.28	0.05	Ribose-5-phosphate isomerase B
zobellia_1724	0.17	0.03	Phosphoglycerate kinase
zobellia_1739	0.38	0.04	dTDP-6-deoxy-D-xylo-hex-3-ulose aminase
zobellia_1759	0.31	0.01	TonB-dependent Receptor

Annexe 2 : Gènes sous-exprimés en présence d'alginate

zobellia_1761	0.39	0.02	Spermidine/putrescine-specific ABC importer, periplasmic component
zobellia_1762	0.32	0.04	Spermidine/putrescine-specific ABC importer, permease component
zobellia_1767	0.30	0.04	Phenylacetic acid degradation NADH oxidoreductase
zobellia_1778	0.05	0.02	ATP synthase, F0 sector C subunit
zobellia_1779	0.17	0.01	ATP synthase, F0 sector B subunit
zobellia_1780	0.23	0.03	ATP synthase, F1 sector delta subunit
zobellia_1781	0.38	0.02	ATP synthase, F1 sector alpha subunit
zobellia_1782	0.25	0.01	ATP synthase, F1 sector gamma subunit
zobellia_1793	0.48	0.04	Tryptophanyl-tRNA synthetase
zobellia_1823	0.49	0.01	Succinate dehydrogenase iron-sulfur subunit
zobellia_1836	0.39	0.01	Thiamine-monophosphate kinase
zobellia_1846	0.41	0.02	Conserved hypothetical periplasmic protein
zobellia_1855	0.37	0.01	Delta-aminolevulinic acid dehydratase
zobellia_1859	0.40	0.01	Coproporphyrinogen-III oxidase
zobellia_1862	0.29	0.04	Porphobilinogen deaminase / Uroporphyrinogen-III synthase
zobellia_1863	0.31	0.01	Glutamyl-tRNA-reductase
zobellia_1877	0.33	0.02	3-hydroxybutyryl-CoA dehydratase
zobellia_1892	0.36	0.01	Conserved hypothetical protein
zobellia_1935	0.36	0.01	Conserved hypothetical periplasmic protein
zobellia_1941	0.36	0.05	Aspartyl-tRNA synthetase
zobellia_2052	0.49	0.05	Conserved hypothetical protein
zobellia_2066	0.39	0.03	3-oxoacyl-[acyl-carrier-protein] synthase
zobellia_2071	0.47	0.03	Conserved hypothetical protein
zobellia_2073	0.27	0.04	(3R)-Hydroxymyristoyl-[acyl-carrier-protein] dehydratase
zobellia_2074	0.37	0.02	3-oxoacyl-[acyl-carrier-protein] synthase III
zobellia_2080	0.32	0.02	KGW repeats protein
zobellia_2097	0.20	0.02	Methionyl-tRNA synthetase, class Ia
zobellia_2103	0.31	0.03	Thiol-disulfide oxidoreductase
zobellia_2121	0.36	0.01	GCN5-related N-acetyltransferase
zobellia_2122	0.49	0.04	Conserved hypothetical protein
zobellia_2123	0.39	0.01	Conserved hypothetical protein
zobellia_2136	0.36	0.01	Riboflavin biosynthesis protein RibF
zobellia_2222	0.17	0.04	Nucleotidyl transferase
zobellia_2239	0.29	0.02	Flavin mononucleotide binding protein
zobellia_2253	0.35	0.01	FixH-like protein
zobellia_2312	0.10	0.04	TonB-dependent Receptor
zobellia_2313	0.08	0.02	Conserved hypothetical lipoprotein
zobellia_2315	0.20	0.05	TonB-dependent Transducer
zobellia_2322	0.49	0.05	PorR protein, PLP-dependent aminotransferase family
zobellia_2327	0.44	0.03	Dihydrolipoyllysine-residue succinyltransferase E2 component of 2-oxoglutarate dehydrogenase complex
zobellia_2339	0.34	0.03	Asparaginyl-tRNA synthetase
zobellia_2342	0.49	0.02	Uridylate kinase
zobellia_2343	0.24	0.05	Glycyl aminopeptidase, family M61
zobellia_2344	0.29	0.02	Elongation factor Ts (EF-Ts)
zobellia_2345	0.16	0.05	Ribosomal protein S2
zobellia_2346	0.19	0.01	Ribosomal protein S9
zobellia_2364	0.36	0.03	Conserved hypothetical protein,
zobellia_2365	0.17	0.01	Cytochrome C monohaem
zobellia_2366	0.36	0.03	Conserved membrane protein,
zobellia_2368	0.36	0.02	Cytochrome c oxidase, subunit I
zobellia_2384	0.15	0.04	Conserved hypothetical membrane protein
zobellia_2407	0.30	0.04	Conserved protein containing calx-beta domain
zobellia_2419	0.38	0.02	Glutamate 5-kinase
zobellia_2420	0.40	0.02	Pyrroline-5-carboxylate reductase
zobellia_2466	0.25	0.02	3-oxoacyl-[acyl-carrier-protein] synthase III
zobellia_2467	0.20	0.02	Glycine cleavage system, protein P component
zobellia_2482	0.50	0.03	Short-chain dehydrogenase/reductase
zobellia_2487	0.48	0.01	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase
zobellia_2497	0.42	0.02	Hypothetical lipoprotein
zobellia_2501	0.21	0.03	Hypothetical membrane protein
zobellia_2520	0.31	0.03	Ribosomal protein L14
zobellia_2521	0.40	0.04	Ribosomal protein L24
zobellia_2523	0.43	0.04	Ribosomal protein S14
zobellia_2525	0.30	0.01	Ribosomal protein L6
zobellia_2526	0.24	0.03	Ribosomal protein L18
zobellia_2531	0.31	0.02	Translation initiation factor IF-1
zobellia_2532	0.38	0.03	Ribosomal protein S13
zobellia_2533	0.26	0.04	Ribosomal protein S11
zobellia_2534	0.11	0.03	Ribosomal protein S4
zobellia_2535	0.15	0.01	DNA-dependent RNA polymerase, alpha subunit
zobellia_2536	0.11	0.02	Ribosomal protein L17
zobellia_2537	0.31	0.01	Carbamoyl-phosphate synthase small chain
zobellia_2538	0.44	0.02	Enolase
zobellia_2542	0.29	0.02	Conserved hypothetical protein

Annexe 2 : Gènes sous-exprimés en présence d'alginate

zobellia_2560	0.30	0.04	Nucleotide-sugar epimerases
zobellia_2561	0.43	0.04	Conserved hypothetical protein
zobellia_2562	0.13	0.02	Arginyl-tRNA synthetase
zobellia_2570	0.39	0.03	Bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase
zobellia_2577	0.32	0.04	Glyoxalase/extradiol ring-cleavage dioxygenase family
zobellia_2585	0.41	0.02	Conserved hypothetical protein
zobellia_2631	0.47	0.02	Two-component system - Response regulator
zobellia_2646	0.37	0.04	Conserved hypothetical membrane protein
zobellia_2681	0.04	0.01	TonB-dependent Receptor
zobellia_2682	0.15	0.03	Conserved hypothetical membrane protein
zobellia_2690	0.14	0.02	Conserved hypothetical periplasmic protein
zobellia_2727	0.22	0.02	GTP-binding protein TypA
zobellia_2733	0.42	0.02	Cell division protein, ATPase component
zobellia_2734	0.48	0.04	TPR repeat-containing protein
zobellia_2743	0.22	0.01	Asparagine synthase B
zobellia_2749	0.48	0.02	Protein-export membrane protein secD
zobellia_2759	0.36	0.04	Cysteinyl-tRNA synthetase
zobellia_2760	0.42	0.01	GTP cyclohydrolase I
zobellia_2779	0.31	0.02	Conserved hypothetical membrane protein
zobellia_2780	0.18	0.04	Triosephosphate isomerase
zobellia_2791	0.26	0.02	Glutamine synthetase II
zobellia_2819	0.30	0.02	Branched-chain-amino-acid aminotransferase
zobellia_2824	0.45	0.03	Co-chaperone protein dnaJ
zobellia_2844	0.26	0.01	Conserved hypothetical membrane protein
zobellia_2845	0.22	0.04	Preprotein translocase subunit
zobellia_2864	0.19	0.02	Diaminopimelate decarboxylase
zobellia_2866	0.29	0.01	Succinyl-CoA ligase [ADP-forming] subunit beta
zobellia_2878	0.49	0.03	Excinuclease ABC, subunit B
zobellia_2901	0.43	0.04	Conserved hypothetical protein
zobellia_2904	0.37	0.03	2,3,4,5-Tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
zobellia_2928	0.45	0.02	LAO/AO transport system kinase ArgK
zobellia_2934	0.24	0.01	Conserved hypothetical lipoprotein
zobellia_3019	0.26	0.01	Transketolase subunit A
zobellia_3039	0.30	0.02	Threonyl-tRNA synthetase
zobellia_3128	0.43	0.03	Possible RNA methylase
zobellia_3129	0.44	0.04	Conserved hypothetical periplasmic protein
zobellia_3130	0.42	0.01	Conserved hypothetical lipoprotein
zobellia_3173	0.47	0.04	GDP-mannose 4,6-dehydratase
zobellia_3216	0.21	0.05	Molybdopterin synthase, large subunit
zobellia_3217	0.16	0.01	Conserved hypothetical periplasmic protein
zobellia_3325	0.38	0.01	Radical SAM superfamily protein
zobellia_3459	0.30	0.03	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase
zobellia_3479	0.48	0.01	Metallo-beta-lactamase superfamily protein
zobellia_3491	0.45	0.01	Lipoprotein-specific ABC transporter, ATPase component
zobellia_3492	0.32	0.02	Conserved hypothetical protein
zobellia_3493	0.40	0.01	Conserved hypothetical protein
zobellia_3505	0.47	0.04	Mucin-desulfating sulfatase, family S1-11
zobellia_3575	0.20	0.03	Conserved hypothetical protein
zobellia_3655	0.28	0.03	Thiamine-phosphate synthase-like protein
zobellia_3656	0.26	0.02	Thiamine-phosphate pyrophosphorylase
zobellia_3674	0.28	0.05	Adenylosuccinate synthase
zobellia_3675	0.39	0.02	Ferric uptake regulation protein
zobellia_3705	0.37	0.02	Xaa-His dipeptidase, family M20
zobellia_3735	0.41	0.04	6-Phosphofructokinase, isoenzyme 1
zobellia_3736	0.18	0.01	Glyceraldehyde 3-phosphate dehydrogenase
zobellia_3737	0.31	0.03	Conserved hypothetical protein
zobellia_3754	0.45	0.01	Conserved hypothetical protein
zobellia_3771	0.22	0.02	Vi polysaccharide biosynthesis protein vipA
zobellia_3779	0.30	0.02	UDP-glucose 6-dehydrogenase
zobellia_3780	0.40	0.04	UDP-glucuronic acid decarboxylase
zobellia_3786	0.48	0.05	Hypothetical protein containing CBM6
zobellia_3787	0.50	0.03	Tyrosine-protein phosphatase
zobellia_3829	0.33	0.01	DNA polymerase III, beta chain
zobellia_3832	0.13	0.04	Conserved hypothetical protein
zobellia_3834	0.50	0.03	Phosphoribosylaminoimidazole-succinocarboxamide synthase
zobellia_3836	0.32	0.01	Cyclic nucleotide-binding protein
zobellia_3853	0.47	0.02	ABC-2 type exporter, permease component
zobellia_3854	0.30	0.00	Membrane fusion protein
zobellia_3855	0.32	0.01	Outer membrane efflux protein
zobellia_3856	0.46	0.01	Multi-pass membrane transporter
zobellia_3868	0.38	0.01	2-hydroxyacid dehydrogenase homolog
zobellia_3875	0.38	0.03	ATP synthase, F0 sector A subunit
zobellia_3876	0.50	0.04	ATP synthase, F0 sector C subunit
zobellia_3877	0.18	0.01	ATP synthase, F0 sector B subunit

Annexe 2 : Gènes sous-exprimés en présence d'alginate

zobellia_3879	0.38	0.01	ATP synthase, F1 sector gamma subunit
zobellia_3941	0.19	0.02	3-Hexulose-6-phosphate isomerase
zobellia_3942	0.07	0.02	3-Hexulose-6-phosphate synthase
zobellia_3975	0.40	0.01	Conserved hypothetical periplasmic protein
zobellia_3982	0.35	0.02	3-Deoxy-D-manno-octulosonate 8-phosphate phosphatase
zobellia_3983	0.34	0.04	Conserved hypothetical protein
zobellia_3984	0.39	0.02	Protein NrfI
zobellia_4001	0.26	0.01	Cystathionine beta-lyase
zobellia_4007	0.23	0.01	Metallopeptidase, family M28
zobellia_4010	0.29	0.03	Dihydrodipicolinate synthetase
zobellia_4011	0.42	0.02	NADP-dependent fatty aldehyde dehydrogenase
zobellia_4012	0.46	0.02	Proline racemase
zobellia_4019	0.44	0.03	Esterase
zobellia_4021	0.45	0.04	2-oxoglutarate decarboxylase/SHCHC synthase
zobellia_4034	0.42	0.05	Conserved hypothetical membrane protein
zobellia_4039	0.36	0.03	Fumarate hydratase, class II
zobellia_4043	0.33	0.01	Arsenate reductase
zobellia_4044	0.30	0.01	Peptide chain release factor 2
zobellia_4052	0.09	0.03	Conserved hypothetical protein
zobellia_4055	0.24	0.02	Glycine C-acetyltransferase
zobellia_4074	0.30	0.01	Conserved hypothetical protein
zobellia_4080	0.14	0.01	Dihydroorotase
zobellia_4091	0.15	0.03	Cation efflux transporter
zobellia_4092	0.37	0.02	cation efflux system protein
zobellia_4110	0.10	0.04	TonB-dependent Receptor
zobellia_4116	0.13	0.05	Phosphoserine aminotransferase
zobellia_4118	0.46	0.01	Ferredoxin
zobellia_4149	0.47	0.04	Acyl-CoA thioester hydrolase yciA
zobellia_4160	0.50	0.01	Glyoxylate/hydroxypyruvate reductase A
zobellia_4161	0.25	0.00	ATP-dependent helicase
zobellia_4180	0.50	0.03	Amidohydrolase
zobellia_4245	0.36	0.02	DNA protection during starvation protein
zobellia_4246	0.36	0.01	Aldo/keto reductase related to aryl-alcohol dehydrogenases
zobellia_4247	0.46	0.02	Short-chain dehydrogenase/reductase
zobellia_4291	0.35	0.02	DEAD/DEAH RNA helicase
zobellia_4318	0.16	0.04	Conserved hypothetical protein
zobellia_4319	0.15	0.04	Conserved hypothetical periplasmic protein
zobellia_4327	0.38	0.04	Alginate lyase, family PL14
zobellia_4330	0.22	0.01	TonB-dependent Receptor
zobellia_4390	0.40	0.02	OmpA-like membrane protein
zobellia_4406	0.31	0.03	FAD-dependent oxidoreductase
zobellia_4412	0.34	0.02	Phosphomannomutase/phosphoglucomutase
zobellia_4435	0.43	0.01	Histidine decarboxylase
zobellia_4454	0.49	0.01	Conserved hypothetical periplasmic protein
zobellia_4462	0.49	0.04	Conserved hypothetical protein
zobellia_4510	0.50	0.02	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase
zobellia_4524	0.17	0.02	Elongation factor P
zobellia_4525	0.34	0.02	Acyl-[acyl-carrier-protein]-:UDP-N-acetylglucosamine O-acyltransferase
zobellia_4526	0.40	0.01	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase / (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
zobellia_4527	0.40	0.02	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase
zobellia_4528	0.41	0.01	Phosphohydrolase family protein
zobellia_4529	0.29	0.04	Two-component system - Response regulator
zobellia_4540	0.49	0.03	Metallopeptidase, family M14
zobellia_4557	0.19	0.02	3-Isopropylmalate dehydratase, small subunit
zobellia_4558	0.38	0.02	2-Isopropylmalate synthase
zobellia_4559	0.10	0.03	3-Isopropylmalate dehydrogenase
zobellia_4567	0.28	0.01	Ketol-acid reductoisomerase
zobellia_4569	0.36	0.01	S-adenosylmethionine-dependent O-methyltransferase
zobellia_4571	0.22	0.01	Dihydroxy-acid dehydratase
zobellia_4603	0.43	0.04	Anthranilate synthase component I-like protein
zobellia_4625	0.32	0.03	PLP-dependent aminotransferase
zobellia_4626	0.31	0.05	6,7-dihydropteridine reductase
zobellia_4653	0.35	0.03	Aldonic acid dehydratase
zobellia_4680	0.29	0.02	Long chain dehydrogenase/reductase
zobellia_4685	0.32	0.03	L-fucose-proton symporter
zobellia_4701	0.40	0.05	Mannonate dehydratase
zobellia_4714	0.46	0.04	Conserved hypothetical periplasmic protein
zobellia_4726	0.50	0.04	Phenylalanyl-tRNA synthetase beta chain
zobellia_4730	0.20	0.03	ABC transporter, ATPase component
zobellia_4732	0.38	0.02	Transaldolase
zobellia_4742	0.38	0.01	Serine hydroxymethyltransferase
zobellia_4752	0.24	0.05	3-Hydroxybutyryl-CoA dehydrogenase
zobellia_4754	0.45	0.05	Nuclease

---

Annexe 2 : Gènes sous-exprimés en présence d'alginate

zobellia_4766	0.43	0.03	Coenzyme A biosynthesis bifunctional protein coaBC
---------------	------	------	--





# Bibliographie

---

- Abbott, D. W., J. M. Eirin-Lopez and A. B. Boraston (2008). "Insight into ligand diversity and novel biological roles for family 32 carbohydrate-binding modules." *Molecular Biology and Evolution* **25**(1): 155-67.
- Abdek-Akher, M., J. K. Hamilton and F. Smith (1951). "The Reduction of Sugars with Sodium Borohydride." *Journal of the American Chemical Society* **73**(10): 4691-92.
- Akimoto, C., H. Aoyagi and H. Tanaka (1999). "Endogenous elicitor-like effects of alginate on physiological activities of plant cells." *Applied Microbiology and Biotechnology* **52**(3): 429-36.
- Akiyama, H., T. Endo, R. Nakakita, K. Murata, Y. Yonemoto and K. Okayama (1992). "Effect of depolymerized alginates on the growth of bifidobacteria." *Bioscience Biotechnology and Biochemistry* **56**(2): 355-6.
- Alkawash, M. A., J. S. Sothill and N. L. Schiller (2006). "Alginate lyase enhances antibiotic killing of mucoid *Pseudomonas aeruginosa* in biofilms." *Apmis* **114**(2): 131-38.
- Allredge, A. L. and M. W. Silver (1988). "Characteristics, dynamics and significance of marine snow." *Progress in Oceanography* **20**(1): 41-82.
- Alonso, C., F. Warnecke, R. Amann and J. Pernthaler (2007). "High local and global diversity of *Flavobacteria* in marine plankton." *Environmental Microbiology* **9**(5): 1253-66.
- Amsler, C. D. and V. A. Fairhead (2006). "Defensive and sensory chemical ecology of brown algae." *Advances in Botanical Research, Vol 43* **43**: 1-91.
- An, Q. D., G. L. Zhang, H. T. Wu, Z. C. Zhang, G. S. Zheng, L. Luan, Y. Murata and X. Li (2009). "Alginate-deriving oligosaccharide production by alginate lyase from newly isolated *Flavobacterium* sp. LXA and its potential application in protection against pathogens." *Journal of Applied Microbiology* **106**(1): 161-70.
- Anderson, K. L. and A. A. Salyers (1989). "Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*." *Journal of Bacteriology* **171**(6): 3199-204.
- Andrade, L. R., R. N. Leal, M. Nosedá, M. E. Duarte, M. S. Pereira, P. A. Mourao, M. Farina and G. M. Amado Filho (2010). "Brown algae overproduce cell wall polysaccharides as a protection mechanism against the heavy metal toxicity." *Marine Pollution Bulletin* **60**(9): 1482-8.
- Aquino, R. S., A. M. Landeira-Fernandez, A. P. Valente, L. R. Andrade and P. A. Mourao (2005). "Occurrence of sulfated galactans in marine angiosperms: evolutionary implications." *Glycobiology* **15**(1): 11-20.
- Archer, K. J. and S. E. Reese (2010). "Detection call algorithms for high-throughput gene expression microarray data." *Briefings in Bioinformatics* **11**(2): 244-52.
- Azam, F. (1998). "Microbial control of oceanic carbon flux: The plot thickens." *Science* **280**(5364): 694-96.
- Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyerreil and F. Thingstad (1983). "The Ecological Role of Water-Column Microbes in the Sea." *Marine Ecology-Progress series* **10**(3): 257-63.
- Azam, F. and R. A. Long (2001). "Sea snow microcosms." *Nature* **414**(6863): 495, 97-8.
- Azam, F. and F. Malfatti (2007). "Microbial structuring of marine ecosystems." *Nature Reviews Microbiology* **5**(10): 782-91.
- Bai, J., J. Wang, F. Xue, J. Li, L. Bu, J. Hu, G. Xu, Q. Bao, G. Zhao, X. Ding, J. Yan and J. Wu (2010). "proTF: a comprehensive data and phylogenomics resource for prokaryotic transcription factors." *Bioinformatics* **26**(19): 2493-5.
- Baldauf, S. L. (2008). "An overview of the phylogeny and diversity of eukaryotes." *Journal of Systematics and Evolution* **46**(3): 263-73.
- Barbeyron, T., A. Gerard, P. Potin, B. Henrissat and B. Kloareg (1998). "The kappa-carrageenase of the marine bacterium *Cytophaga drobachiensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases." *Molecular Biology and Evolution* **15**(5): 528-37.

- Barbeyron, T., S. L'Haridon, E. Corre, B. Kloareg and P. Potin (2001). "*Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of [*Cytophaga*] *uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov." *International Journal of Systematic and Evolutionary Microbiology* **51**(3): 985-97.
- Barbeyron, T., G. Michel, P. Potin, B. Henrissat and B. Kloareg (2000). "iota-Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of kappa-carrageenases." *Journal of Biological Chemistry* **275**(45): 35499-505.
- Bartell, P. F., T. E. Orr and G. K. Lam (1966). "Polysaccharide depolymerase associated with bacteriophage infection." *Journal of Bacteriology* **92**(1): 56-62.
- Barthelmebs, L., B. Lecomte, C. Divies and J. F. Cavin (2000). "Inducible metabolism of phenolic acids in *Pediococcus pentosaceus* is encoded by an autoregulated operon which involves a new class of negative transcriptional regulator." *Journal of Bacteriology* **182**(23): 6724-31.
- Basta, G., L. Osticioli, M. E. Rossodivita, P. Sarchielli, C. Tortoioli, P. Brunetti and R. Calafiore (1995). "Method for Fabrication of Coherent Microcapsules - a New, Potential Immunoisulatory Barrier for Pancreatic-Islet Transplantation." *Diabetes Nutrition & Metabolism* **8**(2): 105-12.
- Behr, M. A., M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolnik, S. Rane and P. M. Small (1999). "Comparative genomics of BCG vaccines by whole-genome DNA microarray." *Science* **284**(5419): 1520-3.
- Bengtsson, M. M. and L. Ovreas (2010). "*Planctomycetes* dominate biofilms on surfaces of the kelp *Laminaria hyperborea*." *BMC Microbiology* **10**: 261.
- Benjamini, Y. and Y. Hochberg (1995). "Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B-Methodological* **57**(1): 289-300.
- Benner, R., J. D. Pakulski, M. Mccarthy, J. I. Hedges and P. G. Hatcher (1992). "Bulk Chemical Characteristics of Dissolved Organic-Matter in the Ocean." *Science* **255**(5051): 1561-64.
- Berteau, O., I. McCort, N. Goasdoue, B. Tissot and R. Daniel (2002). "Characterization of a new alpha-L-fucosidase isolated from the marine mollusk *Pecten maximus* that catalyzes the hydrolysis of alpha-L-fucose from algal fucoidan (*Ascophyllum nodosum*)." *Glycobiology* **12**(4): 273-82.
- Bitton, R., M. Ben-Yehuda, M. Davidovich, Y. Balazs, P. Potin, L. Delage, C. Colin and H. Bianco-Peled (2006). "Structure of algal-born phenolic polymeric adhesives." *Macromolecular Bioscience* **6**(9): 737-46.
- Bolstad, B. M., R. A. Irizarry, M. Astrand and T. P. Speed (2003). "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias." *Bioinformatics* **19**(2): 185-93.
- Boraston, A. B., D. N. Bolam, H. J. Gilbert and G. J. Davies (2004). "Carbohydrate-binding modules: fine-tuning polysaccharide recognition." *Biochemical Journal* **382**(Pt 3): 769-81.
- Boyd, A. and A. M. Chakrabarty (1994). "Role of alginate lyase in cell detachment of *Pseudomonas aeruginosa*." *Applied and Environmental Microbiology* **60**(7): 2355-9.
- Boyd, J. and J. R. Turvey (1978). "Structural Studies of Alginic Acid, Using a Bacterial Poly-Alpha-L-Guluronate Lyase." *Carbohydrate Research* **66**(Oct): 187-94.
- Boyen, C., Y. Bertheau, T. Barbeyron and B. Kloareg (1990a). "Preparation of guluronate lyase from *Pseudomonas alginovora* for protoplast isolation in *Laminaria*." *Enzyme and Microbial Technology* **12**(11): 885-90.
- Boyen, C., B. Kloareg, M. Polnefuller and A. Gibor (1990b). "Preparation of Alginate Lyases from Marine Mollusks for Protoplast Isolation in Brown Algae." *Phycologia* **29**(2): 173-81.
- Braun, P. and J. LaBaer (2003). "High throughput protein production for functional proteomics." *Trends in Biotechnology* **21**(9): 383-8.
- Braun, T. F. and M. J. McBride (2005). "*Flavobacterium johnsoniae* GldJ is a lipoprotein that is required for gliding motility." *Journal of Bacteriology* **187**(8): 2628-37.
- Brunet, J. F., F. Denizot and P. Golstein (1988). "A Differential Molecular-Biology Search for Genes Preferentially Expressed in Functional Lymphocytes-T - the CtlA Genes." *Immunological Reviews* **103**: 21-36.

- Burke, L., Y. Kura, K. Kassem, C. Revenga, M. Spalding and D. McAllister (2001). Pilot Analysis of Global Ecosystems. Coastal Ecosystems. Washington.
- Callebaut, I., G. Labesse, P. Durand, A. Poupon, L. Canard, J. Chomilier, B. Henrissat and J. P. Mornon (1997). "Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives." *Cellular and Molecular Life Sciences* **53**(8): 621-45.
- Cantarel, B. L., P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard and B. Henrissat (2009). "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics." *Nucleic Acids Research* **37**: D233-8.
- Cashin, P., L. Goldsack, D. Hall and R. O'Toole (2006). "Contrasting signal transduction mechanisms in bacterial and eukaryotic gene transcription." *FEMS Microbiology Letters* **261**(2): 155-64.
- Caswell, R., J. Williams, A. Lyddiatt and S. Busby (1992). "Overexpression, purification and characterization of the *Escherichia coli* MelR transcription activator protein." *Biochemical Journal* **287**(2): 493-9.
- Chen, W., K. Honma, A. Sharma and H. K. Kuramitsu (2006). "A universal stress protein of *Porphyromonas gingivalis* is involved in stress responses and biofilm formation." *FEMS Microbiology Letters* **264**(1): 15-21.
- Chhabra, S. R., K. R. Shockley, S. B. Connors, K. L. Scott, R. D. Wolfinger and R. M. Kelly (2003). "Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*." *Journal of Biological Chemistry* **278**(9): 7540-52.
- Chin, W. C., M. V. Orellana and P. Verdugo (1998). "Spontaneous assembly of marine dissolved organic matter into polymer gels." *Nature* **391**(6667): 568-72.
- Christensen, B. E., M. Indergaard and O. Smidsrod (1990). "Polysaccharide Research in Trondheim." *Carbohydrate Polymers* **13**(3): 239-55.
- Chruszcz, M., M. Domagalski, T. Osinski, A. Wlodawer and W. Minor (2010). "Unmet challenges of structural genomics." *Current Opinion in Structural Biology* **20**(5): 587-97.
- Cock, J. M., L. Sterck, P. Rouze, D. Scornet, A. E. Allen, G. Amoutzias, V. Anthouard, F. Artiguenave, J. M. Aury, J. H. Badger, B. Beszteri, K. Billiau, E. Bonnet, J. H. Bothwell, C. Bowler, C. Boyen, C. Brownlee, C. J. Carrano, B. Charrier, G. Y. Cho, S. M. Coelho, J. Collen, E. Corre, C. Da Silva, L. Delage, N. Delaroque, S. M. Dittami, S. Doulebeau, M. Elias, G. Farnham, C. M. M. Gachon, B. Gschloessl, S. Heesch, K. Jabbari, C. Jubin, H. Kawai, K. Kimura, B. Kloareg, F. C. Kupper, D. Lang, A. Le Bail, C. Leblanc, P. Lerouge, M. Lohr, P. J. Lopez, C. Martens, F. Maumus, G. Michel, D. Miranda-Saavedra, J. Morales, H. Moreau, T. Motomura, C. Nagasato, C. A. Napoli, D. R. Nelson, P. Nyvall-Collen, A. F. Peters, C. Pommier, P. Potin, J. Poulain, H. Quesneville, B. Read, S. A. Rensing, A. Ritter, S. Rousvoal, M. Samanta, G. Samson, D. C. Schroeder, B. Segurens, M. Strittmatter, T. Tonon, J. W. Tregear, K. Valentin, P. von Dassow, T. Yamagishi, Y. Van de Peer and P. Wincker (2010). "The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae." *Nature* **465**(7298): 617-21.
- Condemine, G. and J. Robertbaudouy (1991). "Analysis of an *Erwinia chrysantemi* gene cluster involved in pectin degradation." *Molecular Microbiology* **5**(9): 2191-202.
- Corre, S. and D. Prieur (1990). "Density and Morphology of Epiphytic Bacteria on the Kelp *Laminaria digitata*." *Botanica Marina* **33**(6): 515-23.
- Cosse, A. (2007). Analyse du transcriptome des réponses de défense de l'algue brune *Laminaria digitata*, Institut des Sciences et Industries du Vivant et de l'Environnement. **Thèse de doctorat**.
- Cosse, A., C. Leblanc and P. Potin (2007). "Dynamic defense of marine macroalgae against pathogens: from early activated to gene-regulated responses." *Advances in Botanical Research* **46**: 221-66.
- Cowman, M. K., M. F. Slahetka, D. M. Hittner, J. Kim, M. Forino and G. Gadelrab (1984). "Polyacrylamide-gel electrophoresis and Alcian Blue staining of sulphated glycosaminoglycan oligosaccharides." *Biochemical Journal* **221**(3): 707-16.
- Craigie, J. S., E. R. Morris, D. A. Rees and D. Thom (1984). "Alginate Block Structure in *Phaeophyceae* from Nova-Scotia - Variation with Species, Environment and Tissue-Type." *Carbohydrate Polymers* **4**(4): 237-52.

- Cronshaw, J., A. Myers and R. D. Preston (1958). "A Chemical and Physical Investigation of the Cell Walls of Some Marine Algae." *Biochimica et Biophysica Acta* **27**(1): 89-103.
- Dabin, J. (2008). Etude structurale et fonctionnelle des polysaccharidases de *Rhodospirellula baltica*. Paris, Université Pierre et Marie Curie. **Thèse de doctorat**.
- Davies, G. and B. Henrissat (1995). "Structures and mechanisms of glycosyl hydrolases." *Structure* **3**(9): 853-9.
- Davies, G. J., K. S. Wilson and B. Henrissat (1997). "Nomenclature for sugar-binding subsites in glycosyl hydrolases." *Biochemical Journal* **321**(2): 557-9.
- Davis, T. A., B. Volesky and A. Mucci (2003). "A review of the biochemistry of heavy metal biosorption by brown algae." *Water Research* **37**(18): 4311-30.
- Deboer, H. A., L. J. Comstock and M. Vasser (1983). "The Tac Promoter - a Functional Hybrid Derived from the Trp and Lac Promoters." *PNAS* **80**(1): 21-25.
- Dereeper, A., V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J. F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J. M. Claverie and O. Gascuel (2008). "Phylogeny.fr: robust phylogenetic analysis for the non-specialist." *Nucleic Acids Research* **36**: W465-9.
- DeRisi, J., L. Penland, P. O. Brown, M. L. Bittner, P. S. Meltzer, M. Ray, Y. Chen, Y. A. Su and J. M. Trent (1996). "Use of a cDNA microarray to analyse gene expression patterns in human cancer." *Nature Genetics* **14**(4): 457-60.
- Dodd, D., Y. H. Moon, K. Swaminathan, R. I. Mackie and I. K. Cann (2010). "Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic *Bacteroidetes*." *Journal of Biological Chemistry* **285**(39): 30261-73.
- Donati, I. and S. Paoletti (2009). Material properties of alginate. *Alginate: biology and application*. B. H. A. Rehm.
- Doubet, R. S. and R. S. Quatrano (1984). "Properties of Alginate Lyases from Marine Bacteria." *Applied and Environmental Microbiology* **47**(4): 699-703.
- Draget, K. I., O. Smidsrod and G. Skjak-Braek (2005). Alginates from algae. *Polysaccharides and polyamides in the food industry. Properties, production and patents*. A. Steinbüchel and S. K. Rhee.
- Duarte, C. C., J. J. Middelburg and N. Caraco (2005). "Major role of marine vegetation on the oceanic carbon cycle." *Biogeosciences* **2**: 1-8.
- Dumon-Seignovert, L., G. Cariot and L. Vuillard (2004). "The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3)." *Protein Expression and Purification* **37**(1): 203-6.
- Eftekhari, F. and D. P. Speert (1988). "Alginase treatment of mucoid *Pseudomonas aeruginosa* enhances phagocytosis by human monocyte-derived macrophages." *Infection and Immunity* **56**(11): 2788-93.
- Eisenthal, R., M. J. Danson and D. W. Hough (2007). "Catalytic efficiency and kcat/KM: a useful comparator?" *Trends in Biotechnology* **25**(6): 247-9.
- Eitinger, T., D. A. Rodionov, M. Grote and E. Schneider (2011). "Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions." *FEMS Microbiology Reviews* **35**(1): 3-67.
- Elyakova, L. A. and V. V. Favorov (1974). "Isolation and Certain Properties of Alginate Lyase Vi from Mollusk *Littorina* Sp." *Biochimica et Biophysica Acta* **358**(2): 341-54.
- Emanuelsson, O., S. Brunak, G. von Heijne and H. Nielsen (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." *Nature Protocols* **2**(4): 953-71.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." *Acta Crystallographica Section D Biological Crystallography* **60**(Pt 12 Pt 1): 2126-32.
- Entner, N. and M. Doudoroff (1952). "Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*." *Journal of Biological Chemistry* **196**(2): 853-62.
- Field, C. B., M. J. Behrenfeld, J. T. Randerson and P. Falkowski (1998). "Primary production of the biosphere: integrating terrestrial and oceanic components." *Science* **281**(5374): 237-40.
- Fox, B. G., C. Goulding, M. G. Malkowski, L. Stewart and A. Deacon (2008). "Structural genomics: from genes to structures with valuable materials and many questions in between." *Nature Methods* **5**(2): 129-32.

- Friedberg, D., M. Midkiff and J. M. Calvo (2001). "Global versus local regulatory roles for Lrp-related proteins: *Haemophilus influenzae* as a case study." *Journal of Bacteriology* **183**(13): 4004-11.
- Fuerst, J. A. (1995). "The *Planctomycetes*: emerging models for microbial ecology, evolution and cell biology." *Microbiology* **141** ( Pt 7): 1493-506.
- Fujibaya, S., H. Habe and K. Nisizawa (1970). "Heterogeneity of Alginate in Special Reference to Enzymatic Degradation." *Journal of Biochemistry* **67**(1): 37-&.
- Gaboriaud, C., V. Bissery, T. Benchetrit and J. P. Mornon (1987). "Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences." *FEBS Letters* **224**(1): 149-55.
- Gacesa, P. (1987). "Alginate-modifying enzymes: a proposed unified mechanism of action for the lyases and epimerases." *FEBS Letters* **212**(2): 199-202.
- Gallegos, M. T., R. Schleif, A. Bairoch, K. Hofmann and J. L. Ramos (1997). "Arac/XylS family of transcriptional regulators." *Microbiology and Molecular Biology Reviews* **61**(4): 393-410.
- Garron, M. L. and M. Cygler (2010). "Structural and mechanistic classification of uronic acid-containing polysaccharide lyases." *Glycobiology* **20**(12): 1547-73.
- Gattuso, J.-P., M. Frankignoulle and R. Wollast (1998). "Carbon and carbonate metabolism in coastal aquatic ecosystems." *Annual Review of Ecology and Systematics* **29**(1): 405-34.
- Gazeau, F., S. V. Smith, B. Gentili, M. Frankignoulle and J. P. Gattuso (2004). "The European coastal zone: characterization and first assessment of ecosystem metabolism." *Estuarine Coastal and Shelf Science* **60**(4): 673-94.
- Geanacopoulos, M. and S. Adhya (1997). "Functional characterization of roles of GalR and GalS as regulators of the gal regulon." *Journal of Bacteriology* **179**(1): 228-34.
- Gilbert, W. and B. Muller-Hill (1966). "Isolation of the lac repressor." *PNAS* **56**(6): 1891-8.
- Gimmestad, M., H. Ertesvag, T. M. Heggeset, O. Aarstad, B. I. Svanem and S. Valla (2009). "Characterization of three new *Azotobacter vinelandii* alginate lyases, one of which is involved in cyst germination." *Journal of Bacteriology* **191**(15): 4845-53.
- Giordano, M., A. Norici and R. Hell (2005). "Sulfur and phytoplankton: acquisition, metabolism and impact on the environment." *New Phytologist* **166**(2): 371-82.
- Gouet, P., X. Robert and E. Courcelle (2003). "ESPrIPT/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins." *Nucleic Acids Research* **31**(13): 3320-3.
- Gram, L., H.-P. Grossart, A. Schlingloff and T. Kiorboe (2002). "Possible Quorum Sensing in Marine Snow Bacteria: Production of Acylated Homoserine Lactones by *Roseobacter* Strains Isolated from Marine Snow." *Applied and Environmental Microbiology* **68**(8): 4111-16.
- Greenblatt, J. and R. Schleif (1971). "Arabinose C protein: regulation of the arabinose operon in vitro." *Nature New Biology* **233**(40): 166-70.
- Groisillier, A., C. Herve, A. Jeudy, E. Rebuffet, P. Pluchon, Y. Chevlot, D. Flament, C. Geslin, I. Morgado, D. Power, M. Branno, H. Moreau, G. Michel, C. Boyen and M. Czjzek (2010). "MARINE-EXPRESS: taking advantage of high throughput cloning and expression strategies for the post-genomic analysis of marine organisms." *Microbial Cell Factories* **9**(1): 45.
- Guindon, S. and O. Gascuel (2003). "A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood." *Systematic Biology* **52**(5): 696-704.
- Gury, J., L. Barthelmebs, N. P. Tran, C. Divies and J. F. Cavin (2004). "Cloning, deletion, and characterization of PadR, the transcriptional repressor of the phenolic acid decarboxylase-encoding padA gene of *Lactobacillus plantarum*." *Applied and Environmental Microbiology* **70**(4): 2146-53.
- Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." *Journal of Molecular Biology* **166**(4): 557-80.
- Hansell, D. A., C. A. Carlson, D. J. Repeta and R. Schlitzer (2009). "Dissolved organic matter in the ocean: A controversy stimulates new insights." *Oceanography* **22**(4): 202-11.
- Hansen, J. B., R. S. Doubet and J. Ram (1984). "Alginase enzyme production by *Bacillus circulans*." *Applied and Environmental Microbiology* **47**(4): 704-9.
- Hao, S., L. Su, X. Guo, T. Moyana and J. Xiang (2005). "A novel approach to tumor suppression using microencapsulated engineered J558/TNF-alpha cells." *Experimental Oncology* **27**(1): 56-60.

- Hashimoto, W., S. Kawai and K. Murata (2010). "Bacterial supersystem for alginate import/metabolism and its environmental and bioenergy applications." *Bioengineered Bugs* **1**(2): 97-109.
- Hashimoto, W., O. Miyake, K. Momma, S. Kawai and K. Murata (2000). "Molecular Identification of Oligoalginate Lyase of *Sphingomonas* sp. Strain A1 as One of the Enzymes Required for Complete Depolymerization of Alginate." *Journal of Bacteriology* **182**(16): 4572-77.
- Hatch, R. A. and N. L. Schiller (1998). "Alginate Lyase Promotes Diffusion of Aminoglycosides through the Extracellular Polysaccharide of Mucoid *Pseudomonas aeruginosa*." *Antimicrobial Agents and Chemotherapy* **42**(4): 974-77.
- Haug, A. (1961). "The affinity of some divalent metals to different types of alginates." *Acta Chemica Scandinavica* **15**: 1794-5.
- Haug, A. and B. Larsen (1963). "Solubility of Alginate at Low pH." *Acta Chemica Scandinavica* **17**(6): 1653.
- Haug, A. and B. Larsen (1969). "Biosynthesis of alginate. Epimerisation of d-mannuronic to l-guluronic acid residues in the polymer chain." *Biochimica et Biophysica Acta (BBA) - General Subjects* **192**(3): 557-59.
- Haug, A., B. Larsen and O. Smidsrod (1967). "Studies on the sequence of uronic acid residues in alginic acid." *Acta Chemica Scandinavica* **21**(3): 691-704.
- Haug, A., B. Larsen and O. Smidsrod (1974a). "Uronic Acid Sequence in Alginate from Different Sources." *Carbohydrate Research* **32**(2): 217-25.
- Haug, A., B. Larsen and O. Smidsrod (1974b). "Uronic acid sequence in alginate from different sources." *Carbohydr. Res* **32**(2): 217-25.
- Hedges, J. I., J. A. Baldock, Y. Gelin, C. Lee, M. Peterson and S. G. Wakeham (2001). "Evidence for non-selective preservation of organic matter in sinking marine particles." *Nature* **409**(6822): 801-4.
- Hehemann, J. H., G. Correc, T. Barbeyron, W. Helbert, M. Czjzek and G. Michel (2010a). "Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota." *Nature* **464**(7290): 908-12.
- Hehemann, J. H., G. Michel, T. Barbeyron and M. Czjzek (2010b). "Expression, purification and preliminary X-ray diffraction analysis of the catalytic module of a beta-agarase from the flavobacterium *Zobellia galactanivorans*." *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **66**: 413-17.
- Helmann, J. D. (1999). "Anti-sigma factors." *Current Opinion in Microbiology* **2**(2): 135-41.
- Helmann, J. D. (2002). "The extracytoplasmic function (ECF) sigma factors." *Advances in Microbial Physiology* **46**: 47-110.
- Hervé, C., S. E. Marcus and J. P. Knox (2011). "Monoclonal antibodies, carbohydrate-binding modules, and the detection of polysaccharides in plant cell walls." *Methods in Molecular Biology* **715**: 103-13.
- Heyraud, A., C. Gey, C. Leonard, C. Rochas, S. Girond and B. Kloareg (1996). "NMR spectroscopy analysis of oligoguluronates and oligomannuronates prepared by acid or enzymatic hydrolysis of homopolymeric blocks of alginic acid. Application to the determination of the substrate specificity of *Haliotis tuberculata* alginate lyase." *Carbohydrate Research* **289**: 11-23.
- Hoch, J. A. (2000). "Two-component and phosphorelay signal transduction." *Current Opinion in Microbiology* **3**(2): 165-70.
- Hoffman, P., A. Linker, V. Lippman and K. Meyer (1960). "The structure of chondroitin sulfate B from studies with *Flavobacterium* enzymes." *Journal of Biological Chemistry* **235**: 3066-9.
- Hu, X. K., X. L. Jiang, H. M. Hwang, S. L. Liu and H. S. Guan (2004). "Antitumour activities of alginate-derived oligosaccharides and their sulphated substitution derivatives." *European Journal of Phycology* **39**(1): 67-71.
- Huang, W., A. Matte, Y. Li, Y. S. Kim, R. J. Linhardt, H. Su and M. Cygler (1999). "Crystal structure of chondroitinase B from *Flavobacterium heparinum* and its complex with a disaccharide product at 1.7 Å resolution." *Journal of Molecular Biology* **294**(5): 1257-69.
- Hugouvieux-Cotte-Pattat, N., N. Blot and S. Reverchon (2001). "Identification of TogMNAB, an ABC transporter which mediates the uptake of pectic oligomers in *Erwinia chrysanthemi* 3937." *Molecular Microbiology* **41**(5): 1113-23.



- Hugouvieux-Cotte-Pattat, N., G. Condemine, W. Nasser and S. Reverchon (1996). "Regulation of pectinolysis in *Erwinia chrysanthemi*." *Annual Review of Microbiology* **50**: 213-57.
- Indergaard, M. and G. Skjakbraek (1987). "Characteristics of Alginate from *Laminaria digitata* Cultivated in a High-Phosphate Environment." *Hydrobiologia* **151**: 541-49.
- Iwamoto, Y., X. Xu, T. Tamura, T. Oda and T. Muramatsu (2003). "Enzymatically depolymerized alginate oligomers that cause cytotoxic cytokine production in human mononuclear cells." *Bioscience Biotechnology and Biochemistry* **67**(2): 258-63.
- Jackson, P. (1996). "The analysis of fluorophore-labeled carbohydrates by polyacrylamide gel electrophoresis." *Molecular Biotechnology* **5**(2): 101-23.
- Jam, M., D. Flament, J. Allouch, P. Potin, L. Thion, B. Kloareg, M. Czjzek, W. Helbert, G. Michel and T. Barbeyron (2005). "The endo-beta-agarases AgaA and AgaB from the marine bacterium *Zobellia galactanivorans*: two paralogous enzymes with different molecular organizations and catalytic behaviours." *Biochemical Journal* **385**: 703-13.
- Jordan, B. (2002). "Historical background and anticipated developments." *Microarrays, Immune Responses and Vaccines* **975**: 24-32.
- Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen and A. Krogh (2003). "Prediction of lipoprotein signal peptides in Gram-negative bacteria." *Protein Science* **12**(8): 1652-62.
- Kabsch, W. (2010). "Xds." *Acta Crystallographica Section D Biological Crystallography* **66**(Pt 2): 125-32.
- Katoh, K. and H. Toh (2008). "Recent developments in the MAFFT multiple sequence alignment program." *Briefings in Bioinformatics* **9**(4): 286-98.
- Kawada, A., N. Hiura, M. Shiraiwa, S. Tajima, M. Hiruma, K. Hara, A. Ishibashi and H. Takahara (1997). "Stimulation of human keratinocyte growth by alginate oligosaccharides, a possible co-factor for epidermal growth factor in cell culture." *FEBS Letters* **408**(1): 43-6.
- Kidby, D. K. and D. J. Davidson (1973). "A convenient ferricyanide estimation of reducing sugars in the nanomole range." *Analytical Biochemistry* **55**(1): 321-5.
- Kiorboe, T., K. Tang, H.-P. Grossart and H. Ploug (2003). "Dynamics of Microbial Communities on Marine Snow Aggregates: Colonization, Growth, Detachment, and Grazing Mortality of Attached Bacteria." *Applied and Environmental Microbiology* **69**(6): 3036-47.
- Kirchman, D. L. (2002). "A primer on dissolved organic material and heterotrophic prokaryotes in the oceans." *The Ocean Carbon Cycle and Climate: Proceedings of the NATO Asi on Ocean Carbon Cycle and Climate*: 31-63.
- Kirchman, D. L. (2008). Introduction and overview. *Microbial Ecology of the Oceans*. D. L. Kirchman. Hoboken, New Jersey, Wiley & sons: 1-26.
- Kloareg, B. (1984). "Isolation and Analysis of Cell Walls of the Brown Marine Algae *Pelvetia canaliculata* and *Ascophyllum nodosum*." *Physiologie Vegetale* **22**(1): 47-56.
- Kloareg, B. and R. S. Quatrano (1988). "Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides." *Oceanography and Marine Biology Annual Review* **26**: 259-315.
- Kobayashi, T., K. Uchimura, M. Miyazaki, Y. Nogi and K. Horikoshi (2009). "A new high-alkaline alginate lyase from a deep-sea bacterium *Agarivorans* sp." *Extremophiles* **13**(1): 121-29.
- Koehn, J. and I. Hunt (2009). "High-Throughput Protein Production (HTPP): a review of enabling technologies to expedite protein production." *Methods in Molecular Biology* **498**: 1-18.
- Koning, S. M., M. G. Elferink, W. N. Konings and A. J. Driessen (2001). "Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter." *Journal of Bacteriology* **183**(17): 4979-84.
- Kraiwattanapong, J., K. Motomura, T. Ooi and S. Kinoshita (1999). "Characterization of alginate lyase (ALYII) from *Pseudomonas* sp. OS-ALG-9 expressed in recombinant *Escherichia coli*." *World Journal of Microbiology and Biotechnology* **15**(1): 105-09.
- Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." *Journal of Molecular Biology* **372**(3): 774-97.
- Krogh, A., B. Larsson, G. von Heijne and E. L. Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." *Journal of Molecular Biology* **305**(3): 567-80.

- Kujawinski, E. B. (2011). "The impact of microbial metabolism on marine dissolved organic matter. ." *Annual Reviews of Marine Science* **3**: 567-99.
- Kuorelahti, S., P. Jouhten, H. Maaheimo, M. Penttila and P. Richard (2006). "L-galactonate dehydratase is part of the fungal path for D-galacturonic acid catabolism." *Molecular Microbiology* **61**(4): 1060-8.
- Küpper, F. C., E. Gaquerel, E.-M. Boneberg, S. Morath, J.-P. Salaun and P. Potin (2006). "Early events in the perception of lipopolysaccharides in the brown alga *Laminaria digitata* include an oxidative burst and activation of fatty acid oxidation cascades." *Journal of Experimental Botany* **57**(9): 1991-99.
- Küpper, F. C., B. Kloareg, J. Guern and P. Potin (2001). "Oligogulonates elicit an oxidative burst in the brown algal kelp *Laminaria digitata*." *Plant Physiology* **125**(1): 278-91.
- Küpper, F. C., D. G. Müller, A. F. Peters, B. Kloareg and P. Potin (2002). "Oligoalginat recognition and oxidative burst play a key role in natural and induced resistance of sporophytes of Laminariales." *Journal of Chemical Ecology* **28**(10): 2057-81.
- Kvint, K., L. Nachin, A. Diez and T. Nystrom (2003). "The bacterial universal stress protein: function and regulation." *Current Opinion in Microbiology* **6**(2): 140-45.
- Lamppa, J. W., M. E. Ackerman, J. I. Lai, T. C. Scanlon and K. E. Griswold (2011). "Genetically engineered alginate lyase-PEG conjugates exhibit enhanced catalytic function and reduced immunoreactivity." *PLoS One* **6**(2): e17042.
- Lee, S.-M. and J.-H. Lee (2011). "The isolation and characterization of simultaneous saccharification and fermentation microorganisms for *Laminaria japonica* utilization." *Bioresource Technology* **102**(10): 5962-67.
- Li, B., F. Lu, X. Wei and R. Zhao (2008). "Fucoidan: structure and bioactivity." *Molecules* **13**(8): 1671-95.
- Lin, T.-Y. and W. Z. Hassid (1966). "Pathway of Alginic Acid Synthesis in the Marine Brown Alga, *Fucus gardneri* Silva." *Journal of Biological Chemistry* **241**(22): 5284-97.
- Lombard, V., T. Bernard, C. Rancurel, H. Brumer, Pedro M. Coutinho and B. Henrissat (2010). "A hierarchical classification of polysaccharide lyases for glycogenomics." *Biochemical Journal* **432**(3): 437-44.
- Mabeau, S. and B. Kloareg (1987). "Isolation and Analysis of the Cell Walls of Brown Algae - *Fucus spiralis*, *Fucus ceranoides*, *Fucus vesiculosus*, *Fucus serratus*, *Bifurcaria bifurcata* and *Laminaria digitata*." *Journal of Experimental Botany* **38**(194): 1573-80.
- Madgwick, J., A. Haug and B. Larsen (1973). "Alginate Lyase in Brown Alga *Laminaria digitata* (Huds) Lamour." *Acta Chemica Scandinavica* **27**(2): 711-12.
- Madgwick, J., A. Haug and B. Larsen (1978). "Ionic Requirements of Alginate-Modifying Enzymes in Marine Alga *Pelvetia canaliculata* (L) Dcne Et Thur." *Botanica Marina* **21**(1): 1-3.
- Maere, S., K. Heymans and M. Kuiper (2005). "BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks." *Bioinformatics* **21**(16): 3448-9.
- Maki, H., A. Mori, K. Fujiyama, S. Kinoshita and T. Yoshida (1993). "Cloning, sequence analysis and expression in *Escherichia coli* of a gene encoding an alginate lyase from *Pseudomonas* sp. OS-ALG-9." *Journal of General Microbiology* **139**(5): 987-93.
- Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng and S. H. Bryant (2011). "CDD: a Conserved Domain Database for the functional annotation of proteins." *Nucleic Acids Research* **39**: D225-9.
- Martens, E. C., N. M. Koropatkin, T. J. Smith and J. I. Gordon (2009). "Complex glycan catabolism by the human gut microbiota: The *Bacteroidetes* Sus-like paradigm." *Journal of Biological Chemistry* **284**(37): 24673-7.
- Martinez, J., D. C. Smith, G. F. Steward and F. Azam (1996). "Variability in ectohydrolytic enzyme activities of pelagic marine bacteria and its significance for substrate processing in the sea." *Aquatic Microbial Ecology* **10**(3): 223-30.

- Matricardi, P., C. D. Meo, T. Coviello and F. Alhaique (2008). "Recent advances and perspectives on coated alginate microspheres for modified drug delivery." *Expert Opinion on Drug Delivery* **5**(4): 417-25.
- Matsubara, Y., R. Kawada, K. Iwasaki, T. Oda and T. Muramatsu (1998). "Extracellular poly(alpha-L-gulonate)lyase from *Corynebacterium* sp.: purification, characteristics, and conformational properties." *Journal of Protein Chemistry* **17**(1): 29-36.
- Matsuo, Y., H. Imagawa, M. Nishizawa and Y. Shizuri (2005). "Isolation of an algal morphogenesis inducer from a marine bacterium." *Science* **307**(5715): 1598-98.
- Matsuo, Y., M. Suzuki, H. Kasai, Y. Shizuri and S. Harayama (2003). "Isolation and phylogenetic characterization of bacteria capable of inducing differentiation in the green alga *Monostroma oxyspermum*." *Environmental Microbiology* **5**(1): 25-35.
- Matte, A., J. Sivaraman, I. Ekiel, K. Gehring, Z. Jia and M. Cygler (2003). "Contribution of Structural Genomics to Understanding the Biology of *Escherichia coli*." *Journal of Bacteriology* **185**(14): 3994-4002.
- Matthews, B. W. (1968). "Solvent content of protein crystals." *Journal of Molecular Biology* **33**(2): 491-7.
- May, T. B., D. Shinabarger, R. Maharaj, J. Kato, L. Chu, J. D. DeVault, S. Roychoudhury, N. A. Zielinski, A. Berry and R. K. Rothmel (1991). "Alginate synthesis by *Pseudomonas aeruginosa*: a key pathogenic factor in chronic pulmonary infections of cystic fibrosis patients." *Clinical Microbiology Reviews* **4**(2): 191-206.
- McBride, M. J. (2001). "Bacterial gliding motility: Multiple mechanisms for cell movement over surfaces." *Annual Review of Microbiology* **55**: 49-75.
- McBride, M. J., T. F. Braun and J. L. Brust (2003). "*Flavobacterium johnsoniae* GldH is a lipoprotein that is required for gliding motility and chitin utilization." *Journal of Bacteriology* **185**(22): 6648-57.
- McCarren, J., J. W. Becker, D. J. Repeta, Y. Shi, C. R. Young, R. R. Malmstrom, S. W. Chisholm and E. F. DeLong (2010). "Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea." *PNAS* **107**(38): 16420-7.
- McCartney, L., A. W. Blake, J. Flint, D. N. Bolam, A. B. Boraston, H. J. Gilbert and J. P. Knox (2006). "Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules." *PNAS* **103**(12): 4765-70.
- McGraw, D. J. (2006). "The founding of modern marine microbiology: Claude Zobell and his Magnum Opus, 1946." *Newsletter of the Commission of the History of Oceanography* **18**: 5-18.
- McHugh, D. J. (2003). *A guide to seaweed industry*. Rome, Food and Agriculture Organization.
- McKinney, T. K., V. K. Sharma, W. A. Craig and G. L. Archer (2001). "Transcription of the gene mediating methicillin resistance in *Staphylococcus aureus* (*mecA*) is corepressed but not coinduced by cognate *mecA* and beta-lactamase regulators." *Journal of Bacteriology* **183**(23): 6862-8.
- Michel, C. and G. T. Macfarlane (1996). "Digestive fates of soluble polysaccharides from marine macroalgae: involvement of the colonic microflora and physiological consequences for the host." *Journal of Applied Microbiology* **80**(4): 349-69.
- Michel, G., T. Barbeyron, B. Kloareg and M. Czjzek (2009). "The family 6 carbohydrate-binding modules have coevolved with their appended catalytic modules toward similar substrate specificity." *Glycobiology* **19**(6): 615-23.
- Michel, G., K. Pojasek, Y. G. Li, T. Sulea, R. J. Linhardt, R. Raman, V. Prabhakar, R. Sasisekharan and M. Cygler (2004). "The structure of chondroitin B lyase complexed with glycosaminoglycan oligosaccharides unravels a calcium-dependent catalytic machinery." *Journal of Biological Chemistry* **279**(31): 32882-96.
- Michel, G., T. Tonon, D. Scornet, J. M. Cock and B. Kloareg (2010). "The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes." *New Phytologist* **188**(1): 82-97.

- Min, K. H., S. F. Sasaki, Y. Kashiwabara, M. Umekawa and K. Nisizawa (1977). "Fine-Structure of Smg Alginate Fragment in Light of Its Degradation by Alginate Lyases of *Pseudomonas* Sp." *Journal of Biochemistry* **81**(3): 555-62.
- Miroux, B. and J. E. Walker (1996). "Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." *Journal of Molecular Biology* **260**(3): 289-98.
- Moran, M. A. (2008). Genomics and Metagenomics of Marine Prokaryotes. *Microbial Ecology of the Oceans*, John Wiley & Sons, Inc.: 91-129.
- Moran, M. A. (2009). "Metatranscriptomics: Eavesdropping on Complex Microbial Communities." *Microbe* **4**(7): 329-35.
- Muramatsu, T. and K. Egawa (1982). "Alginate-Degrading Enzymes .5. Chemical Composition of Alginate Lyases from the Midgut Gland of *Turbo cornutus* and Chemical Modification of Cys, Trp and Lys Residues." *Agricultural and Biological Chemistry* **46**(4): 883-89.
- Murata, K., T. Inose, T. Hisano, S. Abe, Y. Yonemoto, T. Yamashita, M. Takagi, K. Sakaguchi, A. Kimura and T. Imanaka (1993). "Bacterial alginate lyase - Enzymology, genetics and application." *Journal of Fermentation and Bioengineering* **76**(5): 427-37.
- Murata, K., S. Kawai, B. Mikami and W. Hashimoto (2008). "Superchannel of bacteria: biological significance and new horizons." *Bioscience Biotechnology and Biochemistry* **72**(2): 265-77.
- Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of macromolecular structures by the maximum-likelihood method." *Acta Crystallographica Section D Biological Crystallography* **53**(Pt 3): 240-55.
- Nachin, L., U. Nannmark and T. Nystrom (2005). "Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility." *Journal of Bacteriology* **187**(18): 6265-72.
- Nagata, T. (2008). Organic matter-bacteria interactions in seawater. *Microbial Ecology in the Oceans*. D. L. Kirchman. Hoboken, New Jersey, Wiley & sons: 207-42.
- Nakagawa, A., T. Ozaki, K. Chubachi, T. Hosoyama, T. Okubo, S. Iyobe and T. Suzuki (1998). "An effective method for isolating alginate lyase-producing *Bacillus* sp. ATB-1015 strain and purification and characterization of the lyase." *Journal of Applied Microbiology* **84**(3): 328-35.
- Natsume, M., Y. Kamo, M. Hirayama and T. Adachi (1994). "Isolation and Characterization of Alginate-Derived Oligosaccharides with Root Growth-Promoting Activities." *Carbohydrate Research* **258**: 187-97.
- Nedashkovskaya, O. I., M. Suzuki, M. Vancanneyt, I. Cleenwerck, A. M. Lysenko, V. V. Mikhailov and J. Swings (2004). "*Zobellia amurskyensis* sp nov., *Zobellia laminariae* sp nov and *Zobellia russellii* sp nov., novel marine bacteria of the family Flavobacteriaceae." *International Journal of Systematic and Evolutionary Microbiology* **54**: 1643-48.
- Nelson, S. S. and M. J. McBride (2006). "Mutations in *Flavobacterium johnsoniae* secDF result in defects in gliding motility and chitin utilization." *Journal of Bacteriology* **188**(1): 348-51.
- Nelson, T. E., S. Kirkwood, J. V. Scaletti and F. Smith (1963). "Use of Enzymes in Structural Studies on Polysaccharides .1. Mode of Attack of a Beta-D-(1 -3)-Glucanase on Laminarin." *Canadian Journal of Chemistry* **41**(7): 1671.
- Nicholson, R. L. and R. Hammerschmidt (1992). "Phenolic-Compounds and Their Role in Disease Resistance." *Annual Review of Phytopathology* **30**: 369-89.
- Norderhaug, K. M., S. Fredriksen and K. Nygaard (2003). "Trophic importance of *Laminaria hyperborea* to kelp forest consumers and the importance of bacterial degradation to food quality." *Marine Ecology-Progress series* **255**: 135-44.
- Novotny, A. M. and M. Forman (1975). "Composition and Development of Cell Walls of *Fucus* Embryos." *Planta* **122**(1): 67-78.
- Nystrom, T. and F. C. Neidhardt (1994). "Expression and Role of the Universal Stress Protein, Uspa, of *Escherichia coli* during Growth Arrest." *Molecular Microbiology* **11**(3): 537-44.
- Ochiai, A., W. Hashimoto and K. Murata (2006). "A biosystem for alginate metabolism in *Agrobacterium tumefaciens* strain C58: molecular identification of Atu3025 as an exotype family PL-15 alginate lyase." *Research in Microbiology* **157**(7): 642-9.

- Ochiai, A., T. Itoh, A. Kawamata, W. Hashimoto and K. Murata (2007). "Plant cell wall degradation by saprophytic *Bacillus subtilis* strains: gene clusters responsible for rhamnogalacturonan depolymerization." *Applied and Environmental Microbiology* **73**(12): 3803-13.
- Ochiai, A., M. Yamasaki, B. Mikami, W. Hashimoto and K. Murata (2010). "Crystal structure of exotype alginate lyase Atu3025 from *Agrobacterium tumefaciens*." *Journal of Biological Chemistry* **285**(32): 24519-28.
- Ogura, K., M. Yamasaki, B. Mikami, W. Hashimoto and K. Murata (2008). "Substrate recognition by family 7 alginate lyase from *Sphingomonas* sp. A1." *Journal of Molecular Biology* **380**(2): 373-85.
- Ogura, K., M. Yamasaki, T. Yamada, B. Mikami, W. Hashimoto and K. Murata (2009). "Crystal structure of family 14 polysaccharide lyase with pH-dependent modes of action." *Journal of Biological Chemistry* **284**(51): 35572-9.
- Okazaki, M., K. Furuya, K. Tsukayama and K. Nisizawa (1982). "Isolation and Identification of Alginic Acid from a Calcareous Red Alga *Serraticardia maxima*." *Botanica Marina* **25**(3): 123-32.
- Osawa, T., Y. Matsubara, T. Muramatsu, M. Kimura and Y. Kakuta (2005). "Crystal structure of the alginate (poly alpha-l-guluronate) lyase from *Corynebacterium* sp. at 1.2 Å resolution." *Journal of Molecular Biology* **345**(5): 1111-8.
- Ostgaard, K. (1993). "Determination of Alginate Composition by a Simple Enzymatic Assay." *Hydrobiologia* **261**: 513-20.
- Paetzel, M., A. Karla, N. C. Strynadka and R. E. Dalbey (2002). "Signal peptidases." *Chemical Reviews* **102**(12): 4549-80.
- Pannier, A. K. and L. D. Shea (2004). "Controlled release systems for DNA delivery." *Molecular Therapy* **10**(1): 19-26.
- Perrakis, A., T. K. Sixma, K. S. Wilson and V. S. Lamzin (1997). "wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models." *Acta Crystallographica Section D Biological Crystallography* **53**(Pt 4): 448-55.
- Plumbridge, J. and O. Pellegrini (2004). "Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP." *Molecular Microbiology* **52**(2): 437-49.
- Pomeroy, L. R., P. J. Williams, F. Azam and J. Hobbie (2007). "The microbial loop." *Oceanography* **20**(2): 28-33.
- Popper, Z., G. Michel, C. Herve, D. S. Domozych, W. G. Willats, M. G. Tuohy, B. Kloareg and D. B. Stengel (2010). "Evolution and Diversity of Plant Cell Walls: From Algae to Flowering Plants." *Annual Review of Plant Biology* **62**: 567-90.
- Poretsky, R. S., S. Sun, X. Mou and M. A. Moran (2010). "Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon." *Environmental Microbiology* **12**(3): 616-27.
- Potin, P. (1992). Recherche, production, purification et caractérisation de galactane-hydrolases pour la préparation d'oligosaccharides des parois d'algues rouges. Roscoff, Université de Bretagne Occidentale. **Thèse de doctorat**.
- Potin, P., A. Sanseau, Y. Le Gall, C. Rochas and B. Kloareg (1991). "Purification and characterization of a new k-carrageenase from a marine *Cytophaga*-like bacterium." *European Journal of Biochemistry* **201**(1): 241-47.
- Preiss, J. and G. Ashwell (1962a). "Alginic acid metabolism in bacteria. I. Enzymatic formation of unsaturated oligosaccharides and 4-deoxy-L-erythro-5-hexoseulose uronic acid." *Journal of Biological Chemistry* **237**: 309-16.
- Preiss, J. and G. Ashwell (1962b). "Alginic acid metabolism in bacteria. II. The enzymatic reduction of 4-deoxy-L-erythro-5-hexoseulose uronic acid to 2-keto-3-deoxy-D-gluconic acid." *Journal of Biological Chemistry* **237**: 317-21.
- Preiss, J. and G. Ashwell (1963). "Polygalacturonic acid metabolism in bacteria. I. Enzymatic formation of 4-deoxy-L-threo-5-hexoseulose uronic acid." *Journal of Biological Chemistry* **238**: 1571-83.
- Provasoli, L. (1958). "Effect of Plant Hormones on Ulva." *Biological Bulletin* **114**(3): 375-84.

- Quatrano, R. S. and P. T. Stevens (1976). "Cell wall assembly in *Fucus* zygotes: I. Characterization of the polysaccharide components." *Plant Physiology* **58**(2): 224-31.
- Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan and R. Tobes (2005). "The TetR family of transcriptional repressors." *Microbiology and Molecular Biology Reviews* **69**(2): 326-56.
- Ramsey, D. M. and D. J. Wozniak (2005). "Understanding the control of *Pseudomonas aeruginosa* alginate synthesis and the prospects for management of chronic infections in cystic fibrosis." *Molecular Microbiology* **56**(2): 309-22.
- Rebuffet, E., T. Barbeyron, A. Jeudy, M. Jam, M. Czjzek and G. Michel (2010). "Identification of catalytic residues and mechanistic analysis of family GH82 iota-carrageenases." *Biochemistry* **49**(35): 7590-9.
- Rebuffet, E., A. Groisillier, A. Thompson, A. Jeudy, T. Barbeyron, M. Czjzek and G. Michel (2011). "Discovery and structural characterization of a novel glycosidase family of marine origin." *Environmental Microbiology* **13**(5): 1253-70.
- Reddy, C., M. Gupta, V. Mantri and B. Jha (2008). "Seaweed protoplasts: status, biotechnological perspectives and needs." *Journal of Applied Phycology* **20**(5): 619-32.
- Rehm, B. H. A. and S. Valla (1997). "Bacterial alginates: biosynthesis and applications." *Applied Microbiology and Biotechnology* **48**(3): 281-88.
- Reidl, J., K. Romisch, M. Ehrmann and W. Boos (1989). "MalI, a novel protein involved in regulation of the maltose system of *Escherichia coli*, is highly homologous to the repressor proteins GalR, CytR, and LacI." *Journal of Bacteriology* **171**(9): 4888-99.
- Reyes-Prieto, A., A. P. Weber and D. Bhattacharya (2007). "The origin and establishment of the plastid in algae and plants." *Annual Review of Genetics* **41**: 147-68.
- Rhodes, R. G., M. N. Samarasinghe, A. Shrivastava, J. M. van Baaren, S. Pochiraju, S. Bollampalli and M. J. McBride (2010). "*Flavobacterium johnsoniae* gldN and gldO are partially redundant genes required for gliding motility and surface localization of SprB." *Journal of Bacteriology* **192**(5): 1201-11.
- Richmond, C. S., J. D. Glasner, R. Mau, H. Jin and F. R. Blattner (1999). "Genome-wide expression profiling in *Escherichia coli* K-12." *Nucleic Acids Research* **27**(19): 3821-35.
- Roussel, X. (2009). Enzymologie moléculaire d'une sulfanyl réductase, la sulfirédoxine : caractérisation du mécanisme catalytique. Nancy, Université Henri Poincaré, Nancy I. **Thèse de doctorat**.
- Ruckert, C., J. Milse, A. Albersmeier, D. J. Koch, A. Puhler and J. Kalinowski (2008). "The dual transcriptional regulator CysR in *Corynebacterium glutamicum* ATCC 13032 controls a subset of genes of the McbR regulon in response to the availability of sulphide acceptor molecules." *BMC Genomics* **9**: 483.
- Sabra, W., A. P. Zeng and W. D. Deckwer (2001). "Bacterial alginate: physiology, product quality and process aspects." *Applied Microbiology and Biotechnology* **56**(3-4): 315-25.
- Saier, M. H., Jr. and T. M. Ramseier (1996). "The catabolite repressor/activator (Cra) protein of enteric bacteria." *Journal of Bacteriology* **178**(12): 3411-7.
- Saito, A. and H. Schrempf (2004). "Mutational analysis of the binding affinity and transport activity for N-acetylglucosamine of the novel ABC transporter Ngc in the chitin-degrader *Streptomyces olivaceoviridis*." *Molecular Genetics and Genomics* **271**(5): 545-53.
- Salaün, S. (2009). Interactions entre la macroalgue brune *Laminaria digitata* et ses épibiontes bactériens : Etudes moléculaire et spectroscopiques, et capacité d'adhésion et de formation de biofilm. Roscoff, Université de Bretagne Sud. **Thèse de doctorat**.
- Salgado, L. T., L. P. Cinelli, N. B. Viana, R. T. de Carvalho, P. A. D. Mourao, V. L. Teixeira, M. Farina and G. M. Amado (2009). "A Vanadium Bromoperoxidase Catalyzes the Formation of High-Molecular-Weight Complexes between Brown Algal Phenolic Substances and Alginates." *Journal of Phycology* **45**(1): 193-202.
- Sato, K., M. Naito, H. Yukitake, H. Hirakawa, M. Shoji, M. J. McBride, R. G. Rhodes and K. Nakayama (2010). "A protein secretion system linked to bacteroidete gliding motility and pathogenesis." *PNAS* **107**(1): 276-81.

- Sawabe, T. and Y. Ezura (1996). "Regeneration from *Laminaria japonica* Areschoug (Laminariales, Phaeophyceae) protoplasts isolated with bacterial alginase." *Plant Cell Reports* **15**(12): 892-95.
- Sawabe, T., Y. Ezura and T. Kimura (1992). "Characterization of an alginolytic marine bacterium from decaying Rishiri-Kombu *Laminaria japonica* var. *ochotensis*." *Nippon Suisan Gakkaishi* **58**(1): 141-45.
- Sawabe, T., Y. Oda, Y. Shiomi and Y. Ezura (1995). "Alginate degradation by bacteria isolated from the gut of sea urchins and abalones." *Microbial Ecology* **30**(2): 193-202.
- Sawabe, T., I. Sugimura, M. Ohtsuka, K. Nakano, K. Tajima, Y. Ezura and R. Christen (1998). "*Vibrio halioticoli* sp. nov., a non-motile alginolytic marine bacterium isolated from the gut of the abalone *Haliotis discus hannai*." *International Journal of Systematic Bacteriology* **48**(2): 573-80.
- Schaal, G., P. Riera and C. Leroux (2009). "Trophic significance of the kelp *Laminaria digitata* (Lamour.) for the associated food web: a between-sites comparison." *Estuarine Coastal and Shelf Science* **85**(4): 565-72.
- Schaal, G., P. Riera and C. Leroux (2010). "Trophic ecology in a Northern Brittany (Batz Island, France) kelp (*Laminaria digitata*) forest, as investigated through stable isotopes and chemical assays." *Journal of Sea Research* **63**(1): 24-35.
- Schena, M., D. Shalon, R. W. Davis and P. O. Brown (1995). "Quantitative Monitoring of Gene-Expression Patterns with a Complementary-DNA Microarray." *Science* **270**(5235): 467-70.
- Schmidt, A. and K. Jager (1992). "Open Questions About Sulfur Metabolism in Plants." *Annual Review of Plant Physiology and Plant Molecular Biology* **43**: 325-49.
- Schneider, E. (2001). "ABC transporters catalyzing carbohydrate uptake." *Research in Microbiology* **152**(3-4): 303-10.
- Schoenwaelder, M. E. A. and C. Wiencke (2000). "Phenolic compounds in the embryo development of several northern hemisphere fucoids." *Plant Biology* **2**(1): 24-33.
- Schulze, A. D., A. O. Alabi, A. R. Tattersall-Sheldrake and K. M. Miller (2006). "Bacterial diversity in a marine hatchery: Balance between pathogenic and potentially probiotic bacterial strains." *Aquaculture* **256**(1-4): 50-73.
- Shalon, D., S. J. Smith and P. O. Brown (1996). "A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization." *Genome Research* **6**(7): 639-45.
- Shimokawa, T., S. Yoshida, I. Kusakabe, T. Takeuchi and K. Murata (1997). "Isolation and quantification of alginate-derived oligouronic acids by fluorophore-assisted carbohydrate electrophoresis." *Carbohydrate Research* **299**(1-2): 95-98.
- Shiraiwa, Y., K. Abe, S. F. Sasaki, T. Ikawa and K. Nisizawa (1975). "Alginate Lyase Activities in Extracts from Several Brown Algae." *Botanica Marina* **18**(2): 97-104.
- Simon, A. and E. Biot (2010). "ANAIS: analysis of NimbleGen arrays interface." *Bioinformatics* **26**(19): 2468-9.
- Skerratt, J. H., J. P. Bowman, G. Hallegraef, S. James and P. D. Nichols (2002). "Algicidal bacteria associated with blooms of a toxic dinoflagellate in a temperate Australian estuary." *Marine Ecology-Progress series* **244**: 1-15.
- Skriptsova, A., V. Khomenko and V. Isakov (2004). "Seasonal changes in growth rate, morphology and alginate content in *Undaria pinnatifida* at the northern limit in the Sea of Japan (Russia)." *Journal of Applied Phycology* **16**(1): 17-21.
- Smith, S. V. (1981). "Marine Macrophytes as a Global Carbon Sink." *Science* **211**(4484): 838-40.
- Sonnhammer, E. L., G. von Heijne and A. Krogh (1998). "A hidden Markov model for predicting transmembrane helices in protein sequences." *Proceedings of the International Conference on Intelligent Systems for Molecular Biology* **6**: 175-82.
- Spence, C., W. G. Wells and C. J. Smith (2006). "Characterization of the primary starch utilization operon in the obligate anaerobe *Bacteroides fragilis*: Regulation by carbon source and oxygen." *Journal of Bacteriology* **188**(13): 4663-72.
- Spiro, M. D., B. L. Ridley, S. Eberhard, K. A. Kates, Y. Mathieu, M. A. O'Neill, D. Mohnen, J. Guern, A. Darvill and P. Albersheim (1998). "Biological activity of reducing-end-derivatized oligogalacturonides in tobacco tissue cultures." *Plant Physiology* **116**(4): 1289-98.

- Staufenberger, T., V. Thiel, J. Wiese and J. F. Imhoff (2008). "Phylogenetic analysis of bacteria associated with *Laminaria saccharina*." *FEMS Microbiology Ecology* **64**(1): 65-77.
- Stevens, R. A. and R. E. Levin (1977). "Purification and characteristics of an alginase from *Alginovibrio aquatilis*." *Applied and Environmental Microbiology* **33**(5): 1156-61.
- Stock, A. M., V. L. Robinson and P. N. Goudreau (2000). "Two-component signal transduction." *Annual Review of Biochemistry* **69**: 183-215.
- Stocker, R., J. R. Seymour, A. Samadani, D. E. Hunt and M. F. Polz (2008). "Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches." *PNAS* **105**(11): 4209-14.
- Studier, F. W. (2005). "Protein production by auto-induction in high density shaking cultures." *Protein Expression and Purification* **41**(1): 207-34.
- Suzuki, H., K.-i. Suzuki, A. Inoue and T. Ojima (2006). "A novel oligoalginase from abalone, *Haliotis discus hannai*, that releases disaccharide from alginate polymer in an exolytic manner." *Carbohydrate Research* **341**(11): 1809-19.
- Svanem, B. I. G., W. I. Strand, H. Ertesvåg, G. Skjåk-Bræk, M. Hartmann, T. Barbeyron and S. Valla (2001). "The Catalytic Activities of the Bifunctional *Azotobacter vinelandii* Mannuronan C-5-Epimerase and Alginate Lyase AlgE7 Probably Originate from the Same Active Site in the Enzyme." *Journal of Biological Chemistry* **276**(34): 31542-50.
- Szumant, H., R. A. White and J. A. Hoch (2007). "Sensor complexes regulating two-component signal transduction." *Current Opinion in Structural Biology* **17**(6): 706-15.
- Takase, R., A. Ochiai, B. Mikami, W. Hashimoto and K. Murata (2010). "Molecular identification of unsaturated uronate reductase prerequisite for alginate metabolism in *Sphingomonas* sp. A1." *Biochimica et Biophysica Acta* **1804**(9): 1925-36.
- Tamura, K., J. Dudley, M. Nei and S. Kumar (2007). "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." *Molecular Biology and Evolution* **24**(8): 1596-9.
- Tanaka, K. and S. Sorai (1970). "Hydrolysis of fucoidan by abalone liver alpha-L-fucosidase." *FEBS Letters* **9**(1): 45-48.
- Tao, H., C. Bausch, C. Richmond, F. R. Blattner and T. Conway (1999). "Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media." *Journal of Bacteriology* **181**(20): 6425-40.
- Taylor, B. L. and I. B. Zhulin (1999). "PAS domains: internal sensors of oxygen, redox potential, and light." *Microbiology and Molecular Biology Reviews* **63**(2): 479-506.
- Tenhaken, R., E. Voglas, J. M. Cock, V. Neu and C. G. Huber (2011). "Characterization of GDP-mannose dehydrogenase from the brown alga *Ectocarpus siliculosus* providing the precursor for the alginate polymer." *Journal of Biological Chemistry* **286**(19): 16707-15.
- Thomas, F., T. Barbeyron and G. Michel (2011). "Evaluation of reference genes for real-time quantitative PCR in the marine flavobacterium *Zobellia galactanivorans*." *Journal of Microbiological Methods* **84**(1): 61-6.
- Tobin, J. F. and R. F. Schleif (1990). "Transcription from the rha operon psr promoter." *Journal of Molecular Biology* **211**(1): 1-4.
- Tran, N. P., J. Gury, V. Dartois, T. K. Nguyen, H. Seraut, L. Barthelmebs, P. Gervais and J. F. Cavin (2008). "Phenolic acid-mediated regulation of the padC gene, encoding the phenolic acid decarboxylase of *Bacillus subtilis*." *Journal of Bacteriology* **190**(9): 3213-24.
- Uchimura, K., M. Miyazaki, Y. Nogi, T. Kobayashi and K. Horikoshi (2010). "Cloning and sequencing of alginate lyase genes from deep-sea strains of *Vibrio* and *Agarivorans* and characterization of a new *Vibrio* enzyme." *Marine Biotechnology (NY)* **12**(5): 526-33.
- Vanfossen, A. L., I. Ozdemir, S. L. Zelin and R. M. Kelly (2011). "Glycoside hydrolase inventory drives plant polysaccharide deconstruction by the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*." *Biotechnology and Bioengineering* **108**(7): 1559-69.
- Verdugo, P., A. L. Alldredge, F. Azam, D. L. Kirchman, U. Passow and P. H. Santschi (2004). "The oceanic gel phase: a bridge in the DOM-POM continuum." *Marine Chemistry* **92**: 67-85.
- Watanabe, T. and K. Nishizawa (1982). "Investigation of Alginate Lyase Activities in the Fronds of *Undaria pinnatifida* Treated with Ashes." *Bulletin of the Japanese Society of Scientific Fisheries* **48**(2): 237-41.



- Weickert, M. J. and S. Adhya (1992). "A family of bacterial regulators homologous to Gal and Lac repressors." *Journal of Biological Chemistry* **267**(22): 15869-74.
- Weiner, R. M., L. E. Taylor, B. Henrissat, L. Hauser, M. Land, P. M. Coutinho, C. Rancurel, E. H. Saunders, A. G. Longmire, H. Zhang, E. A. Bayer, H. J. Gilbert, F. Larimer, I. B. Zhulin, N. A. Ekborg, R. Lamed, P. M. Richardson, I. Borovok and S. Hutcheson (2008). "Complete Genome Sequence of the Complex Carbohydrate-Degrading Marine Bacterium, *Saccharophagus degradans* Strain 2-40T." *PLoS Genetics* **4**(5): e1000087.
- Whitman, W. B., D. C. Coleman and W. J. Wiebe (1998). "Prokaryotes: the unseen majority." *PNAS* **95**(12): 6578-83.
- Whittaker, R. H. and G. E. Likens (1973). "Carbon in Biota." *Brookhaven Symposia in Biology*(24): 281-302.
- Wong, T. Y., L. A. Preston and N. L. Schiller (2000). "Alginate lyase: Review of major sources and enzyme characteristics, structure-function analysis, biological roles, and applications." *Annual review of Microbiology* **54**: 289-340.
- Wyss, O., M. G. Neumann and M. D. Socolofsky (1961). "Development and germination of the *Azotobacter* cyst." *Journal of Biophysical and Biochemical Cytology* **10**: 555-65.
- Yamamoto, Y., M. Kurachi, K. Yamaguchi and T. Oda (2007). "Stimulation of multiple cytokine production in mice by alginate oligosaccharides following intraperitoneal administration." *Carbohydrate Research* **342**(8): 1133-37.
- Yamasaki, M., S. Moriwaki, O. Miyake, W. Hashimoto, K. Murata and B. Mikami (2004). "Structure and function of a hypothetical *Pseudomonas aeruginosa* protein PA1167 classified into family PL-7: a novel alginate lyase with a beta-sandwich fold." *Journal of Biological Chemistry* **279**(30): 31863-72.
- Yamasaki, M., K. Ogura, W. Hashimoto, B. Mikami and K. Murata (2005). "A structural basis for depolymerization of alginate by polysaccharide lyase family-7." *Journal of Molecular Biology* **352**(1): 11-21.
- Yasuhara, T. and I. Date (2007). "Intracerebral transplantation of genetically engineered cells for Parkinson's disease: toward clinical application." *Cell Transplantation* **16**(2): 125-32.
- Yoon, H.-J., W. Hashimoto, O. Miyake, M. Okamoto, B. Mikami and K. Murata (2000). "Overexpression in *Escherichia coli*, Purification, and Characterization of *Sphingomonas* sp. A1 Alginate Lyases." *Protein Expression and Purification* **19**(1): 84-90.
- Yoon, H. J., W. Hashimoto, O. Miyake, K. Murata and B. Mikami (2001). "Crystal structure of alginate lyase A1-III complexed with trisaccharide product at 2.0 Å resolution." *Journal of Molecular Biology* **307**(1): 9-16.
- Zablackis, E. and J. Perez (1990). "A Partially Pyruvated Carrageenan from Hawaiian *Grateloupia filicina* (Cryptonemiales, Rhodophyta)." *Botanica Marina* **33**(3): 273-76.
- Zhang, Z., G. Yu, H. Guan, X. Zhao, Y. Du and X. Jiang (2004). "Preparation and structure elucidation of alginate oligosaccharides degraded by alginate lyase from *Vibrio* sp. 510." *Carbohydrate Research* **339**(8): 1475-81.
- Zobell, C. E. (1941). "Studies on marine bacteria. I-The cultural requirements of heterotrophic aerobes." *Journal of Marine Research* **4**(1): 42-75.
- Zobell, C. E. (1946). *Marine Microbiology: A monograph on Hydrobacteriology*, Chronica Botanica.
- Zwart, P. H., P. V. Afonine, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, E. McKee, N. W. Moriarty, R. J. Read, J. C. Sachettini, N. K. Sauter, L. C. Storoni, T. C. Terwilliger and P. D. Adams (2008). "Automated structure solution with the PHENIX suite." *Methods in Molecular Biology* **426**: 419-35.

## Résumé

La bactérie marine *Zobellia galactanivorans* dégrade un grand nombre de polysaccharides de macroalgues. Je me suis intéressé à la dégradation de l'alginate, le composé principal de la paroi des algues brunes, qui est largement utilisé en tant qu'agent gélifiant dans l'industrie. L'annotation du génome de *Z. galactanivorans* a confirmé l'énorme potentiel de cette bactérie pour la dégradation de polysaccharides. Notamment, *Z. galactanivorans* possède 7 alginate lyases putatives. Ces gènes sont organisés en deux clusters comprenant également d'autres acteurs hypothétiques de la dégradation de l'alginate (kinase et déshydrogénases, un régulateur transcriptionnel, un système SusCD-like, etc.). J'ai montré que ces clusters de gènes sont transcrits au sein d'ARN polycistroniques et constituent les premiers opérons alginolytiques décrits. Après l'évaluation de gènes de référence pour *Z. galactanivorans*, une approche transcriptomique ciblée a révélé que l'expression des gènes identifiés est fortement dépendante de la présence d'alginate. Des analyses en génomique comparative et phylogénie ont montré que cette voie métabolique est partiellement conservée chez d'autres espèces bactériennes, mettant en évidence de potentiels événements de transferts de gènes. La compréhension de l'utilisation de l'alginate a été élargie par un profilage transcriptomique à l'échelle du génome entier chez *Z. galactanivorans*. Il a permis d'identifier d'autres gènes répondant à la présence d'alginate et pouvant participer à la perception, la prise en charge ou la dégradation du substrat. J'ai entrepris une stratégie de clonage à moyen débit sur 21 gènes impliqués dans l'utilisation de l'alginate. Deux nouvelles alginate lyases PL7 de *Z. galactanivorans* ont ensuite été surexprimées chez *E. coli*, purifiées et cristallisées. Leur spécificité et mode d'action ont été étudiés, menant à l'identification d'AlyA1 comme endo-guluronate lyase et AlyA5 comme exo-guluronate lyase. Leurs structures 3D ont été résolues. Elles donnent des informations sur les déterminants des propriétés enzymatiques et le mode de reconnaissance du substrat. Tous les résultats ont été intégrés pour proposer un modèle de système alginolytique chez *Z. galactanivorans*.

## Abstract

The marine flavobacterium *Zobellia galactanivorans* is able to degrade a great variety of polysaccharides from macroalgae. I have focused on the degradation of alginate, the main cell wall polysaccharide from brown algae, which is widely used as gelling agent in industrial applications. The annotation of the genome of *Z. galactanivorans* has confirmed the huge capacity of this marine bacterium to degrade polysaccharides. Notably *Z. galactanivorans* possesses seven putative alginate lyases. These genes are organized into two clusters, together with other hypothetical actors of alginate degradation (sugar kinase and dehydrogenases, a transcription factor, a SusCD-like system, etc.). I have shown that these gene clusters are transcribed within polycistronic mRNAs and thus form the first described alginolytic operons. After having evaluated reference genes for *Z. galactanivorans*, a targeted transcriptomic approach revealed that the expression the identified genes is strongly dependent on the presence of alginate. Comparative genomics and phylogenetic analyses showed that this pathway is partly conserved in other bacterial species, unraveling potential gene transfer events. The understanding of alginate utilization was broadened by a whole-genome expression profiling in *Z. galactanivorans*. It allowed the identification of other alginate-responsive genes that might participate in substrate perception, uptake or degradation. I conducted a medium-throughput cloning strategy including 21 target genes involved in alginate utilization. Two novel PL7 alginate lyases from *Z. galactanivorans* were then overexpressed in *E. coli*, purified and crystallized. Their specificity and mode of action have been investigated, leading to the identification of AlyA1 to be an endo-guluronate lyase and AlyA5 an exo-guluronate lyase. Both 3D structures were solved and gave insight into the determinants of substrate recognition and enzymatic properties. All results were integrated to propose a model for the alginolytic system in *Z. galactanivorans*.