



HAL
open science

Recherche et caractérisation de nouvelles enzymes de bactéries marines actives sur les polysaccharides algaux par une approche biochimique et structurale

Maria Matard-Mann

► **To cite this version:**

Maria Matard-Mann. Recherche et caractérisation de nouvelles enzymes de bactéries marines actives sur les polysaccharides algaux par une approche biochimique et structurale. Biochimie [q-bio.BM]. Université Pierre et Marie Curie - Paris VI, 2017. Français. NNT : 2017PA066659 . tel-03465174

HAL Id: tel-03465174

<https://theses.hal.science/tel-03465174v1>

Submitted on 3 Dec 2021

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.

Université Pierre et Marie Curie

Ecole doctorale Complexité du Vivant – ED515

Station Biologique de Roscoff- UMR8227 LBI2M / Equipe Glycobiologie Marine

Recherche et caractérisation de nouvelles enzymes de bactéries marines actives sur les polysaccharides algaux *par une approche biochimique et structurale*

Par Maria Matard-Mann

Thèse de doctorat de Biologie

Dirigée par Mirjam Czjzek et Pi Nyvall Collén

Présentée et soutenue publiquement le 07/12/2017

Devant un jury composé de :

Dr. Magali Remaud-Siméon, PU, Rapporteur

Dr. Charles Tellier, PU, Rapporteur

Dr. Sulzenbacher Gerlind, IR, Examinatrice

Dr. Jean-François Hubert, PU, Examineur

Dr. Christophe Destombe, PU, Président du jury

Dr. Pi Nyvall Collén, Responsable R&D, Co-directrice de thèse

Dr. Mirjam Czjzek, DR, Directrice de thèse



*A celle qui a rendu possible cette aventure, celle qui m'a appris à déployer mes ailes, avant
de s'envoler elle-même vers un autre ciel*

Remerciements

Je tiens à remercier en premier lieu Bernard Kloareg de m'avoir permis de réaliser cette thèse au sein de la Station Biologique de Roscoff, ainsi que Catherine Boyen pour m'avoir accueilli au sein de l'UMR8227-LBI2M.

Je tiens également à remercier tout spécialement le groupe Olmix et son PDG Hervé Balusson pour avoir accepté de soutenir cette thèse dans le cadre d'une bourse CIFRE, ainsi qu'à l'ANRT qui m'a fait confiance en m'accordant cette bourse. Le caractère fondamental de la majorité de mon travail montre bien l'importance de la recherche en amont de tout projet de développement industriel, et je remercie Monsieur Balusson de l'avoir compris. Je remercie également toute l'équipe R&D de Bréhan pour son accueil lors de mes passages par leurs locaux, en particulier Christelle, Marie, Laura, Damien, Clément et Tony qui m'ont permis de me familiariser aux différents aspects de la recherche en industrie. Mais également Amélie pour son aide précieuse pour toutes les démarches administratives. Et enfin, un peu plus que trois fois MERCI à Pi pour son accompagnement, sa patience et sa compréhension pendant ces trois années de collaboration. J'ai vraiment apprécié ton état d'esprit constructif et motivant, quelle que soit la situation.

Retour à la SBR à présent...ça ne va pas être facile de citer tous ceux qui ont croisé ma route en ... cinq ans (dont trois de thèse) ! Je commencerai par celle qui m'a permis de venir à Roscoff, qui m'a accueilli, moi et ma passion débordante pour les algues, qui a su me comprendre et me conseiller, dans mes moments de doutes comme dans mes prises de décisions importantes. Pour tout cela et tout le reste, MERCI Mirjam !

Un grand merci à toute l'équipe GM : Sabine, notre maman-gâteaux, Murielle, la fée du labo *et alex*, la fée des cristaux, Gurvan et Tristan, les bons génies de la forêt numérique peuplée d'arbres phylogénétiques aux branches buissonnantes, Robert, notre Québécois national, capable de venir à bout d'armées entières de clones en buvant une petite gorgée de sirop d'érable, Liz, dont le rire et la paillasse nous donnent toujours le sourire, François, le gentleman-zen du labo 131, Lionel, notre SuperPCman à qui aucun bug ne résiste, Cécile, dont je n'oublierai jamais le talent d'actrice insoupçonné, Kévin, alias Monsieur Bricolage, Aurélie, alias Dionex-Girl, Amandine, alias Pupuce-girl, Anaïs et John, le duo de choc de la cristallo,

l'équipe Algolife : Bobby et ses drôles de dames : Diane, Aline et Bertille, et aussi Sophie et Marie.

Merci aussi à notre « Bureau de filles » : Laure, Amandine, Miriam : notre armoire à thé et à chocolat va me manquer ! Bon courage à mes deux collègues de promo Amandine et Simon, ainsi qu'à la fournée suivante : Laure, Hervé, Adèle, Dung, Miriam, Hetty, Anaïs et à celle d'après : Josselin et Yao.

Merci aussi à ceux qui sont passés par le labo et qui ont compté pour moi : JB, Patty, Flo, Céline, Marie-Mathilde, Laurent, Adrian, Harold, ...

Un immense merci à ma famille et à mes amis pour leur amour et leur soutien: mes parents ; Jacqueline et Pierre, Jess et Sabrina ; Norbert ; Michelle ; Brigitte ; Pascal, Sandrine et Yannick ; Jean *et alice* ; Françoise et Sean ; Pierre, Sonia, Annie et André ; Robert et la chorale du bonheur ; Aurélie et Dung.

MERCI à tous !

*« Raconte-moi la mer
Dis-moi le gout des algues
Et le bleu et le vert
Qui dansent sur les vagues »
« Court le furet file la vague
Je vivrai dans un monde d'algues »
« Algues brunes ou rouges
Dessous la vague bougent
Les goémons »
« Je suis une algue au fil de l'eau
Mais parfois monte la colère
Pouvoir un jour braver le flot
Ne pas toujours se laisser faire »
« La mer
Au ciel d'été confond
Ses blancs moutons
Avec les anges si purs
La mer bergère d'azur
Infinie [...]
Et d'une chanson d'amour
La mer
À bercé mon cœur pour la vie »*

Jean Ferrat, *Raconte-moi la mer* ; Dalida, *Darladirladada* ;
Serges Gainsbourg, *Les Goémons* ; Alain Barrière, *L'algue* ; Charles Trénet, *La Mer*

Sommaire

Remerciements	3
Sommaire	5
Introduction	10
I. Des Hommes et des Algues : historique, état des lieux et perspectives	12
1. Usages des macroalgues: entre tradition et innovation	12
1.1 Usage alimentaire	14
1.2 Usage médicinale : thérapeutique et nutraceutique	18
1.2.1 <i>L'iode</i>	19
1.2.2 <i>Les acides gras</i>	19
1.2.3 <i>Les pigments et les composés phénoliques</i>	20
1.2.4 <i>Les protéines, peptides, acides aminés bioactifs</i>	20
1.2.5 <i>Les polysaccharides</i>	21
1.3 Les algues comme matière première	21
2. Comment pérenniser le succès ?	24
2.1 La ressource en quelques chiffres	24
2.2 La valorisation industrielle: verrous et perspectives	27
II. Les polysaccharides algaux : des molécules complexes aux propriétés multiples	31
1. Rappels et définitions de glycochimie	31
2. Organisation et diversité des polysaccharides algaux	33
2.1 Localisation et spécificités des polysaccharides algaux	33
2.2 Une complexité à plusieurs niveaux	35
2.3 Le cas des carraghénanes	36
3. Nature et propriétés des polysaccharides algaux	38
3.1 Les polysaccharides d'algues vertes	40
3.2 Les polysaccharides d'algues rouges	41
3.2.1 <i>L'agar</i>	41
3.2.2 <i>Les carraghénanes</i>	42
3.3 Chez les algues brunes	43
3.3.1 <i>Les alginates</i>	43
3.3.2 <i>Les fucoïdanes</i>	44
3.3.3 <i>La laminarine</i>	44
III. Les enzymes de dégradation des glucides	46
1. Définition et classification	46
1.1 Généralités	46
1.2 La classification CAZy	47
1.3 Les glycosides hydrolases	48
2. Les CAZymes de bactéries marines	52
2.1 Rôles écologiques	53
2.2 Stratégies de dégradation : les PULs marins	55
2.3 Etat de l'art sur les GH de bactéries marines actives sur polysaccharides algaux	58
3. Applications des CAZymes en biotechnologie bleue	68
3.1 Utilisations en recherche fondamentale	68
3.2 Utilisations en recherche appliquée	69
IV. Problématique scientifique	71
Chapitre 1 : L'approche à moyen débit permet d'optimiser la découverte de nouvelles enzymes de bactéries marines impliquées dans la dégradation de polysaccharides algaux.	76
I. Introduction	76
II. Analyse <i>in silico</i>	76

1. Analyse des séquences brutes	77
2. Recherche de la stratégie de clonage.....	81
3. Conception des amorces pour l'amplification des gènes	81
III. Clonage et tests d'expression	82
1. Préparation des ADN génomiques	82
2. Amplification PCR des inserts	82
3. Préparation du plasmide pFO4 et clonage.....	83
4. Tests d'expression à petite échelle	85
5. Contrôle qualité par séquençage	89
IV. Purification, cristallogénèse et cristallographie aux rayons X	89
1. Production et purification des protéines recombinantes	90
2. Cristallogénèse, diffraction des rayons X et résolution structurale.....	92
V. Détermination des substrats des glycosides hydrolases	94
VI. Caractérisation biochimique des glycosides hydrolases.....	97
1. Stabilité thermique	98
2. Température optimale de l'enzyme.....	99
3. pH optimal.....	99
4. Les paramètres cinétiques	99
Chapitre 2 : La comparaison de nouvelles structures cristallines de kappa-carraghénases révèle des éléments clés de cette sous-famille des GH16	104
Structural insights into marine carbohydrate degradation by family GH16 kappa- carrageenases.....	106
Chapitre 3 : Deux enzymes de bactéries marines spécifiques du β -carraghénane révèlent l'existence d'une nouvelle famille de glycoside hydrolase apparentée aux GH42.....	148
I. Introduction générale sur le β -carraghénane et sa dégradation	148
II. Modèle de dégradation conjointe des ι -, κ - et β -carraghénanes par <i>P.</i> <i>carrageenovora</i>	149
III. Conclusion.....	154
Two marine bacterial enzymes specific of β -carrageenan allow the discovery of a new family of glycoside hydrolases, distantly related to GH42.	157
Chapitre 4 : Caractérisation d'une α -L-fucosidase de la famille GH29 au sein d'un PUL putatif de <i>Zobellia galactanivorans</i>	175
I. Introduction	175
II. Données et hypothèses de départ.....	177
1. Un PUL putatif fucanolytique	177
2. Éléments contradictoires	178
III. Résultats expérimentaux.....	180
1. Constructions clonées.....	180
2. Tests d'activité	181
2.1 Sur substrats synthétiques.....	181
2.2 Sur substrats naturels	184
2.2.1 Polysaccharides algaux	184
2.2.2 Oligosaccharides fucosylés issus du lait.....	188
3. Caractérisation biochimique de la GH29	189
4. Initiation de l'étude structurale	191
IV. Discussion-Conclusion du chapitre	192
1. La 'taille' d'un PUL peut être trompeuse et les substrats complexes	192
2. L'importance des sulfatases	193
3. La technique de détection est cruciale pour le criblage de substrats naturels	193
Conclusion générale et perspectives	195

I. La stratégie moyen débit	196
II. La caractérisation d'enzymes de substrat inconnu	196
III. Les voies de dégradation des carraghénanes	197
1. Les κ -carraghénases de <i>Z. galactanivorans</i> et de <i>P. carrageenovora</i>	197
2. Les GH42-like et la dégradation des carraghénanes par <i>P. carrageenovora</i>	199
IV. Perspectives	200
Annexes	204
Bibliographie I	226
Bibliographie II	245

Introduction générale



Figure 1 : Itinéraire prévu pour l'expédition Tara Pacifique (2016-2018). La fondation Tara Expédition est une initiative française née en 2003 grâce aux créateurs des marques agnès b. et Bourgois. Elle organise des expéditions scientifiques à bord de la goélette Tara dans le but d'étudier et de comprendre l'impact des changements climatiques et de la crise écologique sur les océans. Elle réalise également des campagnes de sensibilisation auprès du grand public et depuis 2015 a reçu le statut d'Observateur spécial aux Nations unies, lui permettant ainsi de participer aux différentes commissions et conférences de l'ONU.

Introduction

Cette thèse a été réalisée dans le cadre d'une bourse CIFRE (Conventions Industrielles de Formation par la Recherche) avec l'entreprise Amadéite du groupe Olmix. Malgré le caractère assez fondamental de ce travail, qui se situe bien en amont dans la chaîne aboutissant à la commercialisation d'un produit, je tiens à souligner l'importance de sa prise en compte par une société comme Olmix qui cherche à valoriser la ressource algale. En effet, l'innovation et les produits qui feront le succès de demain se préparent dès aujourd'hui en allant découvrir les richesses que les ressources marines peuvent nous offrir. Cependant, bien que la mer représente 70 % de la surface du globe, elle reste à ce jour largement méconnue. Et si les grandes campagnes océanographiques comme Tara Pacifique (Figure 1) permettent d'apporter de la visibilité au travail des équipes de recherche tout en augmentant les connaissances du monde marin, l'écart entre la connaissance des écosystèmes terrestres et marins est encore bien grand. Comme nous le verrons au cours de cette introduction, la ressource algale représente une ressource immense en termes de diversité des applications industrielles possibles. Cependant, son coût élevé et son accès limité comparativement aux biomasses végétales terrestres doivent amener à penser différemment les usages qui peuvent en être fait. Ainsi, l'orientation vers des applications à haute valeur ajoutée, nécessitant des volumes de matière première plus faibles, représente la voie la plus soutenable aussi bien d'un point de vue économique qu'écologique. Cependant, pour conquérir de nouveaux marchés en santé, en nutrition ou en cosmétique, les industriels doivent se montrer innovants, mais également apporter des preuves scientifiques aux allégations commerciales mises en avant pour leurs produits. Dans cette optique, les polysaccharides algaux représentent une formidable opportunité. Composés les plus abondants en termes de biomasse algale, ils possèdent des champs d'application multiples et leurs activités biologiques ont fait l'objet de nombreuses études. Cependant, leur diversité et leur complexité structurale rendent leur étude particulièrement ardue. Les enzymes de bactéries marines qui vivent au contact des algues représentent à ce titre un enjeu majeur dans une perspective d'innovation marine. En effet, outils d'une précision incroyable, elles permettent aussi bien d'identifier la structure fine des polysaccharides algaux, que de faciliter leur exploitation industrielle. Ce travail de thèse s'inscrit donc dans cette perspective, et vise à développer les connaissances actuelles sur les enzymes bactériennes marines, mais également à en découvrir de nouvelles, pour, demain, permettre leur utilisation dans la valorisation de la biomasse algale.



Figure 2 : Exemples d'usages du littoral par l'Homme. A, Port de plaisance ; B, Tourisme balnéaire; C, Pêche à pieds ; D, Chargement d'un navire marchand ; E, Récolte de laminaires par un bateau goémonier ; F, Déchet sur une plage ; G, Course sportive « La Barjot » dans le port de Roscoff ; H, Paquebot de la Brittany Ferries dans le port en eau profonde de Roscoff ; I, Groupe d'étudiant en excursion ; J, Caseyser ; K, Construction en front de mer ; L, Thalassothérapie

I. Des Hommes et des Algues : historique, état des lieux et perspectives

La proximité des zones côtières semble avoir été un facteur clé dans le développement et le fonctionnement de la masse cérébrale d'*Homo sapiens*. En effet, les acides gras polyinsaturés et les micronutriments, tels que l'iode, sont indispensables au bon fonctionnement de notre cerveau (Crawford 1992; Cunnane and Crawford 2014). Les produits marins en sont bien plus riches que n'importe quelle source alimentaire terrestre, du fait de leur bioaccumulation dans les réseaux trophiques, des producteurs et fixateurs primaires que sont les algues (micro comme macroalgues), jusqu'aux animaux filtreurs et aux carnivores marins. Aujourd'hui encore, la moitié de la population mondiale vit à moins de 200 km des côtes, environ un tiers (27 %) à moins de 100 km (Kummu *et al.* 2016), et ces 27 % génèrent à eux seuls 42 % du PIB mondial. La multiplicité des activités humaines liées à la ressource marine (Figure 2) fait du littoral un espace privilégié mais aussi fragile, de par les interactions entre les populations qui partagent l'écosystème côtier, qu'elles soient humaines, animales ou végétales. C'est dans cette perspective d'interconnexion entre les Hommes et la Mer que j'aimerais aborder cette étude en commençant par retracer l'évolution des usages des macroalgues par l'Homme, depuis la préhistoire jusqu'à nos jours.

1. Usages des macroalgues: entre tradition et innovation

Le terme de macroalgue désigne un ensemble d'Eucaryotes pluricellulaires capables de réaliser la photosynthèse répartis dans trois phyla distincts : les *Phaeophyceae* (algues brunes), les *Rhodophyceae* (algues rouges) et les *Chlorophyceae* (algues vertes). D'un point de vue évolutif, les algues vertes ont divergé des algues rouges avant de donner naissance aux plantes terrestres. Elles appartiennent donc au même groupe des Archaeplastida. En revanche, les algues brunes sont apparues indépendamment, suite à une endosymbiose secondaire d'une algue rouge unicellulaire par l'ancêtre commun du groupe SAR (Stramenopiles-Alveolate-Rhizaria). D'un point de vue évolutif, les algues brunes sont donc aussi éloignées du phylum des algues rouges et vertes (Archaeplastida) que de celui des animaux (Opisthokontes) (Figure 3). Environ 3000 espèces de macroalgues sont répertoriées au travers du globe, majoritairement en milieu marin. Mais certaines espèces poussent dans des eaux saumâtres voire même en eau douce, témoignant du cheminement évolutif qui donna naissance aux plantes terrestres.

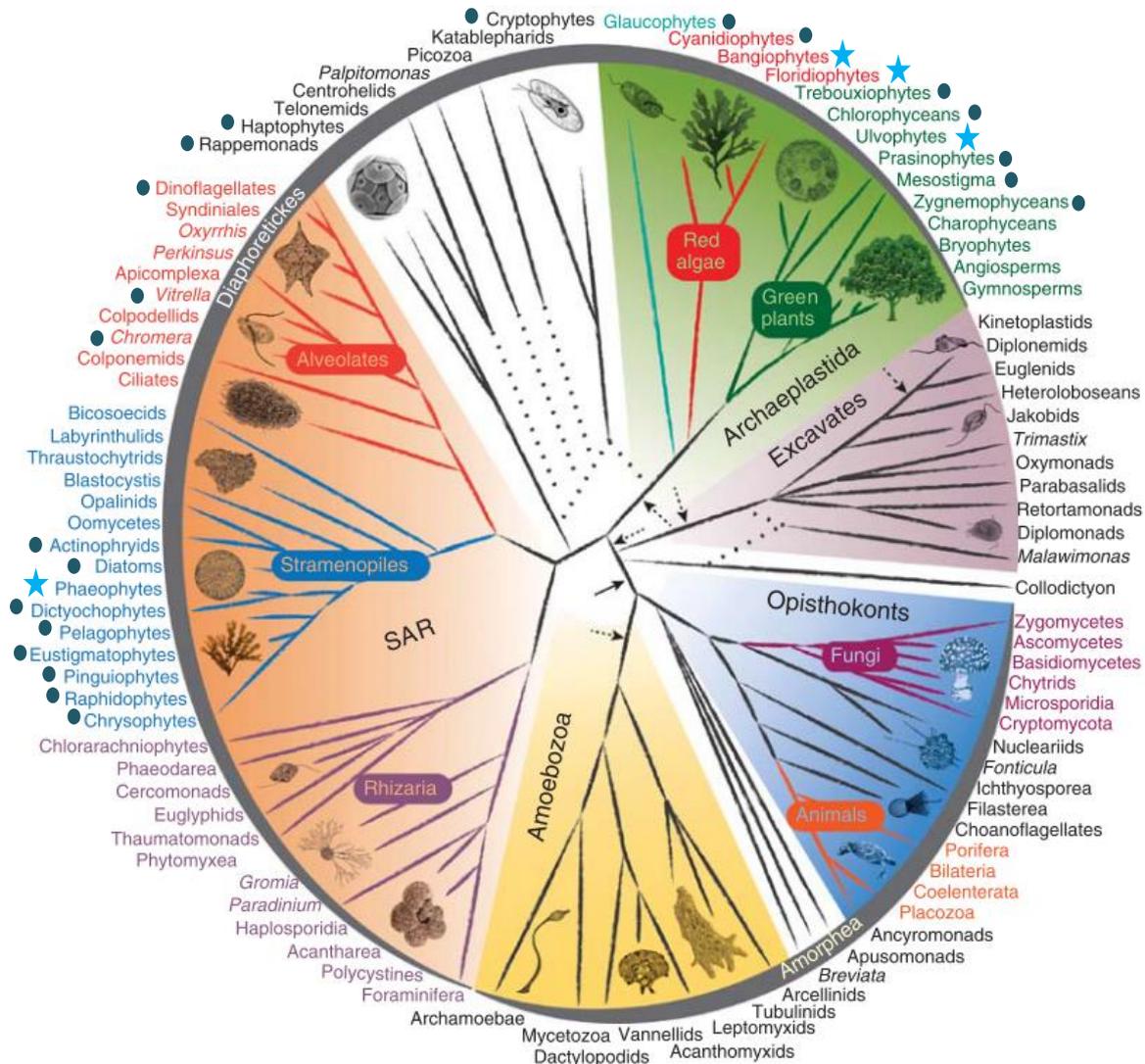


Figure 3 : Arbre phylogénétique des Eucaryotes (Burki *et al.* 2014). Les classes contenant majoritairement des macroalgues sont indiquées par une étoile, celles contenant majoritairement des microalgues par un rond.

Les principales distinctions morphobiologiques entre macroalgues et plantes terrestres sont l'absence de racines et de tissus conducteurs de sèves chez les algues, ainsi qu'une grande diversité de cycles reproductifs, ne faisant jamais intervenir de fleurs ou de graines.

Le mot « algue » vient du latin *alga*, mais des synonymes se retrouvent dans le parler populaire tel que varech sur le littoral normand (issu du scandinave médiéval *vágrek*) ou goémon sur le littoral breton (*gwymon* en gallois et en corrique). Le terme anglais *kelp* viendrait quant à lui de l'anglais médiéval *culp*. Ces terminologies vernaculaires désignent aussi bien les algues vivantes, accrochées à leur substrat, que les algues d'échouage, arrachées par les courants marins et les tempêtes puis déposées par la marée sur les plages. Les types d'usage que les hommes ont fait des macroalgues peuvent s'organiser en trois catégories: Alimentaire, Médicinale et Technologique.

1.1 Usage alimentaire

1.1.1 Alimentation humaine

Les traces les plus anciennes de consommation alimentaire de macroalgues par des populations humaines ont été retrouvées sur le site de Monte Verde, au Chili et datent d'au moins 14 000 ans (Dillehay *et al.* 2008). Les traditions alimentaires de nombreuses populations côtières tout autour du globe font remonter à plusieurs millénaires ce type d'usage. Plus de 3000 ans pour le « *Limu Moui* » du royaume des Tonga (Philpott and Bradford 2006) ou encore à plus de 5000 ans en Irlande (Mouritsen *et al.* 2013). En Amérique du Nord, l'usage ancestral par les amérindiens de Colombie britannique est aussi bien connu (Hansen *et al.* 1981; Turner 2003). La diversité des espèces utilisées peut être impressionnante, comme à Hawaï et en Polynésie en général, où des jardins sous-marins entretenaient la culture de plus de 70 espèces différentes de *limu* (Philpott and Bradford 2006; Hansen *et al.* 1981).

Les premières traces écrites d'un usage alimentaire sont retrouvées dans le recueil chinois de poèmes « Book of Poetry » daté entre -800 et -600 ans avant notre ère (Wood 1974). Dans le plus ancien dictionnaire sino-japonais, édité sur ordre de l'empereur Daigo au X^{ème} siècle, 21 espèces d'algues marines sont recensées comme aliments, avec des recettes associées (Nisizawa *et al.* 1987), dont certaines sont encore d'actualité.

En Europe, ce sont des textes législatifs irlandais du VIII^{ème} siècle ainsi que la saga islandaise d'Egil, datant du X^{ème} siècle, qui en attestent l'usage alimentaire parmi les populations scandinaves (Mouritsen *et al.* 2013). Ce sont principalement la dulse (*Palmaria palmata*), la nori (*Porphyra umbilicalis*) et le pioka (*Chondrus crispus*) qui semblent avoir été consommées traditionnellement en Europe (Figure 4G,F,D). Les Vikings emportaient vraisemblablement de la dulse séchée lors de leurs voyages en mer, car sa richesse en vitamine C leur permettait d'éviter le scorbut. La recette du « *laver bred* » Gallois à base de nori est encore très appréciée aujourd'hui (Hansen *et al.* 1981). Mais contrairement à l'Asie, en Europe, les algues sont souvent associées à des périodes de famines, comme celle que connut l'Irlande entre 1845 et 1852. En témoigne ce dicton des îles Féroé, qui dit d'une personne qui meurt dans la pauvreté : « Elle a été enterrée avec une algue dans la bouche ». Au temps des Romains, les algues véhiculaient déjà une image négative comme le montre la phrase du poète Virgile « *Nihil vilior alga* » (*litt.* « Il n'y a rien qui ait moins de valeur qu'une algue »). Cette association malheureuse fait partie des verrous socio-culturels dont il faut tenir compte dans une optique de développement de la filière des macroalgues en Europe.

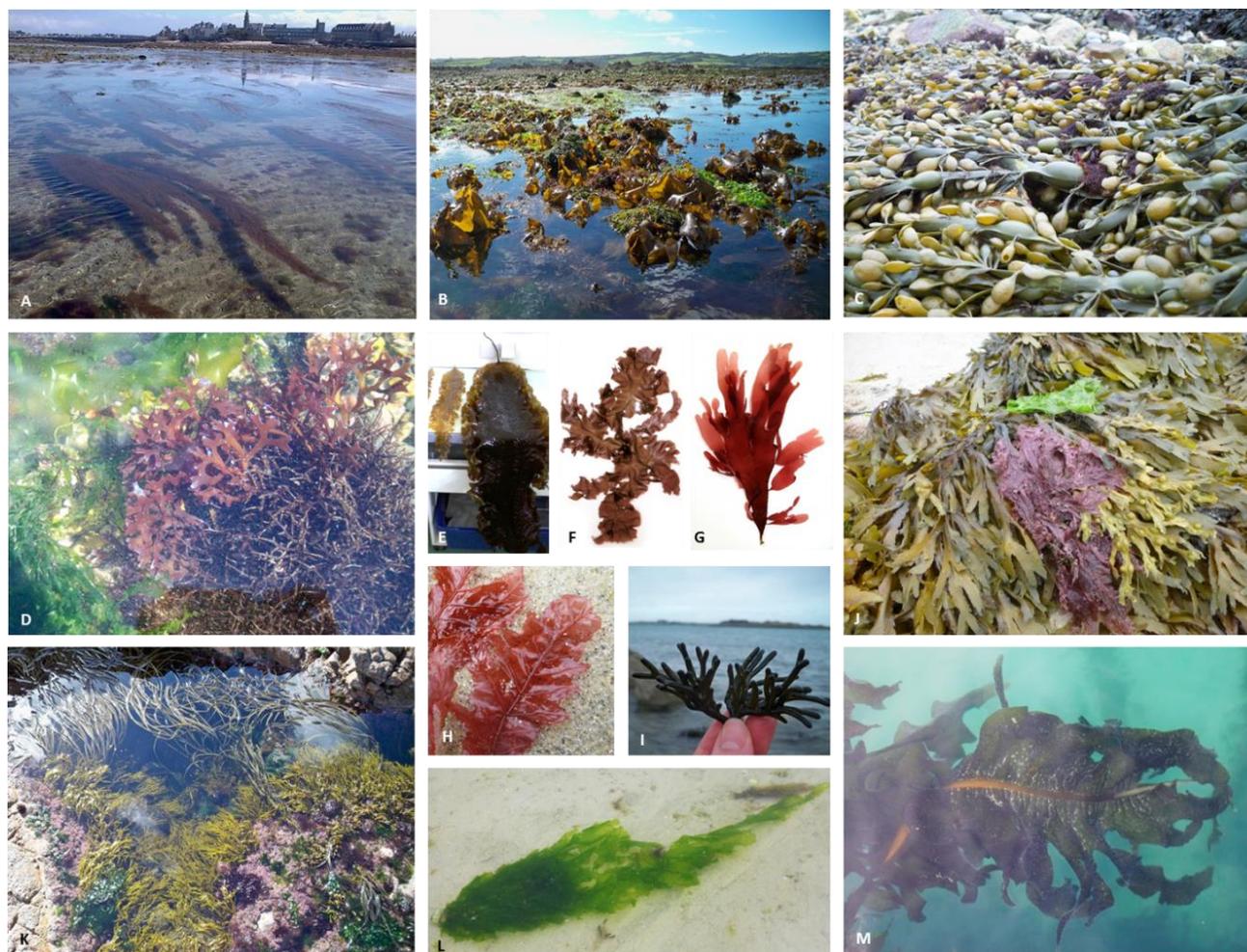


Figure 4: Exemples d'algues couramment rencontrées sur le littoral français de la Manche. A, *Sargassum muticum* ; B, *Laminaria digitata* ; C, *Ascophyllum nodosum* avec l'algue rouge *Polysiphonia lanosa* en épiphyte; D, *Chondrus crispus* (milieu), *Chondracanthus acicularis* (droite) et *Ulva intestinalis* (gauche); E, *Saccharina latissima* ; F, *Porphyra* sp. ; G, *Palmaria palmata*; H, *Delesseria sanguinea* ; I, *Codium* sp. ; J, *Fucus serratus* (à gauche) et *F. vesiculosus* (à droite) ; K, *Himanthalia elongata* (haut), *Dictyota dichotoma* (milieu) et l'algue rouge calcaire *Corallina officinalis* (bords) ; L, *Ulva* sp. ; M, *Undaria pinnatifida*

Cependant, l'influence du courant culturel nippon en Occident, accompagné de sa cuisine à base d'algues (soupes miso, maki-sushi, ...), contribue à l'évolution positive du regard européen sur ces végétaux marins. De plus, leurs qualités nutritionnelles en font des produits très sains d'un point de vue diététique, renforçant leurs attraits gustatifs qui ne sont pas toujours aussi perceptibles aux palais occidentaux qu'à ceux des orientaux.

1.1.2 Alimentation animale

Outre l'alimentation humaine, les algues ont aussi servi à alimenter le bétail depuis des temps très reculés en particulier en Norvège, en Islande et aux îles Féroé (Mouritsen and Richard 2015). Le terme breton de *bezhin-saout* (litt. goémon à vaches) employé par les habitants de l'île de Batz pour désigner la dulse (*Palmaria palmata*), atteste de son usage alimentaire pour le bétail par les populations locales. Cet usage est également attesté chez les Grecques en -100 av JC (Philpott and Bradford 2006) ainsi que chez les Romains (Craigie 2011). En France, les premiers essais contrôlés d'incorporation d'algues aux rations alimentaires des animaux datent de 1917, donnant naissance aux premières entreprises d'exploitation des algues pour l'alimentation animale (Wuitner 1921). Aujourd'hui encore, des entreprises principalement françaises, norvégiennes et canadiennes développent des compléments alimentaires pour les animaux d'élevage en mettant en avant les propriétés immunostimulantes ou prébiotiques des algues telles qu'*Ascophyllum nodosum* ou *Ulva* sp (Figure 4C, L) (Frangoudes 2011; Craigie 2011; Evans and Critchley 2014).

1.1.3 Les hydrocolloïdes algaux

En plus de leurs qualités nutritionnelles et gustatives, un grand nombre d'algues possèdent des propriétés texturantes intéressantes dans divers domaines, allant de la cuisine à la microbiologie, en passant par l'industrie pharmaceutique. L'utilisation traditionnelle de ces propriétés en cuisine se retrouve en Asie, où l'extraction de l'agar fut décrite en Chine en 1658 (Mouritsen and Richard 2015), mais également en Europe, avec le *pioka* breton (*Chondrus crispus*), aussi appelé carrageen ou *Irish moss*, servant de gélifiant dans les recettes traditionnelles, et dont l'importation au nouveau monde par les colons britanniques s'est poursuivie jusqu'en 1835 (Wood 1974). Il faudra pourtant attendre les années 1950 pour voir une véritable industrie des hydrocolloïdes algaux (ou phycocolloïdes) se développer à une échelle mondiale. Les hydrocolloïdes sont des agents de texture incorporés aux produits de grande consommation. De nature protéique (ex. : gélatine) ou polysaccharidique (ex. : gomme de guar), leurs propriétés hydrosolubles leur permettent d'interagir avec les molécules d'eau

du produit, modifiant ainsi ses propriétés rhéologiques. Chez les algues, les hydrocolloïdes sont de nature polysaccharidique et regroupent l'agar-agar, les carraghénanes et l'alginate (Bixler and Porse 2010; Rhein-Knudsen *et al.* 2015; Cunha and Grenha 2016). Leur production mondiale annuelle s'élevait à 100 000 tonnes en 2014, avec un marché d'environ 1,1 milliards de dollars. Ils sont reconnus comme additifs alimentaires et au niveau de la réglementation européenne et possèdent donc un numéro Exxx :

- L'agar-agar (E406) est extrait des algues rouges agarophytes, appartenant principalement aux genres *Gracilaria* et *Gelidium*. La culture de *Gracilaria* est maîtrisée au Chili et en Asie (Chine, Indonésie, Vietnam) ce qui explique son essor spectaculaire depuis les années 1990, qui est passé de 55 % à 80 % en part de la production mondiale d'agar entre 1999 et 2009. Quant au *Gelidium*, il est récolté à l'état sauvage principalement sur les côtes Ouest du Maroc, Sud du Portugal et Nord de l'Espagne. Il est utilisé comme agent gélifiant (glaces, confitures) et épaississant (crèmes).

- Les carraghénanes (E407) sont extraits des algues rouges carraghénophytes, principalement *Kappaphycus alvarezii* ("Eucheuma cottonii" sous son nom commercial) pour l'obtention de carraghénane de type kappa, et *Eucheuma denticulatum* ("Eucheuma spinosum") pour le carraghénane de type iota. La culture de ces deux espèces se développant en eaux chaudes est pratiquée principalement aux Philippines et en Indonésie. *Gigartina skottsbergii* et *Chondrus crispus* sont les deux sources de carraghénane de type lambda, dans leur stade sporophytique, mais leur culture n'étant pas maîtrisée, leur exploitation repose sur une ressource naturelle fragile. Les carraghénanes sont utilisés comme stabilisants, agents gélifiant et épaississant en industrie agro-alimentaire, dans la boulangerie et l'alimentation des animaux de compagnie.

- Les alginates (E400-405) sont extraits d'algues brunes, principalement de l'ordre des *Laminariales* et des *Fucales* (Figure 4B, J), issues de ressources sauvages à l'exception de *Laminaria japonica*. Historiquement, c'est l'Europe qui était le plus gros producteur mondial, suivie des Etats-Unis et du Japon, grâce à ses nombreux champs de laminaires exploitables de façon mécanisée par les pêcheurs goémoniers, principalement en Norvège et en France. Mais depuis les années 1990, la Chine s'est mise à produire des alginates à bas prix grâce à la culture massive de *L. japonica* faisant passer l'Asie au rang de leader mondial. L'alginate est utilisé comme stabilisant et épaississant en agro-alimentaire.

Cependant, les faibles marges réalisées par les producteurs d'hydrocolloïdes combinées aux versatilités des marchés et de l'accès à la ressource, poussent certains d'entre eux à se reconverter dans l'extraction et la vente de produits issus d'algues à plus haute valeur ajoutée,

tels que les compléments alimentaires pour animaux d'élevage ou les produits phytosanitaires (Evans and Critchley 2014).

1.2 Usage médicinale : thérapeutique et nutraceutique

Les premières traces d'usage médicinale des macroalgues se confondent avec l'usage alimentaire sur le site de Monte Verde (Dillehay *et al.* 2008). Les plus anciens témoignages d'utilisation médicinale des algues viennent d'Asie. L'empereur chinois légendaire Shen Nung (env. 3000 ans av. J.C.) aurait rapporté dans son *Materia medica* les propriétés émoullientes, décontractantes musculaires et anti-hématomes des algues. Il les recommandait également pour le traitement de la constipation, des maladies endocriniennes, des cystites et des bronchites chroniques (Mouritsen and Richard 2015). Un autre traité de médecine chinoise (-300 av. J.C.) recommande l'usage d'extraits d'algues comme répulsif pour insectes et les propriétés vermifuges de l'algue rouge *Digenea* sont connues depuis plus de 1000 ans (Stein and Borden 1984). En Occident, c'est en 1200 qu'un médecin de l'école de Salerne note l'intérêt d'utiliser des cendres d'algues pour traiter le goitre, et au XV^{ème} siècle les danois utilisaient le fucus vésiculeux pour traiter les maux de tête. En Ecosse, il était dit « qui mange de la dulce de Guerdie et boit au puits de Kildingie échappera à toute maladie excepté à la peste ». Et en Irlande, la préparation traditionnelle d'infusions de *Chondrus crispus* contre la toux et les rhumes est encore pratiquée (Major 1977; Morrissey *et al.* 2001).

Dans le registre de la « santé végétale », l'usage d'algues est attesté par les écrits de l'agronome romain Columelle au premier siècle (Craigie 2011), mais la pratique en est bien antérieure chez les populations celtes et nordiques. La première initiative commerciale visant à produire des extraits d'algues comme fertilisants, date de la fin de la seconde guerre mondiale. En Grande Bretagne, l'entreprise Maxicrop fut la première à mettre au point et à commercialiser un extrait à base de laminaires récoltées sur les côtes Ecossoises. L'entreprise compléta par la suite son activité en développant des compléments alimentaires à base d'algues pour le bétail (Craigie 2011).

Ces usages traditionnels ont été confirmés et expliqués par les avancées de la science depuis le XIX^{ème} siècle. En effet, la composition unique des algues en font des alliées de choix dans la lutte contre les pathologies animales comme végétales. La liste des effets biologiques démontrés pour les différentes familles de composés issus des algues est résumée dans le tableau en annexe 6. De plus, avec l'émergence des « aliments fonctionnels » dans les années 1980, l'alimentation est devenue un moyen de combattre et de prévenir les maladies humaines

et animales, donnant naissance aux concepts de nutraceutique ou d'aliment (Kalra 2003). S'il faut prendre garde aux allégations mensongères qui profitent du faible encadrement législatif relatif à cette nouvelle tendance (Cardenas 2013), la nutraceutique a cependant le mérite de rappeler que l'alimentation est un facteur essentiel de notre santé, tout comme l'avait déjà souligné Hippocrate, cinq siècles avant notre ère. Dans ce contexte, les algues apparaissent de plus en plus comme des « ingrédients santé » de premier ordre (Holdt and Kraan 2011).

1.2.1 L'iode

Il fut découvert par inadvertance en 1811 par le chimiste français Bernard Courtois, alors qu'il cherchait un moyen de fabriquer du salpêtre à partir de cendres d'algues pour le compte des guerres napoléoniennes. Il commença à être extrait des algues brunes pour ses propriétés thérapeutiques dès les années 1830, même si son rôle exact dans le fonctionnement de la glande thyroïde ne fut découvert qu'entre 1925-1930 par Sir Charles Harington. Depuis, il a été montré que la carence alimentaire en iode entraîne l'apparition de déficiences thyroïdiennes, et peut également provoquer des retards mentaux et de croissance, observés le plus souvent chez des populations vivant loin des côtes ou en régions montagneuses (cf. l'origine de l'expression « crétin des Alpes »). Actuellement, la gestion de la déficience en iode est prise en compte à l'échelle mondiale par la mise en place de politiques de supplémentation alimentaire, en particulier *via* du sel iodé, comme en Suisse ou aux Etats-Unis. Cependant, si les années 1920 ont vu l'apogée de l'extraction d'iode à partir d'algues en Europe, depuis la fin de la seconde guerre mondiale l'iode n'est plus extrait que du caliche chilien, une roche nitrée qui en contient en abondance, ainsi que de certaines eaux souterraines qui en sont enrichies (Kelly 1961).

1.2.2 Les acides gras

Bien que la teneur en lipides des algues soit relativement faible (1 à 5 % de la matière sèche), celles-ci sont majoritairement constituées d'acides gras polyinsaturés (AGPI) à longue chaîne de type oméga 3 (ω -3) et oméga 6 (ω -6), avec un ratio ω -6/ ω -3 relativement faible, à la différence des végétaux terrestres. Ces acides gras sont dits essentiels, car les mammifères sont incapables de les synthétiser et doivent donc les acquérir par la nourriture. Or dans l'alimentation américaine et européenne contemporaine, ce ratio ω -6/ ω -3 est de 50, alors qu'il est de 12 chez les japonais et de 1 chez les esquimaux du Groenland. Cet excès d' ω -6 a été montré comme étant corrélé à la recrudescence des troubles cardiovasculaires dans les pays

occidentaux. En effet, les ω -3 jouent un rôle essentiel dans la composition membranaire des cellules cardiaques. Mais ils sont également impliqués dans le contrôle du ratio entre les protéines de transport du cholestérol, diminuant la proportion de LDL (*Low Density Lipoprotein*) associé à la formation de la plaque d'athérome, au profit du HDL (*High Density Lipoprotein*). Une alimentation enrichie en algues possédant un faible ratio ω -6/ ω -3, comme la dulse où il est à 0,13 ou le wakamé à 0,49 contribuerait donc à prévenir les problèmes cardiovasculaires, ainsi que les problèmes de surpoids qui y sont souvent liés (Mišurcová 2011; Miyashita *et al.* 2011).

1.2.3 Les pigments et les composés phénoliques

Les pigments photosynthétiques sont constitués des chlorophylles, des caroténoïdes (fucoxanthine chez les algues brunes, xanthophylles, α et β -carotène, lutéine,...) et des phycobilliprotéines (phycoérythrine, phycocyanine chez les algues rouges) qui sont constituées d'une apoprotéine et d'un chromophore. De par leur puissant effet antioxydant, ils ont été montrés comme participant à la prévention des maladies cardio-vasculaires, de cancers, de maladies ophtalmiques, ainsi que dans les maladies chroniques impliquant des phénomènes de stress oxydatif (Holdt and Kraan 2011).

Les composés phénoliques sont synthétisés et accumulés majoritairement chez les algues brunes, pouvant représenter jusqu'à 20 % de la matière sèche chez les *Fucales* et 30 % chez les *Dictyotales*. C'est une classe de molécules regroupant les phloroglucinols, les phlorotannins, qui sont des dérivés de haut poids moléculaire du phloroglucinol, ainsi que les composés phénoliques monomériques halogénés, retrouvés occasionnellement chez les algues brunes ainsi que chez quelques algues rouges (ex. : l'algue poivrée, *Osmundea pinnatifida*, dont elle tire sa saveur caractéristique). Tout comme les pigments, c'est leur pouvoir antioxydant qui les rend très intéressants dans la prévention des pathologies liées aux phénomènes d'oxydation. Ils a également été démontré récemment qu'ils possèdent des propriétés antibactériennes et antidiabétiques (Freitas *et al.* 2015).

1.2.4 Les protéines, peptides, acides aminés bioactifs

La teneur protéique des algues rouges, et dans une moindre mesure celle des algues vertes, est relativement élevée, en moyenne 10 à 30 % du poids sec de l'algue (soit le double des algues brunes), avec des valeurs pouvant atteindre 35 à 47 % chez la dulse ou la nori (Mišurcová 2011).

Parmi les protéines, fragments et dérivés peptidiques, certains possèdent des activités biologiques intéressantes. Ainsi, les phycobilliprotéines, évoquées précédemment, présentent de nombreuses propriétés allant des anti-inflammatoires aux antiviraux, neuroprotecteurs et hypocholestérolémiants. Les MAAs (*Mycosporine-like Amino-Acids*) également, possèdent des effets protecteurs contre les UV, et enfin, des peptides endogènes analogues au GABA, sont à l'étude dans le cadre de troubles neurologiques. A noter toutefois qu'en plus de la variabilité spécifique, il existe une variabilité saisonnière importante dans la teneur et la nature des protéines synthétisées dans les macroalgues. De plus leur accès est parfois difficile du fait de l'abondance en polysaccharides. Les stratégies de digestion enzymatique des polysaccharides permettent de multiplier jusqu'à 10 fois les rendements d'extraction (Harnedy and Fitzgerald 2011; Nam 2011; Jeon *et al.* 2011).

1.2.5 Les polysaccharides

La complexité des polysaccharides algaux et les propriétés biophysiques qui leurs sont associées, en font des actifs majeurs possédant de nombreuses applications dans les domaines de la santé humaine, animale et végétale. Elles s'étendent des domaines de la réparation tissulaire aux traitements anticancéreux, en passant par les maladies cardio-vasculaires ou les infections sexuellement transmissibles, et la stimulation immunitaire. Ces applications sont principalement liées aux propriétés hydrocolloïdes des polysaccharides, à la taille des polymères, mais également à leur caractère sulfaté, et à leur composition osidique (Seisun *et al.* 2010; Holdt and Kraan 2011; Rhein-Knudsen *et al.* 2015; Kraan 2012; Cunha and Grenha 2016). Nous reviendrons en détail au paragraphe II de cette introduction sur la diversité structurale des polysaccharides algaux.

1.3 Les algues comme matière première

1.3.1 Entre usages traditionnels et biomatériaux de nouvelle génération

Pour finir, les algues représentent également une matière première aux propriétés multiples et originales. Les habitants des régions côtières tout autour du globe ont utilisé très tôt cette source de fibres végétales pour confectionner divers artefacts (Figure 5). Ainsi, les aborigènes de Tasmanie se servaient des lames de laminaire pour fabriquer des réservoirs servant à transporter l'eau, et jusqu'au XIX^{ème} siècle les populations de Polynésie fabriquaient encore des embarcations légères en utilisant les vésicules aérifères de la laminaire géante *Nereocystis* comme flotteurs (Hornell 1946). Les descendants des Maori en Nouvelle Zélande



Figure 5 : Divers usages des algues comme matière première à travers le temps et l'espace. A, Système pour transporter l'eau réalisé en lame de laminaire par les aborigènes de Tasmanie ; B, C, *Nereocystis* ou laminaire nerf-de-bœuf (« bull-kelp »), aux vésicules aérifères lui permettant de flotter aisément malgré sa grande taille ; D, Confection de pōhā, sacs faits en lame de *Nereocystis* pour préserver la viande de Puffin fuligineux, au sud du détroit de Foveaux en Nouvelle Zélande-Photo prise en 1910-Crédits: Hocken Library, University of Otago ; E, Boîtes de Pétri contenant un milieu de culture bactérien gélifié à l'aire d'agar-agar ; F, Publicité pour des lunettes dont la monture a été réalisée à base d'algues par l'entreprise Algopak, Saint-Malo, France.

avaient réussi à fabriquer des sacs en dédoublant les lames de cette même algue afin de conserver la viande des oiseaux de mer qu'ils chassaient. En Europe, la première exploitation industrielle des algues remonte au XVII^{ème} siècle, où le bicarbonate de sodium (communément appelé « soude ») issu des cendres d'algues brunes fut employé dans la fabrication du verre, afin d'abaisser la température de fusion de la silice (Arzel 1994). Actuellement, des entreprises innovantes, telle qu'Algopak à Saint-Malo en France, ont développé des procédés permettant d'utiliser les fibres d'algues brunes pour créer de nouveaux matériaux de type bioplastiques.

1.3.2 Combustibles et biocarburants

Des résidus d'algues brûlées retrouvés sur certains sites archéologiques laissent à penser qu'elles ont pu être utilisées comme combustible par des populations côtières préhistoriques. Cet usage a d'ailleurs perduré en Bretagne jusqu'au début du XX^{ème} siècle (Arzel 1994). Avec la menace de pénurie des carburants fossiles, la recherche de nouvelles sources de biocarburants a éveillé l'intérêt de nombreuses entreprises pour les algues. Cependant, la teneur en lipides des macroalgues étant très faibles, ce sont vers des espèces de microalgues que les recherches se sont orientées pour la production d'huile végétale carburant. En revanche, la teneur importante en sucres des macroalgues, ainsi que leur taux de croissance élevé, comparé à celui plantes terrestres, représentent en théorie un potentiel non négligeable de biomasse

exploitable dans le cadre d'une filière bioéthanol. Cependant, les études de concept basées sur des projets de culture à grande échelle de laminaires ou d'algues vertes, ont montré que la production de bioéthanol ne serait pas rentable à elle seule face aux coûts de production. Pour que le système soit viable, il faut impérativement mettre en place une voie de valorisation des coproduits vers des compléments alimentaires pour animaux à plus haute valeur ajoutée que le bioéthanol. Des facteurs clés tels que la masse critique à traiter, la variabilité saisonnière de la production, un abaissement des coûts d'exploitation - par rapport aux pratiques actuelles des petites exploitations - ou encore les fluctuations des valeurs du marché des coproduits sont autant de paramètres critiques à prendre en compte car ils complexifient la mise en œuvre à court terme d'une filière bioéthanol basée sur la culture de macroalgues (Philippsen *et al.* 2014; Korzen *et al.* 2015).

1.3.3 Les phycocolloïdes non alimentaires

Enfin, les phycocolloïdes, déjà abordés dans les sections précédentes, ont trouvé des applications très diverses dans de nombreux secteurs industriels autres que l'agro-alimentaire. Une des plus anciennes est l'usage de l'agar comme agent solidifiant des milieux de culture bactériens. Pour la petite histoire, ce serait la femme de Robert Koch qui lui aurait suggéré cette idée, utilisant elle-même l'agar pour réaliser des préparations alimentaires en gelée. Son usage pour solidifier les milieux de culture bactériens permit à Koch de découvrir le bacille de la tuberculose, popularisant ainsi cette technique en microbiologie. Toujours dans le domaine de la recherche, l'agar est également employé pour réaliser des gels d'électrophorèse pour l'ADN, ou des résines de chromatographie. Il sert aussi de matrice pour le clonage et la propagation de plantes, ainsi que de support pour la nourriture d'insectes d'élevage, comme la *Drosophile* ou le vers à soie. L'alginate possède aussi des applications dans les secteurs de la cosmétique et de la pharmaceutique comme adjuvant. Concernant les carraghénanes, ils sont utilisés comme agents liant dans le dentifrice, comme épaississant et stabilisant en cosmétique et comme agent d'encapsulation en pharmaceutique pour contrôler le relargage de molécules actives. Enfin, les alginates extraits des algues brunes représentent la plus grosse biomasse d'hydrocolloïde employée à des fins non alimentaires. En effet, près de 40 % de la production mondiale part dans l'industrie textile comme agent de fixation des pigments. Ils sont aussi utilisés dans l'industrie papetière, dans les moulages dentaires, le garnissage des couches pour bébés et comme agent d'encapsulation. Leur usage en industrie pharmaceutique est également fréquent comme pansements et matrice d'encapsulation pour des médicaments ou des cellules (Seisun *et al.* 2010; Kraan 2012).

Au travers de ces multiples usages se dessinent la richesse et les potentialités de la biomasse algale. Cependant, pour que la croissance de ce marché soit pérenne, plusieurs facteurs clés, inhérents aux spécificités de cette ressource, doivent être pris en compte par les différents acteurs de la filière : exploitants, industriels, pouvoirs publics, scientifiques et consommateurs.

2. Comment pérenniser le succès ?

2.1 La ressource en quelques chiffres

La FAO (*Food and Agriculture Organization of the United Nations*) fournit la principale source de données chiffrées sur la récolte et la vente d'algues à l'échelle mondiale. Depuis 1995, elle publie tous les deux ans un rapport sur « La situation mondiale des pêches et de l'aquaculture ». Elle fournit également une base de données des statistiques mondiales disponibles depuis 1950, permettant de réaliser des représentations graphiques de l'évolution de la production mondiale des ressources aquatiques, en sélectionnant les différents paramètres de suivi désirés (Figures 6, 7, 8)¹.

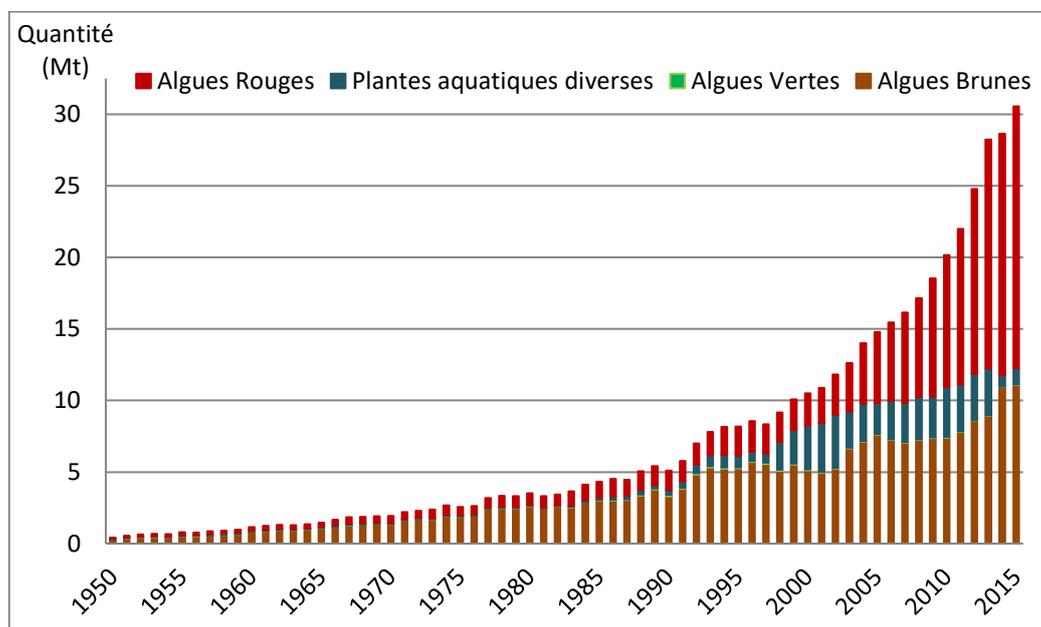


Figure 6 : Evolution de la production mondiale de plantes aquatiques

¹ Afin de construire les figures 6, 7 et 8, les valeurs de récolte de macroalgues disponibles sur le site de la FAO ont fait l'objet de modifications pour les valeurs françaises à l'aide des données trouvées sur le site de l'insee.

L'augmentation exponentielle de la production mondiale de plantes aquatiques depuis les années 50 est principalement liée aux algues rouges, la production des algues brunes suivant une augmentation linéaire et celle des algues vertes restant à un niveau extrêmement faible (Figure 6). Cette augmentation de la production d'algues brunes et rouges est principalement imputable à l'augmentation de la culture des macroalgues en Asie. En effet, même si la récolte en milieu naturel a presque triplée depuis 1950 (Figure 7), elle représente moins de 4 % de la production mondiale actuelle, soit seulement 1,1 des 30,5 Mt de la production mondiale en 2015 (données FAO). Ces données sont à corrélérer avec le caractère limité du stock naturel. En effet, si la ressource est exploitée au maximum sans tenir compte des contraintes de son cycle naturel, un épuisement progressif du stock est inévitable. C'est ce qui s'est passé pour l'algue carraghénophyte *Gigartina skottsbergii* au Chili. Cette espèce est parmi les rares qui produisent du λ -carraghénane, mais sa culture n'étant pas maîtrisée, la population naturelle exploitée de façon intensive est en train de disparaître des côtes chiliennes.

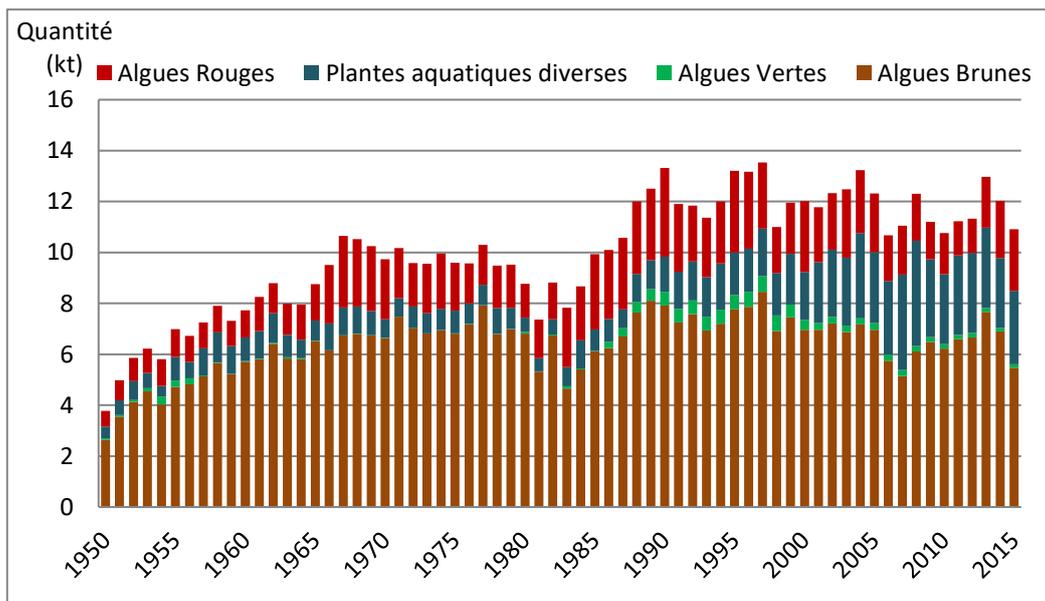


Figure 7 : Evolution de la récolte en milieu naturel à l'échelle mondiale

Pour la ressource en algues brunes sauvages, cette problématique a été prise en compte très tôt par les pêcheurs goémoniers en Europe avec la mise en place de systèmes de quotas et de périodes de récolte. Un système de jachère de 3 ans entre deux collectes sur une même zone a également été mis en place pour la récolte de *Laminaria hyperborea*. L'appui de scientifiques, de l'Ifremer en particulier, a permis d'évaluer et de valider l'efficacité de ce système de gestion de la ressource naturelle, assurant ainsi une pérennité de la biomasse algale sauvage

(Frangoudes and Garineaud 2015). Cependant, avec l'augmentation constante de la demande au niveau du marché mondial, la récolte en milieu naturel ne permet clairement pas de pourvoir aux besoins croissants. Et ce sont les espèces dont la culture est maîtrisée qui permettent d'assurer l'augmentation de la production.

Le marché mondial est dominé par l'Asie, qui représente à elle seule plus de 97 % de la production mondiale avec 29,6 Mt (Figure 8). L'Europe arrive en 3^{ème} position derrière les Amériques avec une production annuelle de 228 kt, soit 0,75 % de la production mondiale. Contrairement à la tendance mondiale, la production européenne est majoritairement basée sur de la récolte en milieu naturel, avec seulement 1,2 % de la production issue de l'aquaculture (données FAO pour l'année 2015). A l'échelle européenne, la France est le deuxième pays producteur (17,5 % sur la période 2004-2014) avec en moyenne 75 000 tonnes par an, derrière la Norvège (58 %) et suivie de l'Irlande (11,4 %) et de l'Islande (7,7 %).

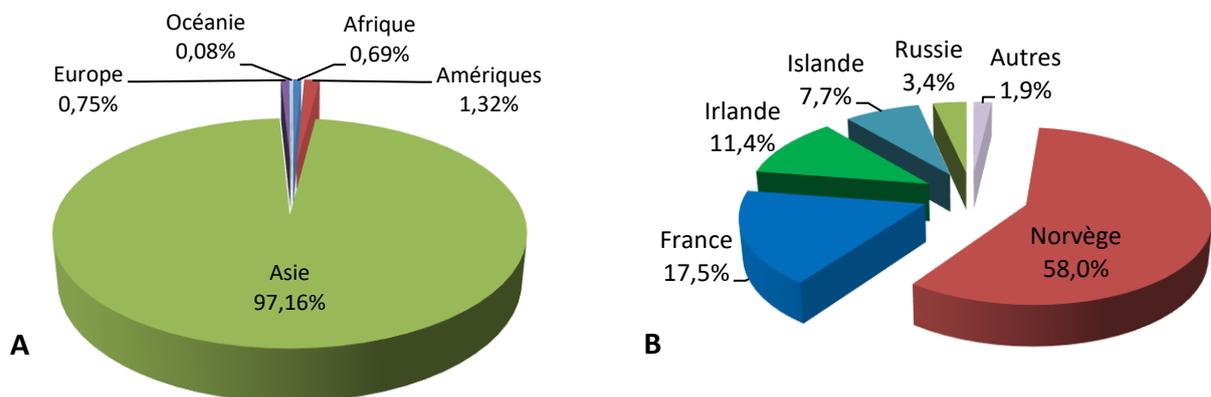


Figure 8 : Répartition de la production de plantes aquatiques A, Mondiale par continent en 2015 (FAO) ; B, Européenne sur la décennie 2004-2014 (FAO et insee)

Les principales espèces exploitées en France² en termes de tonnage sont les laminaires *Laminaria digitata* (41548 t) et *L. hyperborea* (23600 t) récoltées mécaniquement par les bateaux goémoniers au moyen de scoubidoues et de peignes norvégiens respectivement. Les volumes de récolte manuelle des algues de rive sont estimés entre 6000 et 10000 t, répartis entre quatre espèces principales : *Fucus* sp. : 44 %, *Ascophyllum nodosum* : 34 %, *Chondrus crispus* : 8 %, *Palmaria palmata* : 7%. Les 7 % restant incluent entre autres *Himanthalia elongata*, *Porphyra umbilicalis* et *Ulva lactuca*. Les volumes produits par l'algoculture restent encore faibles, entre 50 et 5000 t. A ces volumes, vient s'ajouter la récolte des algues d'échouage qui

² Les données chiffrées pour la France sont issues d'une présentation de Philippe Potin de 2017, compilant des données de 2015 venant de l'Ifremer, du CRPMEM (comité régional des pêches maritimes et des élevages marins) et du SPRARB (syndicat professionnel des récoltants d'algues de rive de Bretagne).

est estimée à plusieurs dizaines de milliers de tonnes annuelles (Hardouin *et al.* 2014). Elle peut se faire directement dans le rideau d'eau ou sur le sable, et ce sont principalement les algues vertes de la famille des *Ulva* sp. (entre 80 et 100 000 t/an récoltées en Bretagne et en Normandie) et les algues rouges *Solieria chordalis* (environ 15 000 t/an récoltées en Vendée) et *Gelidium sesquipedale* (récoltée sur la côte basque) qui sont concernées.

2.2 La valorisation industrielle : verrous et perspectives

Les contraintes industrielles propres à chaque type d'activité seront évidemment différentes entre un producteur de phycocolloïdes ou un producteur de compléments alimentaires pour animaux. Cependant, des contraintes et des besoins communs existent en lien avec la nature de la ressource algale exploitée. L'objectif de ce paragraphe est donc de regrouper les verrous communs à tout type d'activité industrielle et de discuter des mesures nécessaires pour envisager une exploitation durable de la ressource algale dans le contexte Français (Figure 9).

D'un point de vue mondial, l'algoculture représente un enjeu majeur pour empêcher l'épuisement de la ressource naturelle. La comparaison des systèmes européens et asiatiques est assez éloquent, puisque 98,8 % de la production européenne provient de la récolte de ressources sauvages, alors que celle-ci ne représente que 1,6 % de la production asiatique. Ce déséquilibre n'est pas sans conséquences, puisque la production européenne ne suffit pas à remplir les besoins des industriels. A titre d'exemple, la France importe presque deux fois plus d'algues qu'elle n'en produit (130 000 tonnes pour 75 000 tonnes respectivement) (synthèse de résultats 2015, projet Idealg³). Le développement de l'algoculture, surtout dans une optique d'aquaculture multitrophique intégrée, représente donc un objectif crucial pour assurer la pérennité de la ressource (Figure 10). Mais un travail de recherche important est nécessaire afin de maîtriser les cycles reproductifs des espèces poussant dans nos eaux, sélectionner des lignées variétales intéressantes, tout en prenant en compte les questions de brassage génétique des populations. Il faut aussi adapter les techniques de cultures aux conditions plus rigoureuses des côtes européennes comparativement à l'Asie.

Les questions sociétales, telles que les conflits d'usage pour l'espace maritime, mais

³ Idealg est un projet investissement d'avenir porté par l'Université Européenne de Bretagne (UEB) et coordonné par la Station Biologique de Roscoff visant à développer la filière des macroalgues en France. Financé à hauteur de 10 millions d'euros sur la période 2010-2020 par l'Agence National de la Recherche (ANR), il associe des organismes de recherche publics, un centre technique des algues et cinq entreprises spécialisées dans l'aquaculture et la valorisation des algues. <http://www.idealg.ueb.eu/>

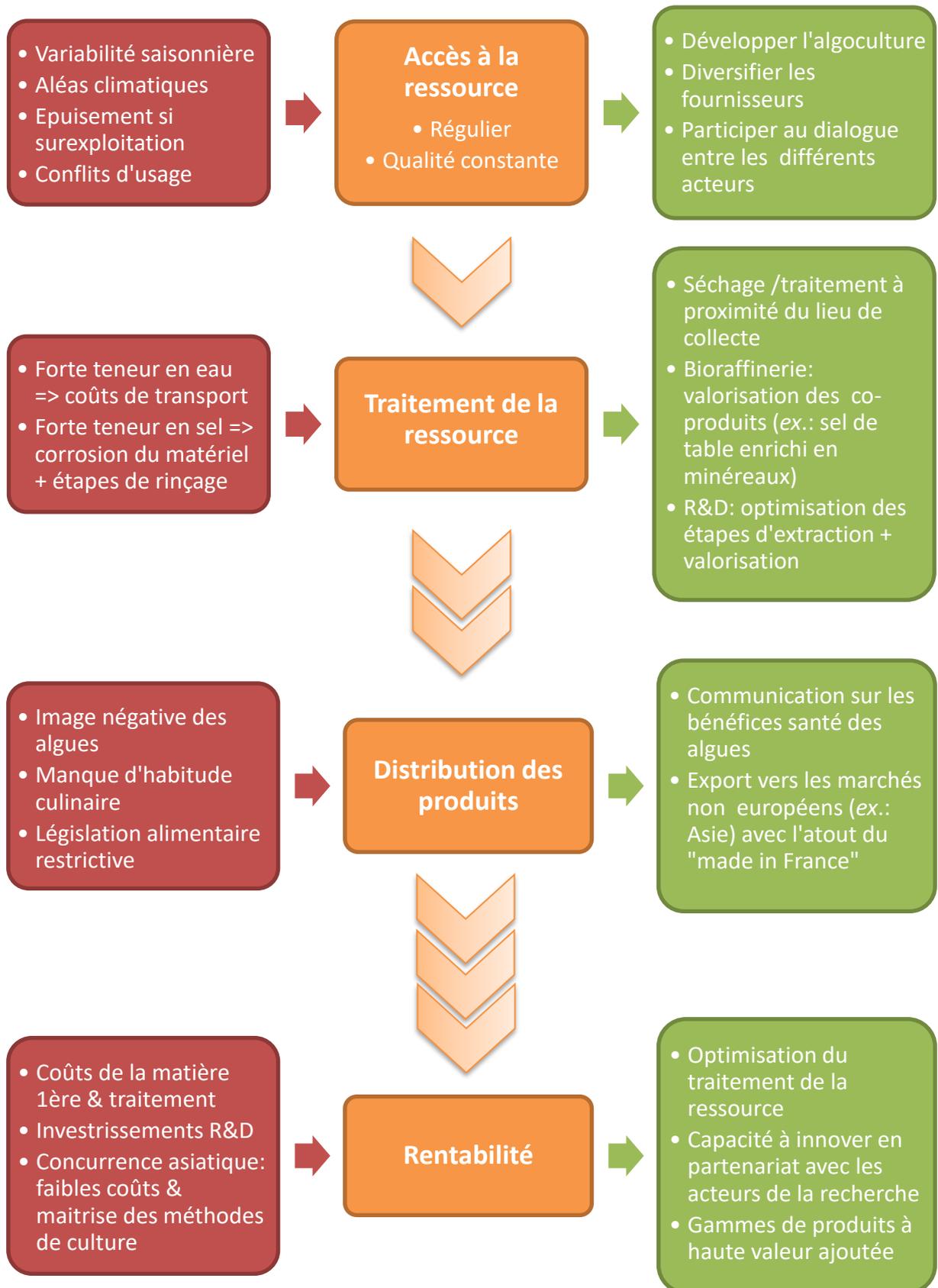


Figure 9 : Représentation schématique des verrous (en rouge) et perspectives (en vert) aux différents besoins industriels (en orange) dans le cadre de l'exploitation des macroalgues.

également l'acceptation sociale de nouvelles pratiques alimentaires associées à la consommation des algues, sont autant de verrous qui restent encore à débloquent. Cependant, des projets intégratifs, tels que le programme Idealg, qui vise à développer la filière macroalgue en Bretagne, ont déjà permis de dresser un état des lieux de la situation et des différents acteurs. Les nombreux projets qui voient le jour dans le cadre de la politique de croissance bleue, à l'échelle nationale comme européenne, devraient également amener à multiplier les échanges entre les différents acteurs du domaine maritime.

Un autre verrou à l'exploitation de la ressource algale est le coût de revient élevé de la matière première (récolte, teneur en eau, volumes de stockage,...). Cette problématique peut être résolue de différentes façons par les industries : en compensant les faibles marges par des gros volumes de production (*ex.* : industrie des phycocolloïdes), mais aussi en cherchant à diversifier la gamme de produits en développant des produits à plus forte valeur ajoutée (*ex.* : gamme nutraceutique humaine), et en ciblant des marchés tels que l'Asie, où l'estampillage « made in France » est considéré comme un gage de qualité, en particulier dans les secteurs de la cosmétique ou de la gastronomie. Pour cela, un investissement en recherche et développement est essentiel, aussi bien en interne que *via* des collaborations avec différents établissements de recherche. Les mesures incitatives de l'Etat français (crédit impôt-recherche, ANRT finançant des bourses de thèse CIFRE public-privé,...) vont d'ailleurs dans ce sens. Un autre élément important à prendre en compte pour contrebalancer les coûts élevés de la matière première réside dans l'élaboration d'une stratégie de valorisation des co-produits tout au long de la chaîne de production. Ainsi, dans le contexte macroalgue, la forte teneur en sel représente souvent un problème pour le traitement de la biomasse. La commercialisation des sels minéraux issus des eaux de rinçage des algues peut ainsi représenter une voie originale de valorisation, tout en facilitant les étapes ultérieures du processus industriel (Magnusson *et al.* 2016).

Enfin, un dernier verrou réside dans la législation. En effet, seules 21 espèces de macroalgues sont autorisées à la consommation humaine en Europe, et l'absence de normalisation à l'échelle mondiale complique encore les choses dès lors que l'exportation à l'international est envisagée, en particulier vers l'Asie. De même, l'usage d'organismes génétiquement modifiés étant extrêmement encadré, le recours à des enzymes produites de façon hétérologues lors du processus de valorisation, nécessite leur expression dans des organismes GRAS (*Generally Recognized as Safe*) comme la levure *Pichia pasteuris*. Enfin, les allégations santé pour les produits destinés à la nutrition animale nécessitent la mise en place d'études souvent lourdes afin de démontrer l'efficacité de la supplémentation sur les animaux.

Pour conclure cette première partie retraçant l'utilisation des macroalgues par l'Homme, j'aimerais insister sur la richesse incroyable que représentent ces espèces marines, et sur leurs potentiels en particulier pour la santé humaine. Certes, il reste encore bien du travail à accomplir pour comprendre le fonctionnement de ces organismes complexes et de leurs interactions avec le milieu, de même que pour extraire au mieux les principes actifs qu'ils contiennent. Mais dans une perspective de domestication, tâchons de tirer les fruits des erreurs commises avec l'agriculture terrestre pour ne pas les reproduire avec l'algoculture. L'apparition des premières maladies au sein des grandes fermes marines asiatiques doit servir de signal d'alarme (Gachon *et al.* 2010). Les méfaits de la monoculture intensive pour l'équilibre des écosystèmes et la vitalité des espèces terrestres sont bien connus (Tschardtke *et al.* 2005; Cheatham *et al.* 2009; Wolfe 2000), et il faudra les prendre en compte dans une optique d'algoculture durable. C'est seulement dans des conditions respectueuses de l'environnement, de la ressource, ainsi que des différents acteurs, qu'une économie bâtie sur la biomasse algale sera viable sur le long terme.

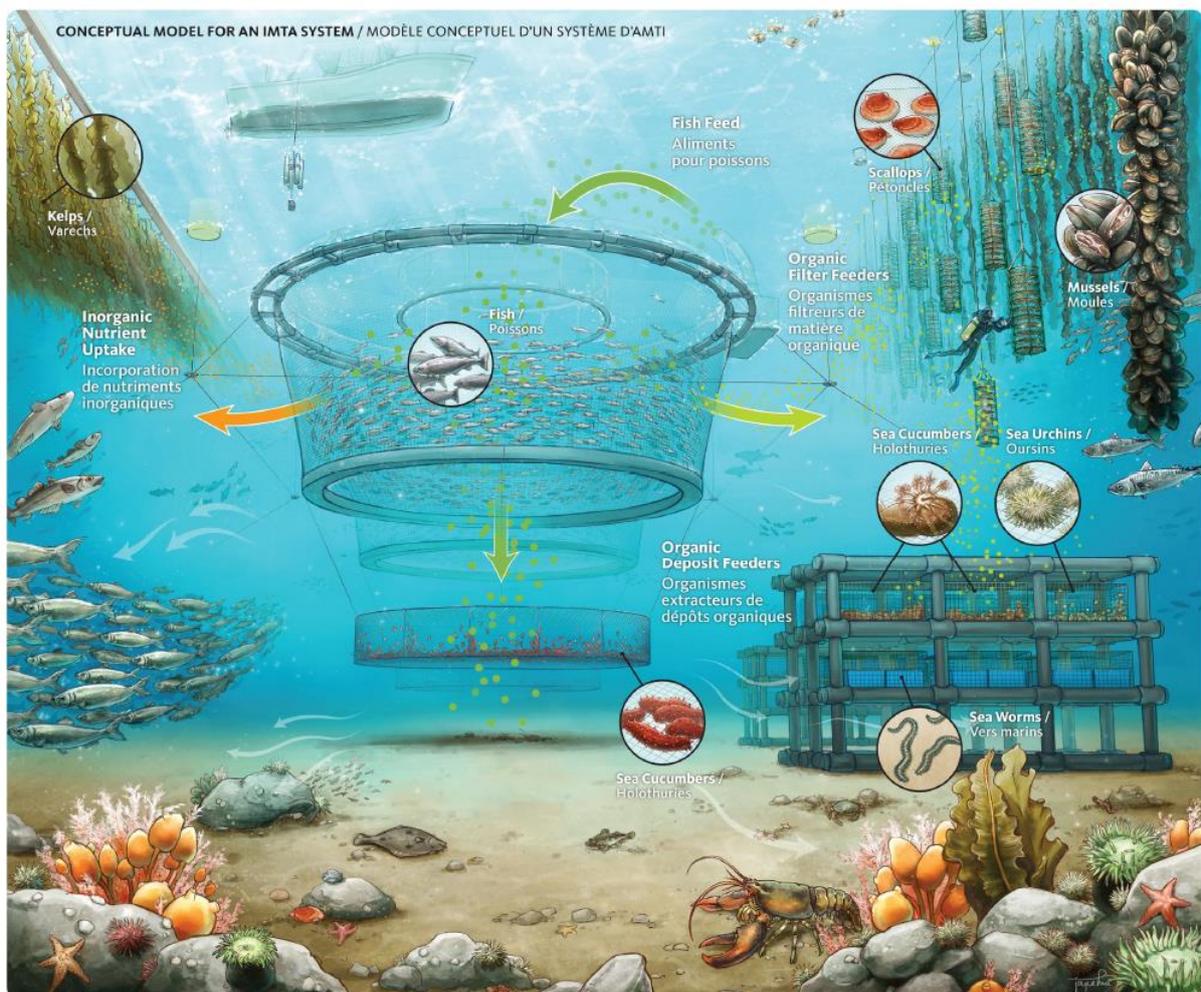


Figure 10 : Modèle conceptuel d'un système d'aquaculture multitrophique intégré (Pêches et Océans Canada 2013)

II. Les polysaccharides algaux : des molécules complexes aux propriétés multiples

Après avoir retracé les grandes lignes l’histoire et des usages de la ressource algale, nous allons à présent nous intéresser plus spécifiquement à une des familles de composés algaux que sont les polysaccharides.

1. Rappels et définitions de glycochimie

Le terme polysaccharide est en réalité un anglicisme, le terme français correct étant polyoside ou polyholoside, mais l’usage en a consacré l’utilisation. Il s’agit de macromolécules obtenues par la polymérisation de plus de dix unités glucidiques simples ou oses. D’un point de vue biochimique, un ose est une molécule de formule brute $C_nH_{2n}O_n$, constituée d’une chaîne de carbones non ramifiée, portant chacun un groupement hydroxyles sauf un qui porte une fonction **carbonyle**. Si cette fonction carbonyle est un aldéhyde, les sucres dérivés appartiendront à la série des **aldoses** (ex. : glycéraldéhyde, glucose, galactose) et si c’est une cétone, à la série des **cétooses** (ex. : dihydroxyacétone, xylulose, fructose). C’est cette fonction carbonyle qui confère aux oses leur pouvoir réducteur. En termes de classification, les oses se distinguent par : 1) leur nombre de carbones, qui va de 3 à 7 dans la plupart des matières naturelles, 2) la nature et la position de leur fonction réductrice, à savoir en C_1 pour les aldoses et en C_2 pour les cétooses, et 3) les positions des différents groupements hydroxyles, générant pour chaque ose à x carbones asymétriques ($x C^*$) 2^x **stéréoisomères** théoriques. La moitié des stéréoisomères sont dans la configuration D, la plus abondante dans la nature, et l’autre moitié dans la configuration L. Les configurations D et L sont **énantiomères** l’une de l’autre et

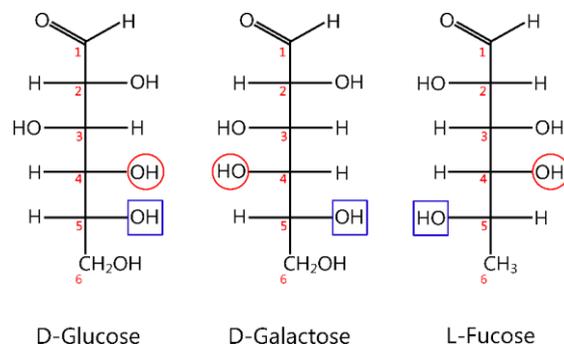


Figure 11 : Représentation de Fischer de trois hexoses. En rouge sont entourées les hydroxyles qui diffèrent entre le galactose et le glucose, et en bleu ceux qui distinguent les configurations D et L des hexoses.

correspondent à des propriétés opposées de déviation du plan de polarisation de la lumière polarisée. Seul le fructose possède l'énantiomère L sous forme majoritaire à l'état naturel.

Des oses ne différant que par la configuration d'un C* sont appelés des **épimères** (ex. : les D-galactose et D-glucose sont épimères en C₄) (Figure 11). En solution, les différentes formes **tautomériques** des oses sont en équilibre dynamique entre des formes cycliques, qui sont les plus stables d'un point de vue thermodynamique, et linéaires, qui sont des états transitoires instables. La cyclisation des oses résulte de la condensation entre la fonction carbonyle, très réactive, et une fonction alcool secondaire de la chaîne carbonée. L'**hémiacétalisation** intramoléculaire génère un hétérocycle de forme **pyranique** s'il a six sommets (forme la plus stable pour les aldoses) et **furanique** s'il en a cinq (forme la plus stable pour les cétooses). Le carbone portant la fonction carbonyle devenant asymétrique lors de la cyclisation, les deux cycles possibles pour une même structure linéaire sont appelés anomères α et β (Figure 12).

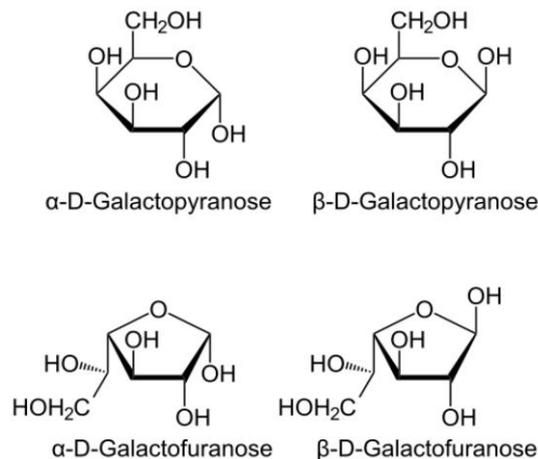


Figure 12 : Représentation de Haworth des quatre tautomères du galactose

La polymérisation de deux oses, par condensation entre l'hydroxyle du carbone anomérique d'un ose et un groupement hydroxyle d'un autre ose, aboutit à la formation d'une **liaison glycosidique**. Si le groupement hydroxyle du second est celui de la fonction hémiacétalique, alors le diholoside formé n'aura pas de pouvoir réducteur (*ex.* : le saccharose). La polymérisation d'un seul type d'ose génère des **polyholosides** ou homoglycanes, tels que l'amylose ou la cellulose, homopolymères de D-glucopyranose liés respectivement en α - et β -1,4. La polymérisation de différents types d'oses génère des **polyhétérosides** ou hétéroglycanes, tels que les alginates, hétéropolymères d'acide D-mannuronique et L-gulonurique liés respectivement en β - et α -1,4, ou les carraghénanes, hétéropolymères de galactose plus ou moins sulfaté, présentant une alternance de liaisons α -1,3 et β -1,4.

Les chaînes de polymère peuvent être linéaires mais aussi ramifiées. Il existe de nombreux types de branchements et de ramifications : l'amylopectine par exemple se distingue de l'amylose par la présence de ramifications de G-glucopyranoses branchées en α -1,6 toutes les 24-30 unités glucose de la chaîne principale. La « décoration » de la chaîne principale par des unités xylose ou fucose par exemple, est également très courante dans les hémicelluloses terrestres et marines comme les xyloglucanes (Freshour *et al.* 2003). Ainsi, chaque groupement hydroxyle d'un ose pouvant théoriquement réaliser une liaison glycosidique avec un autre ose, la diversité structurale des polysaccharides est bien supérieure à celle de l'ADN ou des protéines, avec plus de 10^{12} possibilités d'hexamères glucidiques théoriques contre 20^6 (36×10^6) hexapeptides et 4^6 ($\approx 4 \times 10^3$) hexanucléotides (Laine 1994). A cela il faut ajouter le degré supplémentaire de complexité généré par les partenariats entre protéines ou lipides et les glucides, au sein des glycoprotéines, des protéoglycanes ou des glycolipides, qui sont des éléments clés dans le fonctionnement des organismes vivants (Gagneux and Varki 1999; Varki and Lowe 2009).

2. Organisation et diversité des polysaccharides algaux

2.1 Localisation et spécificités des polysaccharides algaux

Les polysaccharides représentent la principale voie de stockage du carbone fixé au cours de la photosynthèse par les algues. Elles vont ensuite les utiliser de deux façons principales : 1) comme éléments de structure, afin de constituer la paroi en périphérie de la cellule (cellulose, ulvanes, agars, alginates) et 2) comme réserve énergétique, en les stockant dans le cytoplasme (granules d'amidon) ou la vacuole (laminarane) (tableau 2). Les plantes terrestres et les macroalgues possèdent une organisation générale de la paroi similaire, à savoir une phase cristalline structurante, principalement constituée de cellulose, et une phase matricielle servant de liant, constituée d'une multitude de polysaccharides de nature et d'abondance variables selon les différentes lignées végétales (tableau 2). Cependant, deux différences majeures existent entre la paroi des plantes terrestres et celles des algues : la teneur en cellulose et la présence de polysaccharides sulfatés. Les algues possèdent une plus faible teneur en cellulose que les plantes terrestres. En contrepartie, les polysaccharides de la phase matricielle sont plus abondants, et nombre d'entre eux sont sulfatés, ce qui n'est pas le cas des plantes terrestres (Kloareg and Quatrano 1988).

Tableau 1 : Répartition des différents polysaccharides végétaux au sein des cellules
(adapté de Popper *et al.*, 2011)

Poly saccharides	Taxon	Chloroplastida			Rhodophyta	Phaeophyceae
		Embryophyceae	Charophyceae	Chlorophyta		
Phase cristalline	Cellulose					
		-	-	-	Mannane Xylane	-
Phase matricielle	Hemi-cellulose	Xyloglucane, Mannane			Gluco mannane	-
		Xylane		-	Xylane	Xylofuco glucane sulfaté
		-	-	Glucuronane	-	Xylofucogluc- uronane sulfaté
		MLG (<i>mixed- linked glucan</i>)	-	-	Sulfated MLG	-
	Glucane			-	Glucane	
	Acides	Pectines		Ulvane	-	Alginate
Sulfatés	-	-	Agar Carraghénane Porphyrane		Homofucane	
Réserve	Amidon		Inuline Amidon Laminarine	Glycogène floridéen	Laminarine	

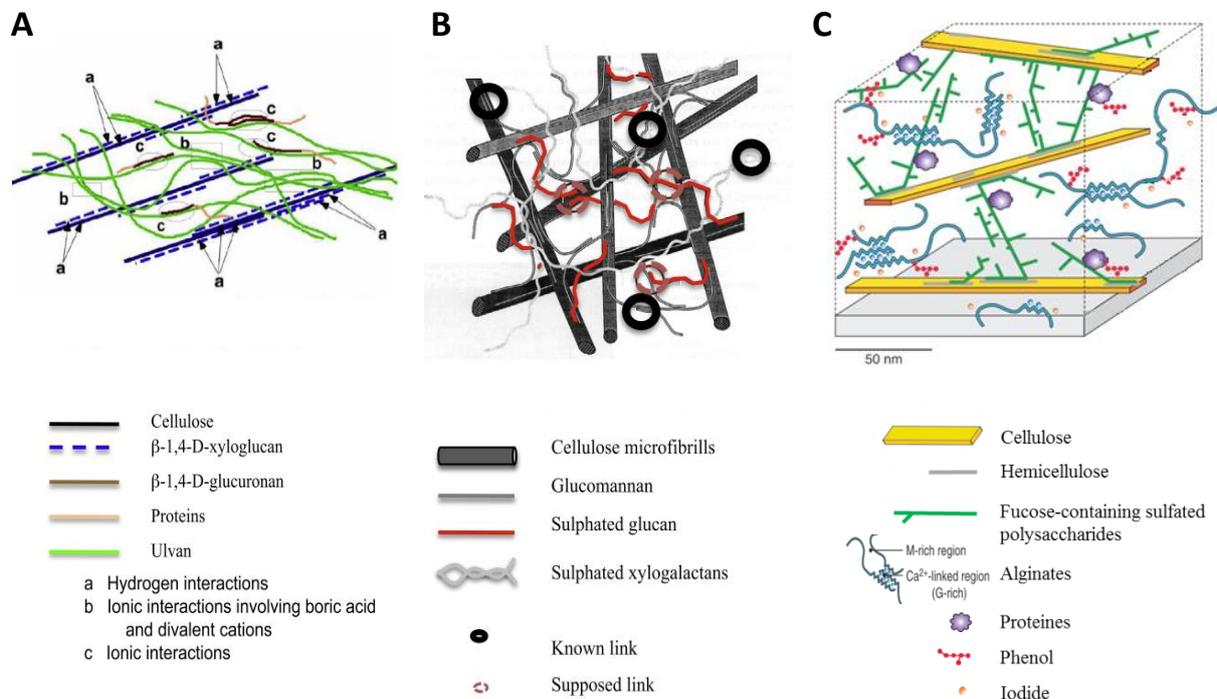


Figure 14 : Représentation schématique de la paroi des algues vertes (A) (Lahaye and Robic 2007), rouges (B) (Lechat 1998) et brunes (C) (d'après Deniaud-Bouët *et al.* 2014)

Grâce au séquençage des génomes des macroalgues brunes *Ectocarpus siliculosus* (Cock *et al.* 2010) et rouges *Chondrus crispus* (Collén *et al.* 2013), il a été montré que la capacité à synthétiser des polysaccharides sulfatés est une caractéristique ancestrale des eucaryotes liée à l'émergence de la vie dans les océans. En témoignent les glycosaminoglycanes (GAGs) de la matrice extracellulaire animale tels que la chondroïtine ou l'héparane sulfate. Cependant, cette capacité aurait été perdue au cours de l'évolution chez les végétaux d'eau douce et terrestres, probablement face à la pauvreté en sulfate du milieu (Michel *et al.* 2010b). La similarité entre les GAGs animaux et les polysaccharides marins sulfatés (PMS) confère à ces derniers des propriétés biomimétiques aux effets bien connus sur le système immunitaire, la coagulation ou les phénomènes de migration, prolifération et régénération cellulaires (voir paragraphe II.3), n'ayant pas d'équivalent chez les végétaux terrestres (Jiao *et al.* 2011; Fleurence and Levine 2016). De façon remarquable, les angiospermes marines, comme *Zostera marina*, qui ont colonisé le milieu marin de façon plus tardive après leur apparition sur terre, ont regagné la capacité à synthétiser des polysaccharides sulfatés (Olsen *et al.* 2016).

2.2 Une complexité à plusieurs niveaux

La vision actuelle de l'architecture pariétale des macroalgues est encore partielle. Différents modèles ont été proposés pour représenter la structuration globale des parois des trois groupes de macroalgues (Figure 13), basés sur des analyses de la composition de différents extraits algaux (aqueux, acides, basiques) (Lahaye and Robic 2007; Deniaud-Bouët *et al.* 2014). Mais de nombreuses questions restent encore en suspens, tant au niveau des interactions entre les composés pariétaux qu'au niveau de la composition fine des chaînes polysaccharidiques. En effet, la variabilité structurale des polysaccharides est fonction de nombreux paramètres : espèce, stade de développement (gamétophyte/sporophyte), âge de l'individu, type de tissu (crampon/stype/lame), géographie et conditions environnementales (température, luminosité, broutage,...) (Kloareg and Quatrano 1988; Robic *et al.* 2009b). A cela s'ajoute la complexité interne des polysaccharides. En effet, sous les noms de familles de polysaccharides, comme les carraghénanes, les fucanes ou les ulvanes, sont regroupées une multitude de sous-familles qui peuvent co-exister au sein d'un même polymère, sous forme de polysaccharides dits hybrides.

Afin d'illustrer cette complexité, nous allons détailler à titre d'exemple le cas des carraghénanes, qui est à ce jour la famille la mieux caractérisée.

2.3 Le cas des carraghénanes

Les carraghénanes sont des polymères de galactose liés alternativement par des liaisons α -1,3 (unité D) et β -1,4 (unité G). Les unités D et G (nomenclature proposée par Knutsen *et al.* 1994) forment l'unité disaccharidique de base qui peut être modifiée par la présence d'un pont 3,6-anhydro (unité DA), la présence de groupements sulfate (unités D6S, DA2S, G4S,...), d'acide pyruvique (P) ou encore de méthyls (M) (Figure 14). Par convention, si les oligosaccharides sont issus de la coupure de la liaison β -1,4, ils appartiendront à la série des néo-carrabioses, et s'ils sont issus de la coupure de la liaison α -1,3 à la série des carrabioses.

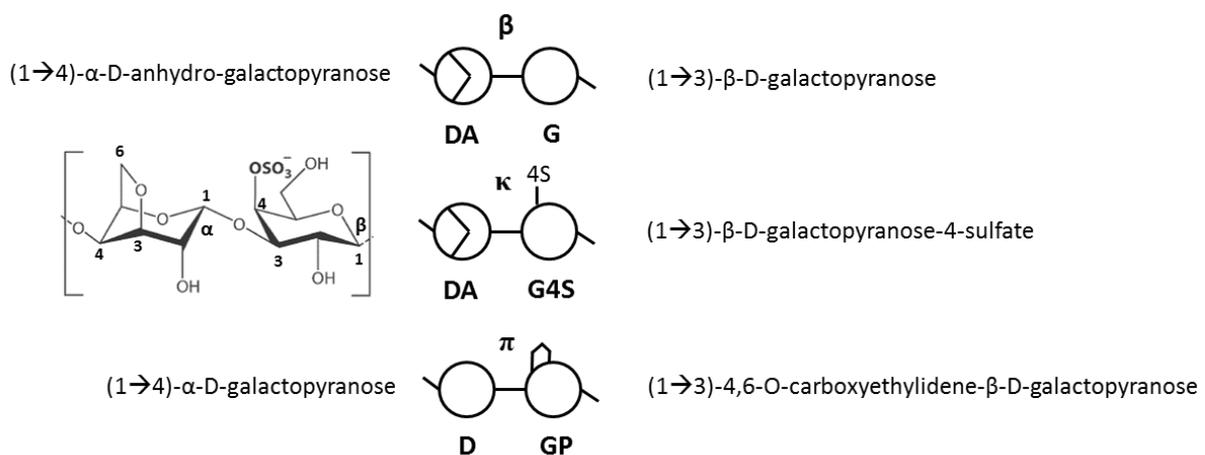


Figure 15 : Schématisation des néo-carrabioses β , κ et π selon la nomenclature de (Knutsen *et al.* 1994). La structure chimique correspondant au néo- κ -carrabiose est donnée pour illustration, ainsi que les noms systématiques de différentes unités galactoses.

Une nomenclature plus ancienne basée sur les lettres grecques (Rees *et al.* 1969) est également couramment utilisée pour classer les disaccharides en sous-familles en fonction de leur degré de sulfatation et de la présence de modifications chimiques. La Figure 15 illustre les différents néo-carrabioses définissant les 15 sous-familles communément admises. Un regroupement des différents types de carraghénanes en trois familles majeures a également été proposé (Craigie 1990) : la famille des κ possède une unité G4S, la famille des λ une unité D2S et la famille des β ne possède aucune des deux unités précédentes.

Lors de la biosynthèse des carraghénanes, la sulfatation enzymatique des unités D et G par des sulfotransférases serait la première étape de modification aboutissant à la formation de précurseurs (entourés en bleu foncé sur la Figure 15), pouvant ensuite être convertis en formes matures par l'action de sulfurylases pour générer le pont 3,6-anhydro sur l'unité DA (entourés en bleu clair) (Craigie 1990). La formation du 3,6-anhydrogalactose (symbolisée par les flèches

sur la Figure 15) peut aussi être réalisée par un traitement chimique alcalin lors de l'exploitation industrielle des carraghénanes, afin de convertir les précurseurs μ - et ν - respectivement en κ - et ι -carraghénanes. Les extraits algaux ainsi enrichis en 3,6-anhydrogalactose présentent de meilleures propriétés gélifiantes (Seisun *et al.* 2010). En effet, l'augmentation des interactions hydrophobes résultant de la présence du pontage 3,6-anhydro favorise la formation d'hélices de carraghénanes indispensables à la formation d'hydrogels.

Cependant, les enzymes algales de biosynthèse du carraghénane sont encore très peu connues (Ficko-Blean *et al.* 2015). Grâce au séquençage en 2013 de la première macroalgue rouge, *Chondrus crispus*, de nombreux gènes codant des enzymes du métabolisme des polysaccharides sulfatés ont été mis en évidence : 65 glycosyltransférases, responsables de l'étape de polymérisation des glucides, 9 carbohydrate-sulfotransférases, impliquées dans la première étape de sulfatation des glucides, et 12 galactose-6-sulfurylases, responsables de la cyclisation du 3-6, anhydrogalactose (Collén *et al.* 2013). Toutefois, les voies de biosynthèse proposées pour les carraghénanes demeurent encore relativement hypothétiques, très peu d'études ayant porté sur le sujet (Wong and Craigie 1978; Zinoun *et al.* 1997; Genicot-Joncour *et al.* 2009).

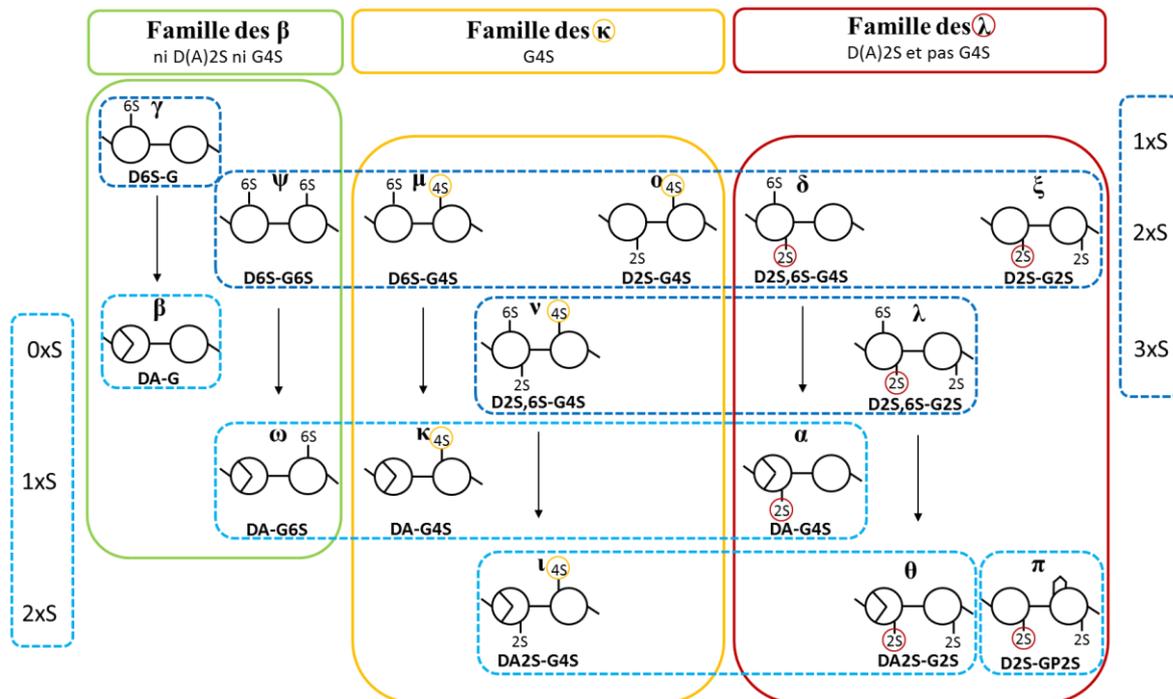


Figure 16 : Organisation des différentes sous-familles de carraghénanes, d'après (Craigie 1990). Les familles β , κ et λ , sont entourées en vert, en jaune et en rouge respectivement. Les sous-familles sont indiquées par les lettres grecques au-dessus des motifs de néo-carrabiose, et sont organisées en fonction du nombre de groupements sulfates et de leur caractère précurseur (bleu foncé) ou mûré (bleu clair). Les flèches indiquent la bioconversion possible d'une sous-famille à une autre, soit démontrée par des enzymes de type sulfohydrolases, soit réalisée chimiquement par addition de soude.

3. Nature et propriétés des polysaccharides algaux

La paroi végétale représente l'interface entre les cellules photosynthétiques et leur environnement, cellulaire ou extérieur. Tout comme son homologue chez les organismes animaux - la matrice extracellulaire - la paroi va jouer de multiples rôles. Support de la cohésion cellulaire et de la croissance tissulaire, elle apporte également protection et résistance aux tissus face aux différents stress biotiques et abiotiques (dessiccation, salinité, rayonnement UV, broutage, agents pathogènes). Les constituants polysaccharidiques y jouent un rôle essentiel grâce à leur structure biochimique. Leurs propriétés osmotiques, liées à leurs charges, à la présence de contre-ions et à leur capacité à former des hydrogels, leur permettent de réguler les échanges aqueux et ioniques des cellules (Kloareg and Quatrano 1988). Ils confèrent également aux tissus algaux leurs propriétés de résistance, de souplesse et de flottabilité, indispensables à la vie en milieu marin. La présence de nombreux groupements hydroxyles le long du squelette carboné confère également aux polysaccharides le rôle de barrière protectrice contre l'oxydation intracellulaire liée aux UV (von Sonntag 1980; Stern and Jedrzejas 2008). Enfin, lors de leur dégradation par des organismes brouteurs ou pathogènes, la libération d'oligosaccharides, dits éliciteurs, représente un signal d'alarme pour les cellules végétales, déclenchant la mise en place d'une machinerie cellulaire de défense (Cosse *et al.* 2007; Thomas *et al.* 2014).

Ces propriétés osmotiques, antioxydantes et de communication cellulaire vont pouvoir être utilisées dans de nombreux domaines allant de l'agroalimentaire et de l'agriculture à la pharmaceutique en passant par la cosmétologie et la conception de biomatériaux (Figure 16). Dans les applications impliquant la reconnaissance/communication cellulaire, c'est à la fois le caractère chargé des polysaccharides ainsi que leur taille qui semblent corrélés à leurs propriétés. Les interactions électrostatiques aspécifiques sont impliquées dans les propriétés anticoagulantes ou bactériostatiques notamment. Mais des mécanismes plus complexes, mettant en jeu la reconnaissance des motifs de sulfatation et la nature des oses, pourraient être impliqués dans les propriétés antibactériennes, antivirales et anticancéreuses. Les détails de ces mécanismes de reconnaissance cellulaire sont encore peu connus. En effet, leur étude est compliquée en raison notamment du manque de caractérisation précise des substrats algaux utilisés dans les études *in vitro* comme *in vivo*, lié à la complexité de leurs structures (Jiao *et al.* 2011).

Polysaccharides de paroi des algues

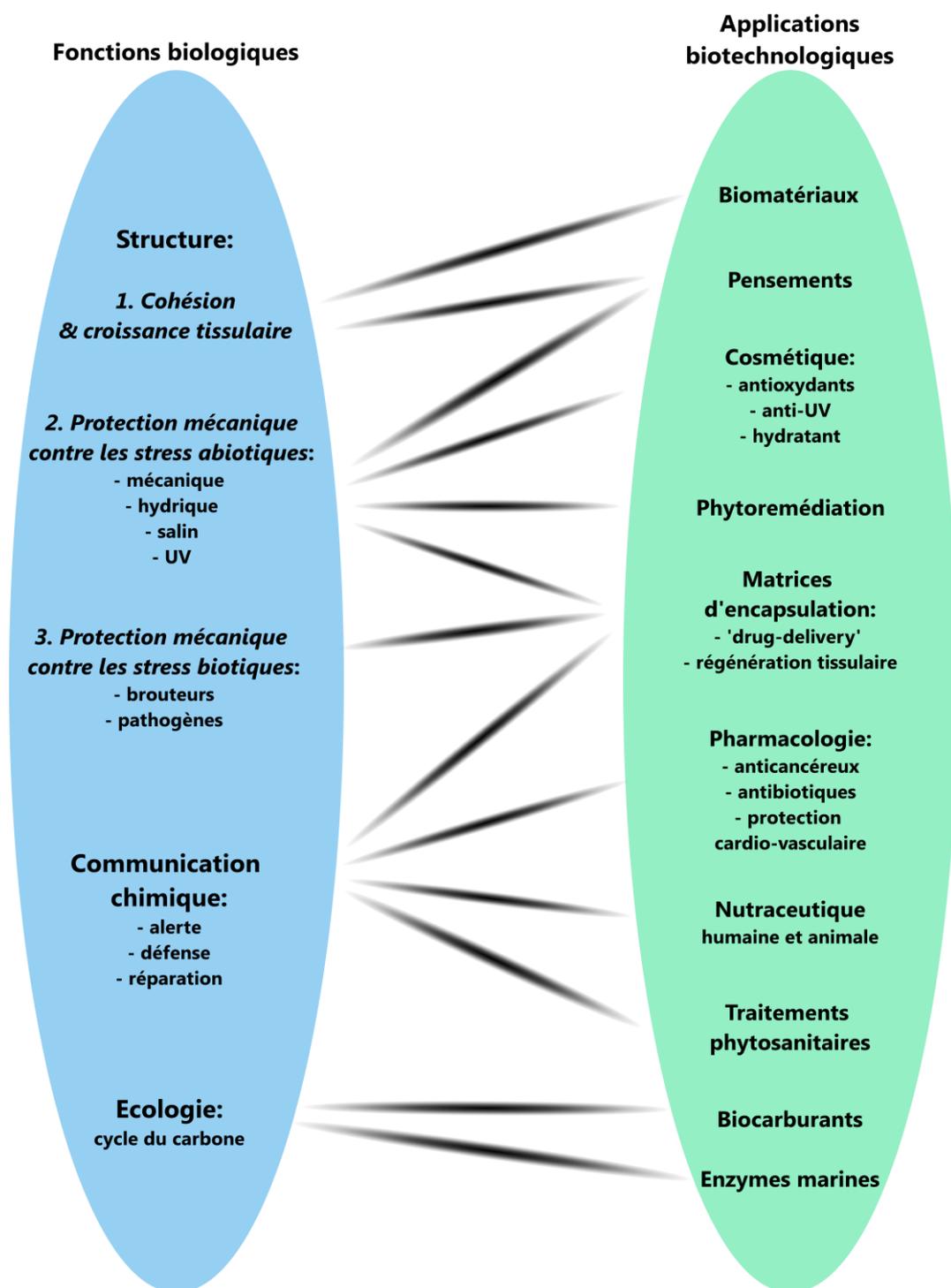


Figure 17 : Corrélation entre les rôles et propriétés biologiques des polysaccharides algaux et leurs applications biotechnologiques.

Un bref aperçu de cette complexité structurale des polysaccharides au sein des trois groupes de macroalgues, et des fonctions qui leur sont associées va être exposé dans les sous-parties suivantes.

3.1 Les polysaccharides d'algues vertes

Les polysaccharides majoritaires des algues vertes sont les ulvanes. Ils représentent environ 30 % du poids sec de l'algue. Ces polymères sont faiblement ramifiés et principalement composés de rhamnose, d'acides glucuronique et iduronique ainsi que de xylose, agencés selon divers motifs disaccharidiques. Les deux motifs principaux sont les acides ulvanobiourinique-3-sulfate de type A (β -D-GlcA (1 \rightarrow 4) α -L-Rha 3S \rightarrow 1) et B (α -L-IdoA (1 \rightarrow 4) α -L-Rha 3S \rightarrow 1) (Figure 17).

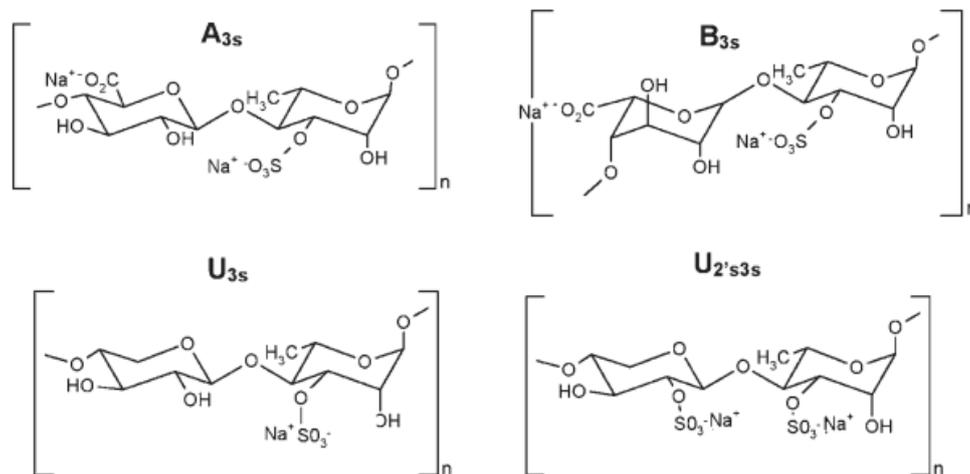


Figure 18 : Principaux motifs disaccharidiques des ulvanes (Robic *et al.* 2009a) **extraits d'*Ulva sp.***. A_{3s} et B_{3s}: acides ulvanobiouriniques; U_{3s} et U_{2's3s}: ulvanobioses

Les ulvanes peuvent former des gels thermoréversibles sous l'action de l'acide borique, de cations divalents et de la température (Quemener *et al.* 1997; Lahaye and Robic 2007). Les propriétés biologiques des ulvanes relevées dans la littérature sont répertoriées dans le tableau xx en annexe (Alves *et al.* 2013; Lahaye and Robic 2007; Violle *et al.* 2017). L'originalité majeure des ulvanes vis-à-vis des autres polysaccharides algaux est sa richesse en acides iduroniques et glucuroniques, qui sont des constituants majeurs des GAGs animaux. Ses applications biomédicales en lien avec les pathologies musculo-squelettiques ou cutanées (Bhatnagar and Bhatnagar 2015), la conception de composés anti-vieillessement cutané (Adrien *et al.* 2017) ou la synthèse d'analogues de l'héparine sont donc tout particulièrement prometteuses. Le rhamnose, qui rentre également dans la composition des ulvanes, peut aussi

servir de précurseur dans la synthèse chimique de composés aromatiques et de dérivés acétylés et benzoylés aux propriétés antioxydantes augmentées (Qi *et al.* 2006). Le retard relatif de l'exploitation des ulvanes par rapport aux autres polysaccharides algaux représente toutefois une opportunité pour l'innovation industrielle et la création de marchés de niches, comme en témoignent les récentes études sur les voies de bioraffinerie et de valorisation des algues vertes (Magnusson *et al.* 2016; Glasson *et al.* 2017).

3.2 Les polysaccharides d'algues rouges

3.2.1 L'agar

L'agar peut représenter jusqu'à 42 % de la matière sèche chez des espèces comme *Gelidium*. Il s'agit en fait d'un mélange d'agarose (Figure 18), un polymère linéaire d'agarobioses aux propriétés gélifiantes, et d'un composé communément appelé agaropectine, une version chargée de l'agarose, modifiée par la présence de groupements sulfates, méthyles et acides pyruviques (Seisun *et al.* 2010). L'agarose est constitué d'une alternance de résidus D-galactose et 3,6-anhydro-L-galactose liés respectivement en β -1,4 et α -1,3.

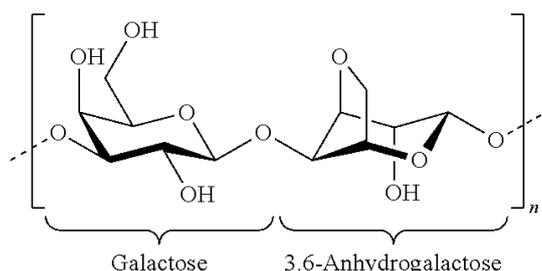


Figure 19 : Structure chimique de l'unité de répétition de l'agar, l'agarobiose.

Environ 10 % de la production mondiale sert à la production d'agar technique pour les usages biotechnologiques cités au paragraphe I.1.3. Malgré les effets hypoglycémiants, anti-agrégation des globules rouges ou anti-tumoraux de l'agar, et antioxydants des oligo agaroses (Kraan 2012), les 90 % restants sont actuellement absorbés par l'industrie agroalimentaire pour les propriétés gélifiantes de l'agar (paragraphe I.1.1). Bien qu'il soit encore matière à débats, le mécanisme communément admis pour expliquer le processus de gélification repose sur la succession de deux étapes (Figure 19A), la première aboutissant à la formation de doubles hélices, qui s'agrègent ensuite pour former un réseau hydrocolloïdal structuré (Fernandez *et al.* 2007).

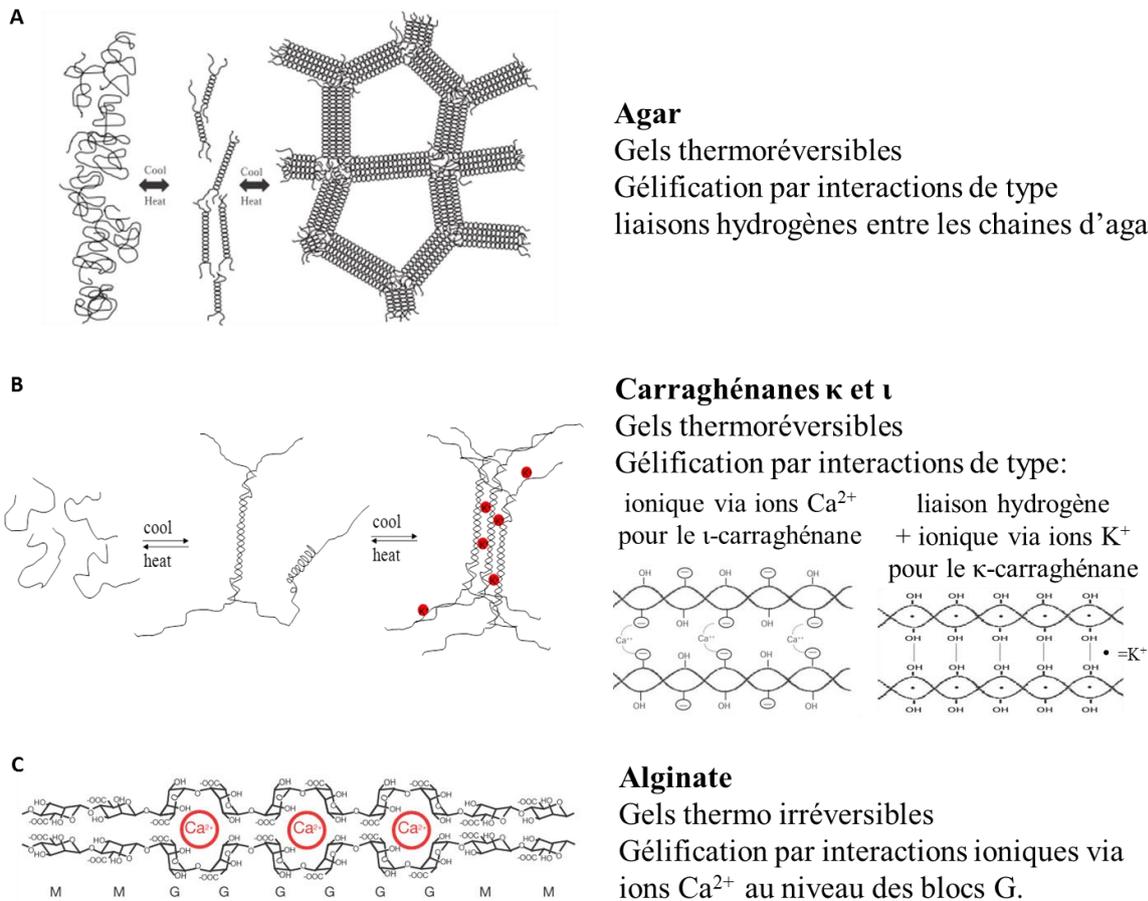


Figure 20 : Modèles proposés pour expliquer les mécanismes de gélification des phycocolloïdes exploités industriellement.

3.2.2 Les carraghénanes

Les carraghénanes peuvent représenter jusqu'à 88 % de la matière sèche de l'algue, chez des espèces comme *Kappaphycus*. Seuls les κ - et ι -carraghénanes possèdent la capacité de gélifier en solution aqueuse pour former des hydrogels. Le λ -carraghénane, fortement sulfaté mais dépourvu de pont 3,6-anhydro, possède seulement une forte viscosité. La présence du 3,6-anhydro-D-galactose semble en effet primordiale pour la gélification des carraghénanes. Même si le mécanisme exact de gélification est encore sujet à débat, la conformation chaise ${}^1\text{C}_4$ imposée par le pont 3,6-anhydro de l'unité DA du κ -carraghénane (ou DA2S du ι -carraghénane), en alternance avec la conformation chaise ${}^4\text{C}_1$ du galactose sulfaté (G4S), semble en effet favoriser la formation d'hélices initiales, avant l'étape suivante d'agrégation des hélices *via* la présence des groupements sulfate et de contre-ions (Figure 19B) (Lawson and Rees 1970; Van De Velde *et al.* 2002).

Outre les débouchés en l'agro-industrie, qui absorbe 80 % de la production mondiale (voir paragraphe I.1.1), 10 % de la production de carraghénanes est valorisé dans le secteur

pharmaceutique. En plus de leur utilisation comme adjuvants (agents de suspension ou stabilisants), leurs propriétés anti-coagulantes, anti-tumorales et anti-virales, en particulier contre les virus sexuellement transmissibles comme le VIH ou le VHS, sont très bien documentées (Kraan 2012; Fleurence and Levine 2016; Necas and Bartosikova 2013). Dans le domaine de la santé végétale, le taux de sulfatation des carraghénanes a été corrélé au pouvoir éliciteur des extraits testés sur des plants de tabac notamment (Mercier *et al.* 2001).

Cependant, les carraghénanes sont sujets à controverses. En effet, leur utilisation comme agents pro-inflammatoires (*pow-oedema*), mais également des études indiquant les effets carcinogéniques des carraghénanes dégradés comme non dégradés (Tobacman 2001) remettent en question leur innocuité alimentaire. Des études cherchant à déterminer l'origine de ces effets pro-inflammatoires sembleraient indiquer que la liaison α -1,3 ainsi que la présence du 3,6-anhydrogalactose, inexistant chez les animaux, seraient des épitopes induisant la réponse inflammatoire de l'organisme (Bhattacharyya *et al.* 2010). Le fort taux de sulfatation pourrait également avoir un impact sur l'activité des sulfatases endogènes humaines et la composition des GAGs de la muqueuse intestinale (Yang *et al.* 2012). Cependant, la question n'est pas encore tranchée.

3.3 Chez les algues brunes

3.3.1 Les alginates

Les alginates représentent 30 à 40 % de la matière sèche de l'algue. Ce sont des copolymères d'acides D-mannuronique (blocs M) et L-guluronique (blocs G, épimères en C₅ des blocs M) liés respectivement en β -1,4 et α -1,4 (Figure 20). La répartition et le ratio des blocs M et G déterminent les propriétés gélifiantes des alginates (Seisun *et al.* 2010). En effet, ce sont les blocs G qui, grâce à leur conformation dite « en boîte à œufs » vont piéger des ions calcium (Grant *et al.* 1973), induisant ainsi la formation d'un réseau colloïdal (Figure 19C).

Outre les applications présentées en I.1.1 et I.1.3, environ 5 % de la production mondiale est utilisée à des fins médicales. De nombreux médicaments utilisent les propriétés mucoprotectrices de l'alginate dans le cadre des troubles digestifs (*ex.* : Gaviscon®). Il est aussi utilisé pour ses propriétés protectrices et réparatrices dans le traitement des plaies ouvertes, mais également pour ses propriétés hémostatiques pour imprégner le matériel médical en chirurgie (Bhatnagar and Bhatnagar 2015; Holdt and Kraan 2011). Ses effets hypolipémiant et hypocholestérolémiant, antioxydants et anti-cancéreux sont également démontrés, ainsi que

ses effets pré-biotiques en complément alimentaire pour les animaux (O'Sullivan *et al.* 2010).

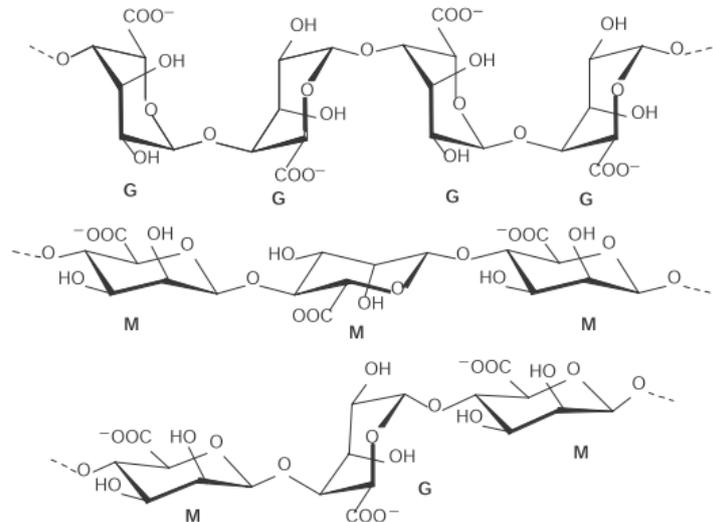


Figure 21 : Structure chimique de l'alginate et de ses blocs M et G (Seisun *et al.* 2010)

3.3.2 Les fucoïdanes

Le terme fucoïdane désigne une famille de fucanes sulfatés extraits des algues brunes. L'autre source de fucanes sulfatés étant les invertébrés marins. Sont inclus dans cette définition les polysaccharides constitués majoritairement de fucoses sulfatés et contenant moins de 10 % d'autres oses (acide glucuronique, mannose, galactose, xylose) (Berteau and Mulloy 2003). Les fucoïdanes peuvent représenter jusqu'à 20 % de la matière sèche chez *Fucus* sp.. Leurs structures varient d'une espèce à l'autre, tant en termes de liaisons glycosidiques de la chaîne principale qu'en termes de nature et de position des ramifications (Figure 21).

Les principaux effets biologiques des fucoïdanes sont anticoagulants, immunostimulants, antiviraux et anti-tumoraux, mais également anti-inflammatoires et antioxydants. Ils possèdent également des propriétés élicitrices en santé végétale (Klarzynski *et al.* 2003). Comme pour les autres polysaccharides sulfatés, le poids moléculaire et le degré de sulfatation jouent un rôle majeur dans leurs activités biologiques (Haroun-Bouhedja *et al.* 2000; Deniaud-Bouët *et al.* 2017).

3.3.3 La laminarine

De tous les polysaccharides issus des macroalgues, la laminarine est le seul qui ne soit pas un constituant pariétal mais un polysaccharide de réserve spécifique des algues brunes, stocké au sein du cytoplasme (Michel *et al.* 2010a). Elle est constituée d'environ 25 résidus de

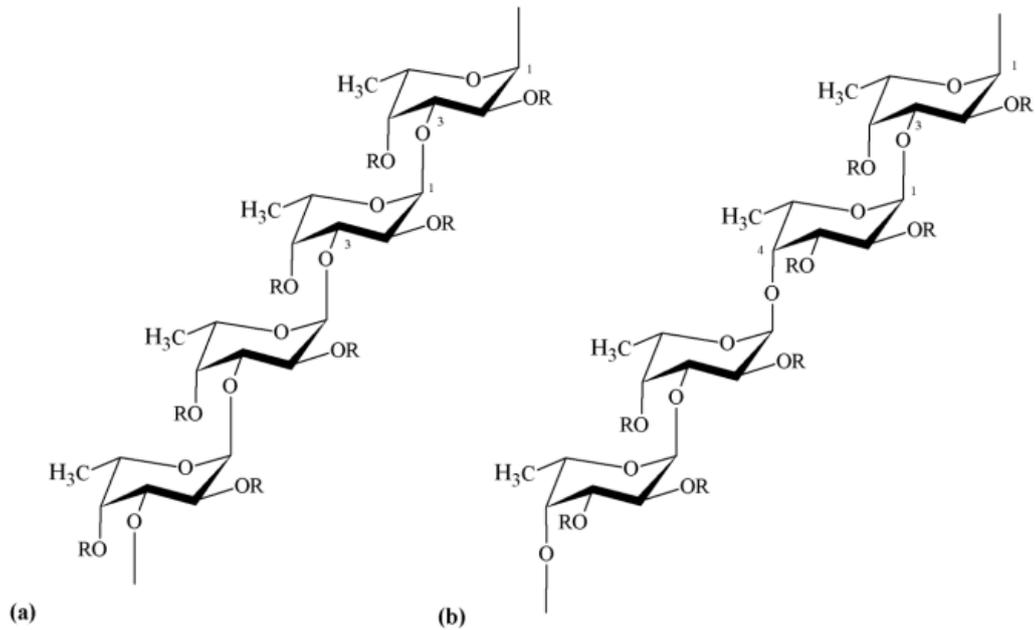


Figure 22 : Exemple de structures de fucoïdanes observées dans les ordres des Laminariales et Chordariales (a) et des Fucales (b). La chaîne principale est constituée de résidus α -L-fucose liés en (1,3) (a) ou en alternance (1,3) et (1,4) (b). R : position potentielle de ramification de type osidique (acide glucuronique, mannose, galactose, xylose, fucose) ou de groupements sulfate ou acétate (Cunha and Grenha 2016)

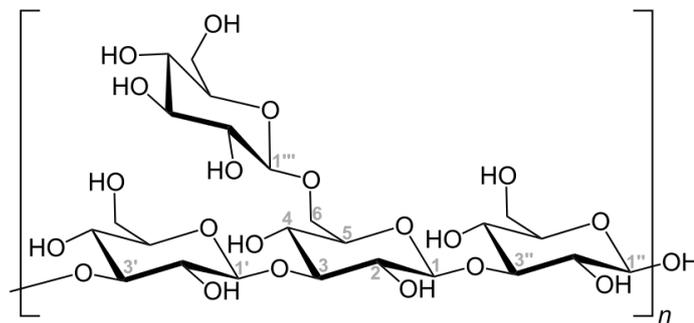


Figure 23 : Structure chimique de l'unité de répétition de la laminarine.

β -1,3-glucane, branchés occasionnellement en β -1,6 (Figure 22) et peut représenter jusqu'à 32 % du poids sec de l'algue (Elyakova and Zvyagintseva 1974).

Ses propriétés élicitrices et antivirales en santé végétale ont été démontrées, et font l'objet d'une commercialisation de produits phytosanitaires, notamment par les laboratoires Goëmar (Arysta LifeScience) en France (Aziz *et al.* 2003; Gauthier *et al.* 2014). Ses propriétés anticoagulantes et immunostimulantes sont également bien connues, de même que ses effets hypocholestérolémiants, hypotenseurs et antibactériens (Holdt and Kraan 2011; Fleurence and Levine 2016).

Ce bref aperçu de la diversité des polysaccharides algaux permet de mesurer la richesse que représente cette ressource. Outre son exploitation par l'homme, cette biomasse représente surtout une source essentielle de carbone pour la faune comme pour la flore bactérienne marine. Cependant, pour y avoir accès, ces organismes non-photosynthétiques ont dû développer des stratégies d'extraction du carbone pariétal. Or, la liaison glycosidique est la liaison chimique la plus stable du vivant, avec une demi-vie de l'ordre de 4,7 à 9,4 millions d'années selon qu'il s'agit respectivement d'anomères β ou α (Wolfenden *et al.* 1998). Il a donc fallu « inventer » des moyens d'accélérer le phénomène, pour le rendre compatible avec les processus métaboliques des organismes producteurs (remodelage des parois végétales, mobilisation des réserves glucidiques) et consommateurs de glucides (recyclage de la matière organique par les bactéries et les eucaryotes non photosynthétiques). Et pour faire face à ce défi de grande ampleur, l'évolution a mis au point des outils hautement efficaces que sont les enzymes de dégradation des glucides.

III. Les enzymes de dégradation des glucides

1. Définition et classification

1.1 Généralités

Bien que l'utilisation des propriétés catalytiques des enzymes par l'Homme remonte probablement aussi loin que la fabrication du premier fromage ou du premier vinaigre, l'identification de leur nature biochimique ainsi que la compréhension de leur mode d'action sont bien plus récents. Le terme « enzyme » a été inventé en 1878 par le physiologiste allemand William Kühne en combinant les mots grecs *en* et *zyme* signifiant « dans » et « levure », pour désigner la substance responsable de la fermentation par les levures. En 1898, Duclaux proposa de faire terminer par le suffixe *-ase* le nom des enzymes afin de les distinguer des substances biologiques dépourvues d'activité catalytique. Et ce n'est qu'en 1926 que la nature protéique des enzymes fut démontrée grâce à la cristallisation de l'uréase par James B. Sumner (Purich 2010). D'un point de vue thermodynamique, les enzymes sont des catalyseurs biologiques qui permettent, comme tout catalyseur, d'abaisser la barrière énergétique séparant l'état initial de l'état final d'une réaction chimique. Leur effet est donc d'accélérer une réaction chimique d'un facteur pouvant aller jusqu'à 10^{17} . D'un point de vue biologique, ce sont des protéines synthétisées par toute cellule vivante et encodées dans le matériel génétique de tout organisme biologique, virus compris. Véritables actionneurs des processus biologiques, leurs fonctions

sont aussi nombreuses et variées que leurs substrats. Au cours de cette thèse, je me suis intéressée à une catégorie d'enzymes bien précise, ayant pour substrat les glucides.

1.2 La classification CAZy

Les enzymes actives sur les glucides sont regroupées sous le terme de CAZymes (*Carbohydrate Active Enzymes*). Cette dénomination englobe aussi bien des enzymes impliquées dans la synthèse - les glycosyltransférases (GT) - que dans la dégradation des glucides : les glycoside hydrolases (GH), les polysaccharide lyases (PL), les carbohydrate-estérases (CE), et des enzymes auxiliaires à activité redox (AA)..

Dans chacune de ces catégories, les enzymes sont regroupées au sein de familles selon leur similarité de séquence en acides aminés et non en fonction leur spécificité de substrat, identifiée par un numéro EC (Henrissat *et al.* 1991). Ainsi au sein d'une même famille, par exemple la famille 16 des GH, peuvent se trouver des enzymes possédant différents numéros EC : lichénases (EC 3.2.1.73), κ -carrageenases (EC 3.2.1.83) ou β -agarases (EC 3.2.1.81) entre autres. La liste actualisée de toutes les familles et de leurs membres connus est disponible dans la base de données CAZy (<http://www.cazy.org/>), qui recense les données de séquençage dans tous les domaines du vivant

Les enzymes regroupées au sein d'une même famille partagent un même repliement structural, ce qui représente un atout pour la détermination structurale de nouvelles enzymes grâce au remplacement moléculaire (Davies and Henrissat 1995). De plus, la machinerie catalytique et le mécanisme moléculaire sont également conservés, permettant ainsi de prédire le mécanisme catalytique à partir de la seule séquence (Henrissat and Callebaut 1995).

Cependant, le nombre de repliements structuraux étant limité par rapport au nombre de familles répertoriées par CAZy, certaines familles vont posséder un même repliement. Si, de plus, elles partagent la même machinerie catalytique et le même mécanisme réactionnel, tout en possédant une similarité de séquence trop faible pour faire partie de la même famille, elles sont regroupées au sein de clans ou de superfamilles. Ainsi, le clan des GH-B est constitué des familles GH7, contenant des cellulases, et GH16, contenant des κ -carraghénases, β -agarases, laminarinases, lichénases et XETs (*xyloglucan endotransglycosylases*). Les différences structurales associées aux résidus catalytiques ont ainsi permis de retracer l'évolution phylogénétique de ce clan : l'insertion d'un résidu supplémentaire (x) au sein du motif conservé contenant les résidus catalytiques ExDx(x)E crée une structure de type « β -bulge » à la place d'un brin β régulier

(Figure 23). Ce β -bulge aurait été présent chez l'ancêtre commun du clan des GH-B, puis perdu au cours de l'évolution par les lichénases et les XETs (Michel *et al.* 2001a).

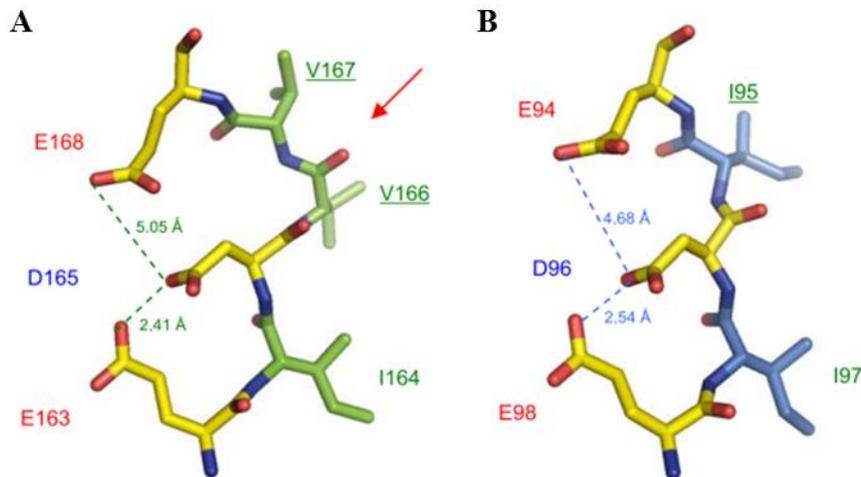


Figure 24 : Organisation structurale des résidus catalytiques au sein du clan des GH-B. A : Détail du site actif de la κ -carraghénase de *P. carrageenovora*, de type β -bulge (EIDVYE), la flèche rouge pointant le résidu V167 supplémentaire induisant la formation du β -bulge ; B : Détail du site actif de la xyloglucanase de *T. majus*, de type brin β régulier (EIDIE).

La classification CAZY, basée sur la séquence protéique, présente donc l'avantage de mieux refléter les caractéristiques structurales des protéines que la classification selon la spécificité de substrat. Elle permet également de retracer le cheminement évolutif entre les enzymes, tout en fournissant des informations sur le mécanisme enzymatique qui est généralement conservé au sein d'une famille. Elle illustre également la difficulté de tirer des conclusions quant à la spécificité de substrat à partir de l'appartenance à une famille.

1.3 Les glycosides hydrolases

Au cours de cette thèse, je me suis plus particulièrement intéressée au groupe des glycosides hydrolases. Il s'agit du groupe historique qui a servi de base de départ à la classification CAZY proposée par Bernard Henrissat en 1991. Les GH comptent à ce jour 145 familles couvrant 191 numéros EC, et près de 480 000 ORF. Ces enzymes catalysent la réaction d'hydrolyse de la liaison glycosidique positionnée par entre les sous-sites -1 et +1. Les sous-sites correspondent aux zones d'interaction entre l'enzyme et les unités glucidiques, et sont numérotés de façon négative du côté non réducteur et positive du côté réducteur de la chaîne polyosidique.

(Davies *et al.* 1997) (Figure 24).

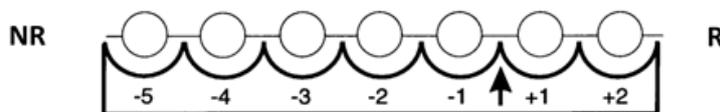


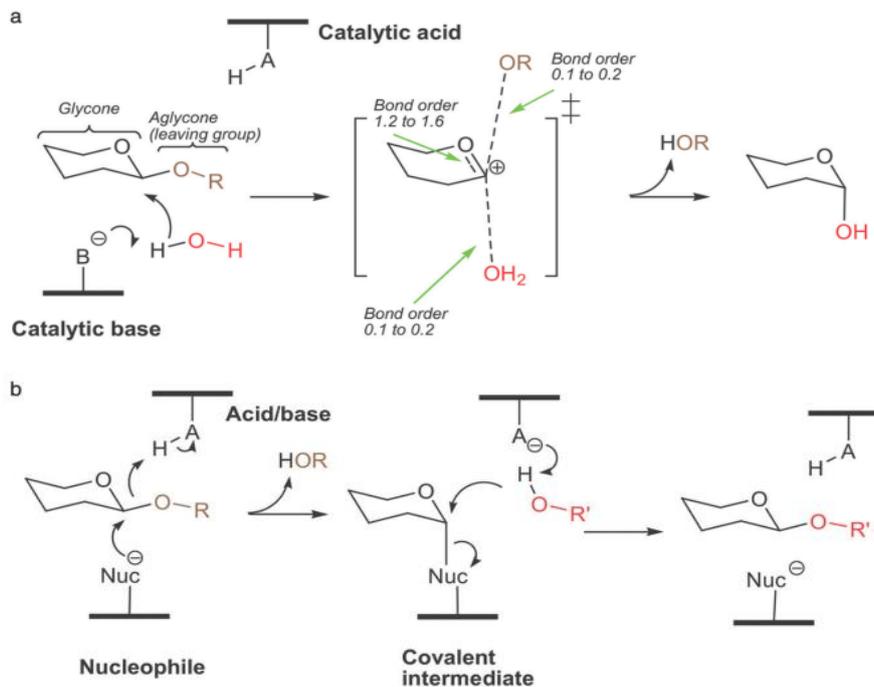
Figure 25 : Représentation schématique et nomenclature des sous-sites catalytiques d'une CAZyme liée à son substrat (adapté de Davies *et al.* 1997)

1.3.1 Mécanisme réactionnel

En fonction du type de mécanisme réactionnel, l'hydrolyse aboutira soit à la rétention soit à l'inversion de la configuration anomérique de l'ose situé du côté non réducteur, selon le mécanisme proposé par Koshland en 1953 (Koshland 1953) (Figure 25). Ainsi, dans le cas des mécanismes de simple déplacement, une molécule d'eau va réaliser l'attaque nucléophile du carbone alpha de l'ose situé du côté non réducteur de la liaison glycosidique (sous-site catalytique -1) grâce à la présence d'un acide aminé catalytique de l'enzyme jouant le rôle de base. Un autre acide aminé catalytique va jouer le rôle d'acide en fournissant un hydrogène au groupe partant du côté réducteur de la liaison glycosidique (sous-site catalytique +1) permettant la rupture de la liaison mais avec inversion de la configuration du sucre en -1. Dans le cas d'un mécanisme de double déplacement, c'est un des acides aminés catalytiques de l'enzyme qui va jouer le rôle de nucléophile et attaquer le C α en formant un intermédiaire covalent avec le sucre. Un autre acide aminé va jouer le rôle d'acide/base en fournissant un proton au groupe partant puis en déprotonant une molécule d'eau qui va apporter son groupement hydroxyle au C α et ainsi libérer le sucre en -1 dans la même configuration anomérique qu'au départ. Si à la place de la molécule d'eau il s'agit d'un autre sucre, l'activité sera de type transglycosidase, aboutissant à la synthèse d'un oligosaccharide de degré de polymérisation supérieur à celui du substrat initial. Les enzymes de type transglycosydases appartiennent à la catégorie des GH procédant par rétention de la configuration anomérique, comme les XETs de la famille des GH16 par exemple (Hrmova *et al.* 2009). Si certaines possèdent une activité transglycosylase stricte, la plupart catalysent les deux types de réaction, hydrolyse et transglycosylation, dans

des proportions variables en fonction des conditions réactionnelles, la concentration et la nature chimique de l'accepteur, molécule d'eau ou oligosaccharide (Sinnott 1990; Bissaro *et al.* 2015).

Figure 26 : Mécanismes enzymatiques des glycosides hydrolases (Davies and Sinnott 2008). (a) Simple déplacement avec inversion de la configuration anomérique du C α et (b) double déplacement avec conservation de la configuration anomérique du C α . La légende *Bond order* se réfère à l'ordre de liaison entre les différents atomes impliqués dans l'état de transition. Ses



valeurs faibles illustrent le caractère fortement dissociatif de cet état de transition.

Le contrôle de l'équilibre thermodynamique entre ces deux activités, notamment par ingénierie enzymatique, représente un enjeu majeur dans les études de glycochimie de synthèse (Mackenzie *et al.* 1998; Zeuner *et al.* 2014; David *et al.* 2017).

Les glycosyl transférases (GT) constituent l'autre catégorie d'enzymes capable de synthétiser une liaison glycosidique. Mais contrairement aux transglycosidases, elles nécessitent la présence d'un sucre phosphaté activé, un « nucléotide-sucre », comme donneur de glycoside (David 1995)

1.3.2 Topologie et mode d'action

Comme toutes les enzymes, les GH peuvent interagir avec leur substrat selon un mode exolytique, c'est-à-dire à partir d'une extrémité de la chaîne polysaccharidique, ou bien selon un mode endolytique, c'est-à-dire à n'importe quel endroit de la chaîne polysaccharidique. Par

ailleurs, certaines GH présentent la capacité de rester fixées à leur substrat entre deux événements catalytiques successifs, générant un déplacement de l'enzyme en sens inverse du groupement partant selon un mécanisme dit processif. Ce mode d'action peut être observé aussi bien chez des enzymes de type exolytique que de type endolytique. Il se prête plus particulièrement à la dégradation de substrats dont la structure semi-cristalline ou les interactions ioniques structurantes rendent l'accès particulièrement difficile aux enzymes, comme la cellulose (Varrot *et al.* 2003) ou encore les hydrogels de κ -carraghénane (Lemoine *et al.* 2009).

Grâce aux données structurales, ces différents modes d'action peuvent être corrélés à la topologie de leur site actif. Trois topologies ont ainsi été définies : en forme de poche, de crevasse (ou gorge) et de tunnel (Figure 26).

La topologie en forme de poche est caractéristique des exo-enzymes, qui reconnaissent spécifiquement les extrémités des chaînes polysaccharidiques. Dans ce cas, la surface de contact entre l'enzyme et le substrat est relativement restreinte, limitant la taille des oligosaccharides clivés (*ex.* : monosaccharides générés par les β -galactosidases ou β -glucosidases) (Wiesmann *et al.* 1995; Czjzek *et al.* 2001).

La topologie en forme de crevasse (ou gorge) présente souvent un long canal à la surface de l'enzyme. Cela offre la possibilité d'une interaction avec plusieurs unités glucidiques et donc une fixation en différents points de la chaîne polysaccharidique. Cette topologie est couramment rencontrée chez les enzymes endolytiques, comme des endocellulases (Spezio *et al.* 1993), des chitinases (Fusetti *et al.* 2002), des alginates lyases (Osawa *et al.* 2005) ou des laminarinases (Labourel *et al.* 2014). Elle peut être associée ou non à un mécanisme processif (Payne *et al.* 2012).

Enfin, la topologie de type tunnel est créée par l'allongement des boucles protéiques bordant la gorge catalytique et la recouvrant en partie. Identifiée initialement chez les cellobiohydrolases, cette topologie relativement rare par rapport aux deux précédentes est caractéristique des enzymes processives (Henrissat and Davies 1997). Elle est présente notamment chez les κ - et ι -carraghénases de *P. carrageenovora* et *A. fortis* respectivement (Michel *et al.* 2001a, 2001b). Le caractère endolytique de certaines de ces enzymes a conduit à proposer un mécanisme d'entrée du substrat via une ouverture temporaire du tunnel, et ce grâce à la flexibilité des boucles bordant le tunnel.

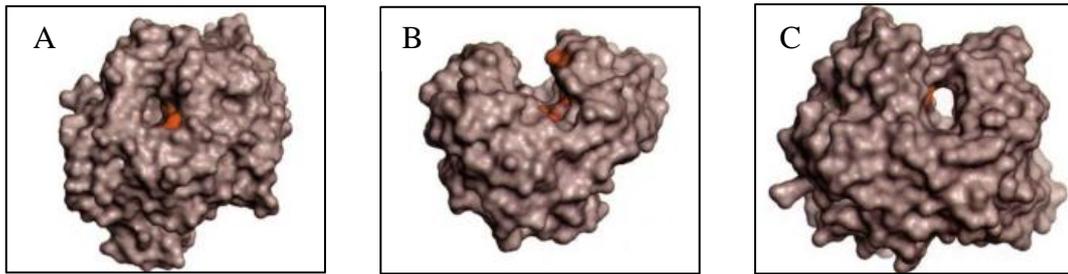


Figure 27 : Topologies associées au site catalytique des GH : poche (A), crevasse ou gorge (B) et tunnel (C), d’après Davies and Henrissat 1995. A : Glucoamylase d’*A. awamori*, B : Endoglucanase E de *T. fusca*, C : Cellobiohydrolase II de *T. reesei*. Les résidus catalytiques sont colorés en rouge.

1.3.3 Modularité

Un module est un domaine fonctionnel protéique, dont le repliement indépendant assure la conservation de l’activité, qu’il soit ou non associé à d’autres modules. En plus des modules possédant une activité catalytique (GH, PL, CE, GT et AA), la base de données CAZy recense également les modules complémentaires dits CBM (*Carbohydrate Binding Modules*). Ces modules sont fréquemment présent au sein d’enzymes actives sur des polysaccharides récalcitrants à la dégradation, telle que la cellulose cristalline ou la chitine (Wang *et al.* 2008; Horn *et al.* 2006b). Dépourvus d’activité catalytique propre, ils possèdent une forte affinité pour leur substrat et facilite donc l’activité du module catalytique auquel ils sont associés (Guillén *et al.* 2010). La modularité est fréquente chez les glycosides hydrolases, et peut engendrer des difficultés pour la résolution structurale. En effet, les modules sont reliés entre eux par des régions charnières (*linkers*) de tailles variables, dont la flexibilité peut poser problème lors du processus de cristallogénèse (cf. chapitre 4). Lors de la stratégie de clonage, il est donc important de tenir compte de cette modularité pour tenter de cloner les différents modules indépendamment (cf. chapitre 1).

2. Les CAZymes de bactéries marines

Dans le milieu marin comme sur terre, les CAZymes impliquées dans la dégradation des polysaccharides, et en particulier les GH, sont synthétisées par deux types d’organismes. Par les organismes photosynthétiques tout d’abord, qui s’en servent au cours des phénomènes de croissance cellulaire et de remodelage tissulaire (xyloglucan transglycosidases /hydrolases et expansines végétales, chitinases associées à la mue des insectes, ...) (Strohmeier *et al.* 2004; Merzendorfer and Zimoch 2003). Mais aussi et surtout par les microorganismes qui leur sont

associés, de façon pathogène, symbiotique ou commensale, et qui vont extraire grâce à ces enzymes le carbone stocké dans les polysaccharides de l'organisme source (Barbeyron *et al.* 2016). Les enzymes étudiées au cours de ce travail de thèse appartiennent à cette seconde catégorie, et proviennent en l'occurrence de bactéries marines vivant au contact des macroalgues.

2.1 Rôles écologiques

La production primaire des océans est comparable à celle des plantes terrestres, et représente 46,2 % CO₂ fixé à l'échelle mondiale (Field 1998). Les producteurs primaires marins sont majoritairement constitués du phytoplancton, qui fixe à lui seul plus de 90 % du CO₂. Mais au sein des écosystèmes côtiers, les macroalgues, bien que fixant proportionnellement moins de carbone à l'échelle mondiale (4,2 %), sont des acteurs essentiels (Gattuso *et al.* 1998). En plus de leur rôle dans la fixation du CO₂ (16 % de la production primaire côtière soit 2 Gt de carbone par an), elles possèdent de multiples fonctions écosystémiques : abri contre les prédateurs, zone nourricière pour les pontes et juvéniles, support de fixation et source de nourriture pour de nombreuses espèces d'Eucaryotes – animaux et végétaux - et de Procaryotes, aussi bien en mer qu'à terre (Bertness *et al.* 1999; Martin *et al.* 2014). Soulignons ici l'importance, trop souvent négligée par les politiques de gestion des plages, de la 'laisse de mer' (algues arrachées de leur substrat et laissées en échouage sur le sable lorsque la mer se retire) dans l'apport de nutriments qui permettent la colonisation et la fixation des dunes par les espèces végétales pionnières du bord de mer (Colombini and Chelazzi 2003).

La plus grande partie de la biomasse de carbone fixé par les algues est ensuite reminéralisée par les organismes vivant à leur contact, aussi bien dans la colonne d'eau en pleine mer, qu'au niveau des environnements côtiers. Les bactéries marines y jouent un rôle primordial (Azam and Malfatti 2007; Teeling *et al.* 2016). En pleine mer, la dégradation des polymères algaux s'accompagne de la libération de matière organique particulaire (POM) et dissoute (DOM) qui s'agrègent en structures désignées sous le terme de « neige marine » (Figure 27). Au cours de son enfoncement inexorable vers les fonds pélagiques, cette matière organique va progressivement être digérée par différentes populations microbiennes qui vont à leur tour alimenter le réseau trophique *via* les protozoaires puis le zooplancton et les poissons. Les éléments les plus récalcitrants constitueront la fraction du carbone atmosphérique réellement piégée hors du cycle du carbone, qui sera stockée dans les fonds marins pour des millions d'années, et qui évoluera vers la formation de carbone fossile (pétrole et gaz). Au

niveau côtier, l'interaction algues-bactéries est aussi très importante, bien qu'encore moins connue qu'au niveau océanique. Les macroalgues offrent des surfaces propices à la colonisation par des communautés bactériennes (Egan *et al.* 2013, 2014). Celles-ci se développent le plus souvent au sein de biofilms et entretiennent des relations étroites avec leurs hôtes, incluant la synthèse de molécules influençant la morphogénèse de l'algue, ou sa protection contre des bactéries pathogènes. Cependant, en cas d'altération de l'intégrité de l'hôte, ces mêmes bactéries peuvent se révéler opportunistes et accélérer le processus de dégradation de l'algue. De façon plus générale, les bactéries marines contribuent de façon prépondérante au recyclage du carbone fixé par les algues, en synthétisant une grande diversité d'enzymes adaptées à la diversité des polysaccharides algaux.

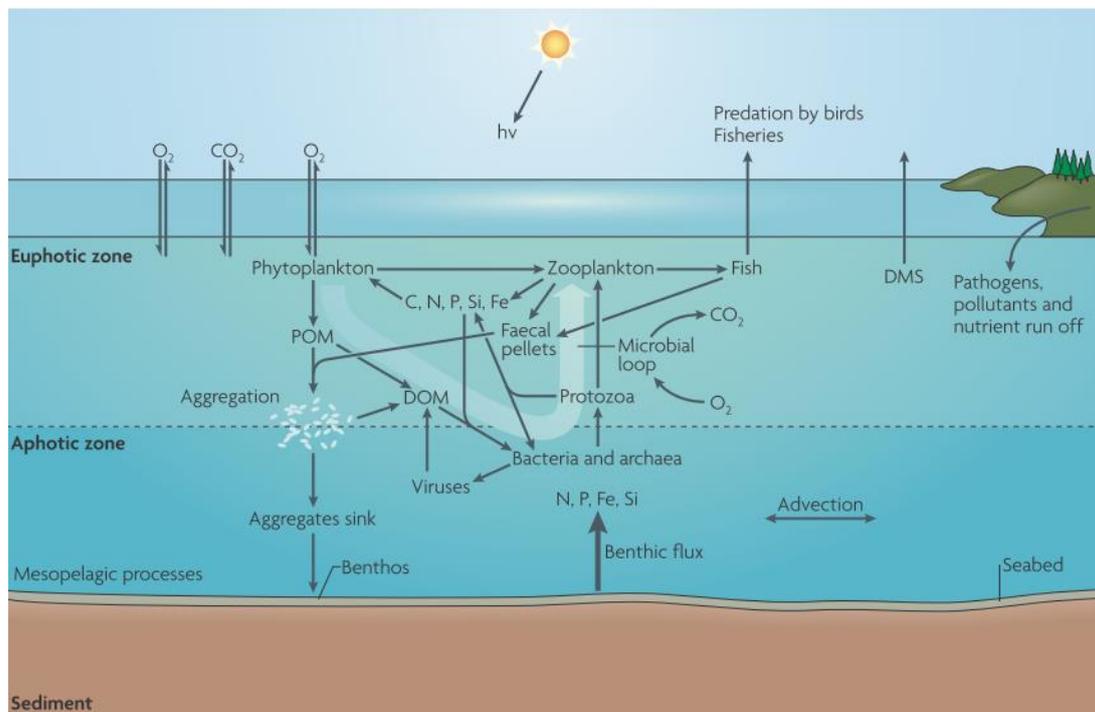


Figure 28 : Structure du réseau microbien dans l'écosystème marin (Azam and Malfatti 2007). La matière organique fixée par le phytoplancton est en grande partie convertie en matière organique particulaire (POM) et dissoute (DOM) qui est ensuite consommée par la communauté microbienne.

Cependant, la diversité des outils enzymatiques à leur disposition varie fortement selon les bactéries. En effet, les travaux de Teeling *et al.* sur les communautés bactériennes se développant dans le contexte des blooms phytoplanctoniques, ont montré qu'il existe une succession temporelle des espèces bactériennes elle-même corrélée à la richesse de ces bactéries en protéines du catabolisme des glucides (Teeling *et al.* 2012, 2016). Ainsi, les bactéries du phylum des *Bacteroidetes* sont les premières à croître et possèdent la plus grande diversité de

CAZymes et de transporteurs transmembranaires dédiés aux polysaccharides. Les *Proteobacteria*, qui arrivent dans un second temps, possèdent un répertoire plus restreint. Elles tirent donc profit de l'action des enzymes des *Bacteroidetes* qui ont préalablement réduit la complexité des substrats en les rendant accessibles à des populations bactériennes qui n'auraient pas pu les exploiter sans leur action préalable.

En ce qui concerne l'écologie bactérienne des macroalgues, il a été montré une variabilité importante des espèces bactériennes en fonction du phylum algal (*Rhodophyta*, *Chlorophyta*, *Phaeophyta*) mais une faible variabilité à l'échelle des phyla bactériens, avec une prédominance des *Proteobacteria* et des *Bacteroidetes*, tout comme observé au niveau des communautés bactériennes associées au phytoplancton (Martin *et al.* 2014). La récente comparaison des génomes de 126 bactéries marines hétérotrophes (BMH) par Barbeyron *et al.* a permis de confirmer les observations de Teeling *et al.* sur les différences de répertoire enzymatique entre les différents genres bactériens, mais également en fonction du type d'habitat (Barbeyron *et al.* 2016). Ainsi, chez les bactéries associées aux macroalgues analysées dans cette étude, le phylum des *Bacteroidetes* arrive toujours en tête en termes de diversité et d'abondance en CAZymes (en moyenne : 118 GH, 15 PL, 50 S), suivi des *Planctomycetes* (en moyenne : 66 GH, 1 PL, 19 S) et enfin des *Proteobacteria* (en moyenne : 25 GH, 1 PL, 6 S).

La diversité des enzymes associées à la dégradation des polysaccharides algaux est donc liée, d'une part à la niche écologique (macroalgues, grands fonds marins, etc.), mais également au phylum algal auquel sont associées les bactéries, reflétant ainsi la co-évolution de la flore bactérienne marine en réponse à la complexité et à l'accessibilité de la source de carbone.

2.2 Stratégies de dégradation : les PULs marins

Bien que les génomes d'origine terrestre soient majoritaires parmi les organismes séquencés, le développement des techniques à haut débit a généré un accroissement rapide du nombre de génomes de bactéries marines séquencés, facilitant ainsi l'étude des voies de dégradations des polysaccharides algaux. En effet, parallèlement à ce qui a été observé chez les bactéries terrestres, les bactéries marines rassemblent souvent l'ensemble des gènes codant pour la dégradation d'un type de polysaccharide au sein de groupes de gènes ('clusters') co-régulés par un même facteur de transcription. Chez les *Bacteroidetes*, ces clusters sont désignés par le terme de PUL (*Polysaccharide Utilization Locus*), et sont de type opéron, c'est-à-dire qu'un

seul ARNm polycistronique est synthétisé et code pour l'ensemble des gènes du PUL (Figure 28).

Initialement mis en évidence chez la bactérie intestinale humaine *Bacteroidetes thetaiotaomicron* dans le contexte de la dégradation de l'amidon (Sus pour *Starch Utilizing System*) (Anderson and Salyers 1989a, 1989b; D'Elia and Salyers 1996; Bjursell *et al.* 2006), le paradigme des PULs a ensuite été élargi à l'ensemble des substrats polysaccharidiques pouvant être dégradés par les *Bacteroidetes*, quel que soit le milieu terrestre ou marin (Thomas *et al.* 2011; Grondin *et al.* 2017). Au niveau marin, différents PULs spécialisés dans la dégradation de l'alginate (Thomas *et al.* 2012), du porphyrane (Hehemann *et al.* 2012b), de l'agar (Hehemann *et al.* 2012a), ou de la laminarine (Kabisch *et al.* 2014) ont été décrits à ce jour.

Plusieurs critères permettent d'identifier des structures de type PUL au sein des génomes des *Bacteroidetes*, et servent de base à la prédiction de leur présence, notamment par l'outil PULDB (*Polysaccharide-Utilization Loci DataBase*) développé par le site CAZy (Terrapon *et al.* 2015). Le premier indicateur est la présence d'un couple de gènes codant pour des protéines de type SusC/SusD, impliquées dans le transport (SusC : protéine canal) et la reconnaissance (SusD) d'oligosaccharides. Un régulateur transcriptionnel et plusieurs CAZymes sont également codés au sein d'un PUL, tous impliqués dans la voie de dégradation d'un même substrat et permettent de prédire la localisation des PULs et de pouvoir ainsi caractériser les voies complètes de dégradation de polysaccharides (Larsbrink *et al.* 2014; Thomas *et al.* 2012; Hehemann *et al.* 2012a).

Cependant, avec l'augmentation du nombre de PULs identifiés, de nouvelles variantes viennent élargir la définition canonique du PUL, écornant au passage le caractère nécessaire du couple SusC/SusD. En effet, si des cas d'absence de l'un des membres ou la distance intergénique importante entre les gènes du couple *susC/susD* ont déjà été documentés chez les bactéries intestinales (Terrapon *et al.* 2015), la description récente d'un PUL carraghénolytique chez la bactérie marine *Zobellia galactanivorans* (Ficko-Blean *et al.*, sous presse) est un des rares cas décrits (Xie *et al.* 2007) de translocation du couple *susC/susD* dans une région génomique bien en aval du PUL.

Le concept de PUL, bien qu'ayant été découvert et défini dans la limite du genre des *Bacteroidetes* pourrait cependant être élargi à d'autres genres bactériens présentant des regroupements similaires de gènes impliqués dans le catabolisme des glucides.

En effet, les SusC-like protéines appartiennent à la famille des protéines membranaires TonB dépendantes (TBD) qui regroupent les TBDRs (*TonB-Dependent Receptors*) et les TBDTs

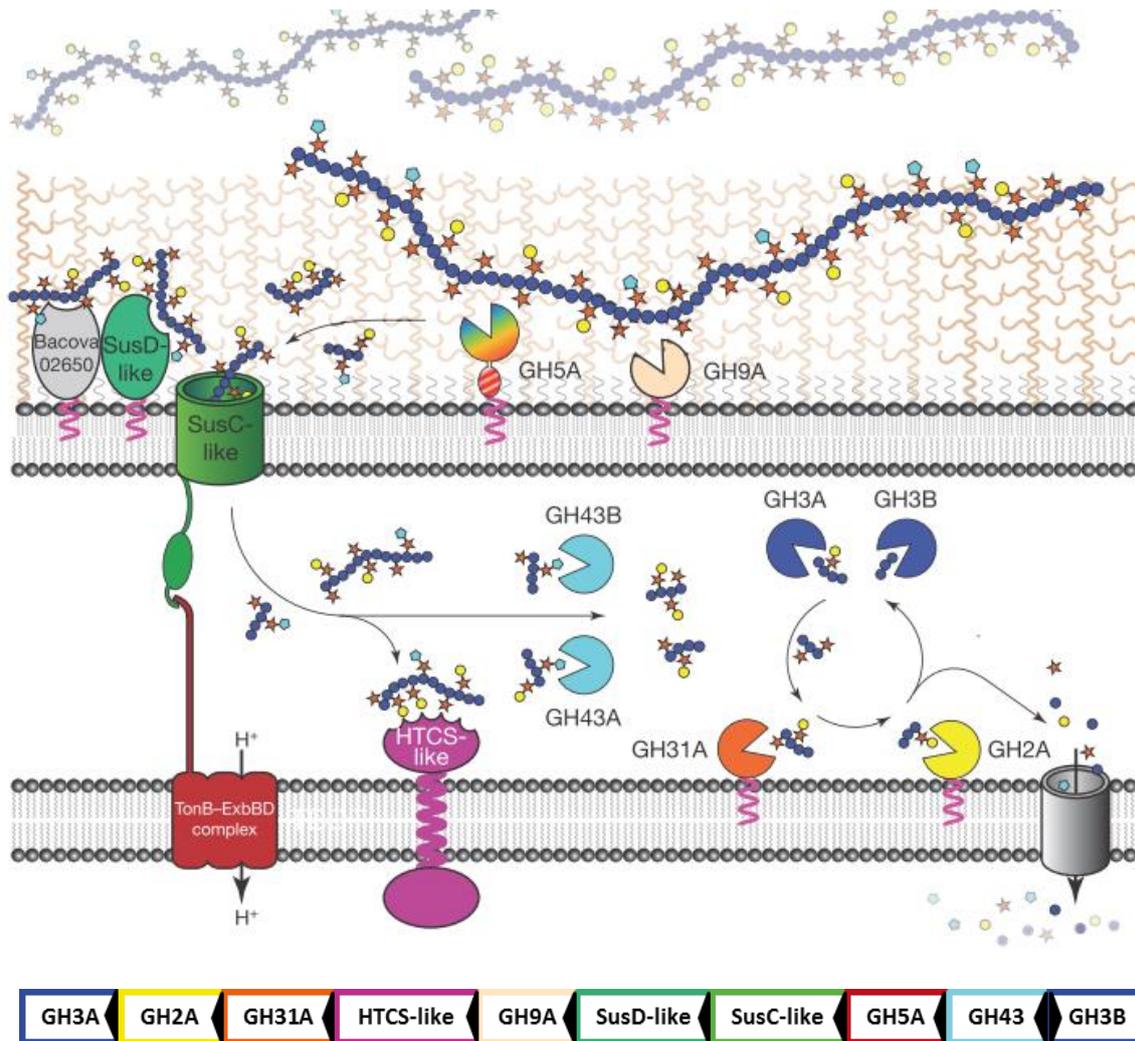


Figure 29 : Schéma de la localisation des différentes protéines impliquées dans la dégradation du xyloglucan chez *B. thetaiotaomicron* (Larsbrink *et al.* 2014). Les gènes du PUL sont entourés avec le code couleur de leur protéine correspondante sur le schéma. Les flèches noires indiquent le sens du cadre de lecture.

(pour *TonB-dependent transducers* si elles possèdent une extension C-terminale impliquée dans une cascade d'interactions protéiques aboutissant à l'induction de gènes). Les protéines TBD sont largement répandues au sein des bactéries à Gram négatif. Initialement identifiées pour leur rôle dans l'assimilation du fer (Hartney *et al.* 2011), leur implication dans l'assimilation des glucides est plus récente.

Il a été mis en évidence chez des Proteobacteria un nombre élevé de gènes codant des protéines TBD, dont certains co-localisent avec des gènes codant des CAZymes et parfois même avec un régulateur transcriptionnel. Ainsi, le terme de CUT (pour 'carbohydrate utilization locus containing TBDDT') a été proposé pour désigner cet équivalent de PUL mais chez des bactéries n'appartenant pas aux Bacteroidetes (Blanvillain *et al.* 2007).

De façon analogue, il a été montré très récemment que des bactéries à Gram positif du genre Firmicutes ont également développé des stratégies de regroupements de gènes impliqués dans le catabolisme des glucides, baptisés gpPULs (pour ‘Gram-positives PULs’) (Sheridan *et al.* 2016).

Les structures de type PUL représentent donc des stratégies évolutives aux spécificités variables selon les genres bactériens, mais reflétant surtout une capacité d’adaptation des bactéries aux différentes niches écologiques présentes sur terre. Avec l’augmentation croissante des données de séquençage, de nombreux travaux ont mis en évidence l’importance des phénomènes de transferts latéraux de gènes entre bactéries partageant des niches écologiques semblables. Les cas de conservation de gènes de PULs, voire de PULs entiers, en synténie entre des genres bactériens différents sont fréquemment observés, et contribuent à améliorer la compréhension des voies de dégradation des glucides chez de nouvelles espèces. Des phénomènes de remaniement génétique (‘reshuffling’) sont aussi fréquemment observés, complexifiant alors la tâche d’identification d’une voie complète de dégradation. Mais, au même titre que la présence d’éléments génétiquement mobiles (transposons, plasmides) spécialisés dans la dégradation de glucides ou la conservation synténique de PULs entre espèces différentes, ils témoignent de la dynamique évolutive qui sous-tend les processus d’acquisition et de dégradation des glucides par les bactéries (Martens *et al.* 2014).

2.3 Etat de l’art sur les GH de bactéries marines actives sur polysaccharides algaux

Plutôt que d’énumérer sous forme d’un catalogue rébarbatif l’état des connaissances en matière de GH marines, la présentation synthétique sous forme de tableau présentant le type de substrat, la famille de GH et le niveau de caractérisation (protéine native, clonée, caractérisation biochimique et/ou structurale), ainsi que les références bibliographiques majeures associées m’a paru être une solution plus appréciable (Tableau 3). J’y ai dressé la liste des enzymes actives sur polysaccharides algaux caractérisées d’un point de vue biochimique et/ou structural, d’après la liste disponible dans la base de données CAZy. Seules les souches d’origine marine (bactéries et archées) sont prises en compte, ainsi que les souches de bactéries intestinales humaines pour lesquelles des cas de transferts latéraux de gènes d’origine marine ont été mis en évidence (*Bacteroidetes thetaiotaomicron*, *Bacteroides plebeius*). Etant donné que ma thèse s’est focalisée sur les enzymes de type GH, ce sont elles qui font l’objet de cet inventaire bibliographique. Cependant j’ai tenu à ajouter quelques références pour les familles de

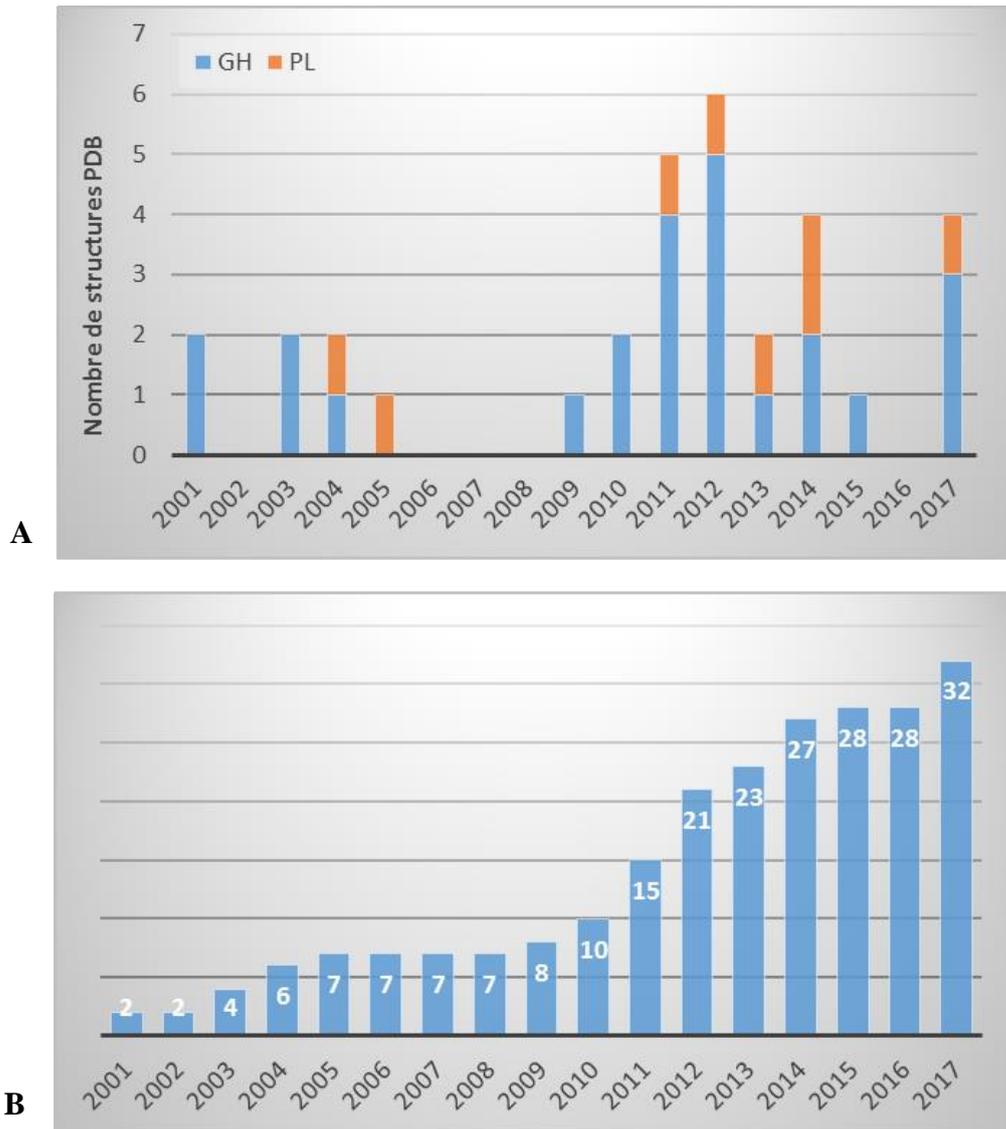


Figure 30 : Evolution du nombre de structures de CAZymes marines depuis 2001. A : En bleu les GH et en orange les PL ; B : total cumulé

	GH marines			CAZy					
	caractérisées C	structure S	% S/C	déposées D	caractérisées C	% C/D	structure S	% S/D	% S/C
2008	40	5	12,5	103111	9691	9,4	801	0,8	8,3
2013	80	18	22,5	337143	12730	3,8	1419	0,4	11,1
2017	100	24	24,0						

Tableau 2 : Statistiques calculées à partir de la base de données CAZy comparant le pourcentage de structures PDB disponibles en fonction du nombre de séquences protéiques déposées dans les bases de données officielles et ayant fait l'objet d'une caractérisation biochimique et/ou structurale.

polysaccharides lyases actives sur les ulvanes et la laminarine. Pour les PL, seules les références structurales sont fournies en détail. Le nombre d'enzymes caractérisées d'un point de vue biochimique est toutefois précisé à titre indicatif pour chaque famille.

A l'heure de l'écriture de ces lignes, 100 GH et 53 PL de bactéries actives sur polysaccharides algaux sont caractérisées mais à des degrés divers. Une seule a fait l'objet d'études uniquement sur enzyme native (β -agarase de *Pseudoalteromonas fortis*), toutes les autres ont été clonées et caractérisées au niveau biochimique et/ou structural (4 structures ont été déposées à la PDB sans publication associée caractérisant l'activité biochimique associée). Cependant le nombre de structures d'origine marine déposées dans la PDB (protein data bank), bien qu'en nette

augmentation depuis les années 2010, est encore relativement faible : 24 GH et 8 PL, soit 21,6 % de structures résolues par rapport au nombre d'enzymes caractérisées (Figure 29). Pour comparaison, les statistiques de la base de données CAZy, enzymes et CBM confondus, montrent qu'en 2008 seules 8,3 % des séquences caractérisées possédaient une structure déposées dans la PDB, et ce chiffre passait à 11,1 % en 2013 (Lombard *et al.* 2014). Pour les GH marines, le pourcentage de résolution structurale est passé de 22,5 % en 2013 à 24 % en 2017 (Tableau 2). Cela indique une couverture structurale des GH marines plus de deux fois supérieure à la moyenne des CAZymes et modules associés.

Grâce à l'augmentation du nombre de génomes de bactéries marines séquencés, de nombreuses fonctions enzymatiques sont prédites par alignement des séquences. Cela permet d'une part d'augmenter le peuplement des familles de GH existantes et de créer de nouvelles familles, parfois propres au milieu marin (ex. : GH82, GH117, GH118).

Enfin cela permet d'étudier l'histoire évolutive des voies de dégradation des glucides *via* les nombreux événements remaniement (duplication, transposition) et de transferts génétiques (plasmides, conservation synténique de clusters) dont les génomes conservent l'empreinte (Thomas *et al.* 2011; Kabisch *et al.* 2014; Zhu *et al.* 2017).

Cependant, la proportion d'enzymes caractérisées au niveau biochimique reste très inférieure au nombre prédit grâce à l'homologie de séquence (Tableau 2). En effet, la caractérisation biochimique représente un goulet d'étranglement de par le manque de techniques permettant une approche à moyen et haut débit. Néanmoins, le développement d'approches de miniaturisation via micro-puces ComPP (pour *Comprehensive microarray Polymer Profiling*) par exemple représente une voie prometteuse pour le criblage à haut débit de substrats

polysaccharidiques, mais qui reste limité pour le moment à des protéines non enzymatiques, de type CBM, SusD-like ou anticorps monoclonaux (Moller *et al.* 2007; Salmeán *et al.* 2017).

Pourtant, la compréhension des mécanismes enzymatiques et des interactions structure/fonction entre les enzymes marines et leurs substrats algaux revêt un intérêt tout particulier en termes de transfert de connaissances vers la santé humaine. En effet, comme évoqué au paragraphe II de cette introduction, de nombreux polysaccharides algaux présentent la particularité d'être sulfatés, tout comme la majorité des GAGs de la matrice extracellulaire (MEC) animale. Or de nombreuses pathologies humaines impliquent une altération de l'homéostasie de la MEC (blessures, arthrite rhumatoïde, cancers,...) souvent associée à une activité anormale des enzymes de remodelage de la MEC (Frantz *et al.* 2010). Les polysaccharides marins sulfatés représentent donc de bons modèles pour l'étude théorique des interactions enzymes/glucides, mais également des biomimétiques de choix qui, grâce à une meilleure connaissance des enzymes de modification (GH, PL, CE, S) pourraient participer à la médecine de demain (Senni *et al.* 2006, 2011).

Pages suivantes :

Tableau 3 : Inventaire des enzymes actives sur polysaccharides algaux référencées dans la base de données CAZy. GH : glycosides hydrolase, PL : polysaccharide lyase. Les noms de bactéries non marines mais dont les enzymes sont issues d'un transfert horizontal de gènes provenant de bactéries marines sont colorées en orange. Les noms des bactéries dont certaines enzymes ont fait l'objet d'étude durant cette thèse sont indiqués en gras. Dans la colonne des références, celles correspondant à des dépôts dans la PDB sans publication associées à une caractérisation biochimiques sont annotées (**PDB**).

NB1 : La liste des références bibliographiques associées au tableau 3 est séparée de la liste bibliographique principale, et s'intitule « Bibliographie II ».

NB2 : l'activité fucosidase de *T. maritima* n'a jamais été démontrée sur polysaccharide algaux. Cependant, la capacité de cette bactérie vivant dans les grands fonds marins à dégrader différentes sources polysaccharidiques (glucomannane, xyloglucane, cellulose, ...) (Chhabra *et al.* 2002) qui se retrouvent également dans la paroi des algues, nous a amené à l'inclure dans ce tableau.

Substrat	Famille	Protéine	Organisme	Genbank	PDB	Références	
Carraghénases							
β	GH16		<i>Paraglaciecola hydrolitica</i> S66			Schultz-Johansen <i>et al.</i> , in prep	
		GH42-like		<i>Paraglaciecola hydrolitica</i> S66			Schultz-Johansen <i>et al.</i> , in prep
				<i>Pseudoalteromonas carrageenovora</i> 9^T ATCC 43555			McLean and Williamson 1981 Nat Matard-Mann <i>et al.</i> , in prep
κ	GH16	PcCgkA	<i>Pseudoalteromonas carrageenovora</i> 9 ^T ATCC 43555	CAA50624.1	1DYP 5OCQ	Barbeyron <i>et al.</i> 1994; Michel <i>et al.</i> 2001a Matard-Mann <i>et al.</i> , in press	
		ZgCgkA	<i>Zobellia galactanivorans</i> Dsij	CAZ94309.1	5OCR	Potin <i>et al.</i> 1991; Barbeyron <i>et al.</i> 1998 Matard-Mann <i>et al.</i> , in press	
			<i>Pseudoalteromonas porphyrae</i> LL1	ADD92366.1		Liu <i>et al.</i> 2011	
			<i>Zobellia</i> sp. ZM-2			Liu <i>et al.</i> 2013	
		Cgk-K142a	<i>Pseudoalteromonas tetraodonis</i> JAM-K142	BAJ61957.1		Kobayashi <i>et al.</i> 2012	
		Cgk-K142b		BAJ61958.1			
		CgkS	<i>Shewanella</i> sp. P1-14-1 Kz7	AHN05534.1		Wang <i>et al.</i> 2015	
		CgkX	<i>Pseudoalteromonas</i> sp. QY203	AFV39914.1		Xu <i>et al.</i> 2015	
ι	GH82	AfCgiA	" <i>Alteromonas fortis</i> " ATCC 43554	CAC07801.1	1H80 1KTW 3LMW	Barbeyron <i>et al.</i> 2000; Michel <i>et al.</i> 2001b, 2003; Rebuffet <i>et al.</i> 2010	
		ZgCgiA1	<i>Zobellia galactanivorans</i> Dsij	CAC07822.1		Barbeyron <i>et al.</i> 2000	
		ZgCgiA2		CAZ96312.1		Rebuffet <i>et al.</i> 2010	
		ZgCgiA3		CBW46642.1			
		CgiF	<i>Flavobacterium</i> sp. YS-80-122	AEV89930.1		Li <i>et al.</i> 2017	
		CgiA	<i>Microbulbifer thermotolerans</i> JAMB-A94	BAJ40863.1		Hatada <i>et al.</i> 2011	
		CgiA_Ce	<i>Cellulophaga</i> sp. QY3	AEV89930.1		Ma <i>et al.</i> 2013	
		CgiB_Ce		AGN70890.1			
λ	GHnc	CglA	<i>Pseudoalteromonas</i> sp. CL19	BAF35571.1		Ohta and Hatada 2006	
		PcCglA	<i>Pseudoalteromonas carrageenovora</i> 9T ATCC 43555	CAL37005.1		Guibet <i>et al.</i> 2007	

Activité enzymatique	Famille	Protéine	Organisme	Genbank	PDB	Références	
Agarases							
α	GH96	AaAgaA	<i>Alteromonas agarlytica</i> DSM 12513	AAF26838.1		Potin <i>et al.</i> 1993; Flament <i>et al.</i> 2007	
		Aga33	<i>Thalassomonas=Thalassotalea agarivorans</i> JAMB A33	BAF44076.1		Hatada <i>et al.</i> 2006	
		ZgAgaA	<i>Zobellia galactanivorans</i> Dsij	CAZ98338.1	104Y 1URX	Allouch <i>et al.</i> 2003, 2004; Jam <i>et al.</i> 2005	
ZgAgaB	AAF21821.1	104Z 4ATF		Allouch <i>et al.</i> 2003; Jam <i>et al.</i> 2005 Hehemann <i>et al.</i> 2012a			
ZgAgaD	CAZ98378.1	4ASM		Hehemann <i>et al.</i> 2012a			
β	GH16	AagA	<i>Pseudoalteromonas gracilis</i> B9	AAF03246.1		Schroeder <i>et al.</i> 2003	
			<i>Pseudoalteromonas</i> sp. AG4	ADD60418.1		Oh <i>et al.</i> 2010b	
			<i>Pseudoalteromonas atlantica</i> ATCC 19262	AAA91888.1		Morrice <i>et al.</i> 1983Nat	
		AgaH71	<i>Pseudoalteromonas hodoensis</i> H7	AIF29515.1		Park <i>et al.</i> 2014b	
		AgaA	<i>Pseudoalteromonas</i> sp. AG52	ADA72015.1		Oh <i>et al.</i> 2010a	
		AgaA	<i>Pseudoalteromonas</i> sp. CY24	AAN39119.2		Lu <i>et al.</i> 2009	
		AgaA6	<i>Pseudoalteromonas</i> sp. JT-6	ACF40223.1		Tao <i>et al.</i> 2008	
		AgaYT	<i>Flammeovirga yaeyamensis</i> YT	AEK80424.1		Yang <i>et al.</i> 2011	
		AgaG1	<i>Alteromonas</i> sp. GNUM-1	AGW43026.1		Kin and Hong 2012	
		BuGH16Bwt	<i>Bacteroides uniformis</i> NP1			5T9X	Pluvignage <i>et al.</i> 2017 PDB
		BpGH16A	<i>Bacteroides plebeius</i> DSM 17135	EDY95404.1			Hehemann <i>et al.</i> 2012b
		MtAgaA	<i>Microbulbifer thermotolerans</i> JAMB-A94	BAD29947.1		3WZ1	Takagi <i>et al.</i> 2015
		AgaA7	<i>Microbulbifer elongatus</i> JAMB-A7	BAC99022.1			Ohta <i>et al.</i> 2004b
		Aga16B	<i>Saccharophagus degradans</i> 2-40	ABD80437.1			Ekborg <i>et al.</i> 2006
		AgaY1	<i>Saccharophagus</i> sp. AG21	AFR90184.1			Lee <i>et al.</i> 2013b
		AgaB34	<i>Agarivorans albus</i> YKW-34	ABW77762.1			Fu <i>et al.</i> 2009
		AgaJA3	<i>Agarivorans</i> sp. JA-1 / HZ105	ADY17921.1			Jeon <i>et al.</i> 2012
		AgaA	<i>Agarivorans</i> sp. LQ48	ACM50513.1			Long <i>et al.</i> 2010
		AgaG1	<i>Alteromonas</i> sp. GNUM-1	AGW43026.1			Chi <i>et al.</i> 2014
		YM01-3	<i>Catenovulum agarivorans</i> YM01	AGU13985.1			Cui <i>et al.</i> 2014
		AgaB	<i>Flammeovirga</i> sp. SJP92	ANN44251.1			Dong <i>et al.</i> 2016
		AgaG4	<i>Flammeovirga yaeyamensis</i> MY04	ACY02061.1			Han <i>et al.</i> 2013
		AgaA, D	<i>Vibrio</i> sp. PO-303				Dong <i>et al.</i> 2007b, 2007a
		AgaV	<i>Vibrio</i> sp. V134				Zhang and Sun 2007
		AgaY	<i>Janthinobacterium</i> sp. SY12	ABO93616.2			Shi <i>et al.</i> 2008

Activité enzymatique	Famille	Protéine	Organisme	Genbank	PDB	Références		
β	Agarases							
	GH50	AgaD Aga50D	<i>Saccharophagus degradans</i> 2-40	ABD81904.1	4BQ2-5	Pluvinage <i>et al.</i> 2013		
		AgaA Aga50A		ABD80438.1		Ekborg <i>et al.</i> 2006		
		AgaA34	<i>Agarivorans albus</i> YKW-34	P85974.1		Fu <i>et al.</i> 2008		
		AgWH50A	<i>Agarivorans gilvus</i> WH0801	AFP32918.1		Liu <i>et al.</i> 2014a		
		AgWH50C		AHM94172.1		Liu <i>et al.</i> 2014b		
		AgrA	<i>Agarivorans</i> sp. AG17	ACU52709.1		Nikapitiya <i>et al.</i> 2010		
		Hz1	<i>Agarivorans</i> sp. HZ105	ADY17918.1		Hu <i>et al.</i> 2009		
		Hz2		ADY17919.1		Lin <i>et al.</i> 2012		
			<i>Agarivorans</i> sp. JA-1	ABK97391.1		Lee <i>et al.</i> 2006		
			<i>Agarivorans</i> sp. JAMB-AII	BAD99519.1		Ohta <i>et al.</i> 2005		
			<i>Agarivorans</i> sp. QM38	ABM90422.1		Du <i>et al.</i> 2011		
			<i>Alteromonas</i> sp. E-1	BAE97587.1		Kirimura <i>et al.</i> 1999		
			<i>Pseudoalteromonas</i> sp. NJ21	AHC72907.1		Li <i>et al.</i> 2015		
			<i>Saccharophagus</i> sp. AG21	AFZ95003.1		Lee <i>et al.</i> 2013a		
			<i>Thalassotalea=Thalassomoas agarivorans</i> BCRC 17492	AGT98631.1		Liang <i>et al.</i> 2014		
			<i>Vibrio</i> sp. CN41	ADM25828.1		Liao <i>et al.</i> 2011		
			AgaA	<i>Vibrio</i> sp. JT0107	BAA03541.1		Sugano <i>et al.</i> 1993	
			AgaB		BAA04744.1		Sugano <i>et al.</i> 1994	
			AgaB	<i>Vibrio</i> sp. PO-303	BAG71427.1		Araki <i>et al.</i> 1998	
			GH86	AgrA	<i>Pseudoalteromonas atlantica</i> T6c	ABG40858.1		Belas 1989
				AgaO	<i>Microbulbifer thermotolerans</i> JAMB-A94	BAD86832.1		Ohta <i>et al.</i> 2004a
				AgaC	<i>Saccharophagus degradans</i> 2-40	ABD81910.1		Ekborg <i>et al.</i> 2006
				AgaE		ABD81915.1		
			GH118	AgaC	<i>Vibrio</i> sp. PO-303	BAF03590.1		Dong <i>et al.</i> 2006
				AgaB	<i>Pseudoalteromonas</i> sp. CY24	AAQ56237.1		Ma <i>et al.</i> 2007
				AgaXa	<i>Catenovulum</i> sp. X3	ADF46235.1		Xie <i>et al.</i> 2012

Activité enzymatique	Famille	Protéine	Organisme	Genbank	PDB	Références	
Porphyranases							
β	GH16	ZgPorA	<i>Zobellia galactanivorans</i> Dsij	CBM41182.1	3ILF 4ATE	Hehemann <i>et al.</i> 2010	
		ZgPorB		CAZ95074.1	3JUJ		
		AagA	<i>Pseudomonas</i> sp. ND137	BAB79291.1		Aoki and Kamei 2006	
	GH86	BpGH16B	<i>Bacteroides plebeius</i> DSM 17135	EDY95423.1	4AWD	Hehemann <i>et al.</i> 2012b	
		BpGH86A		EDY95427.1	4AW7		
3,6-anhydro-L-galactosidases							
	GH117	ZgAhgA	<i>Zobellia galactanivorans</i> Dsij	CAZ98798.1	3P2N	Rebuffet <i>et al.</i> 2011	
		Zg3615		CAZ97753.1	4U6D	Ficko-Blean <i>et al.</i> Sous presse	
		SdNABH	<i>Saccharophagus degradans</i> 2-40	ABD81917.1	3R4Y,Z	Ha <i>et al.</i> 2011	
		BpGH117	<i>Bacteroides plebeius</i> DSM 17135	EDY95405.1	4AK5-7	Hehemann <i>et al.</i> 2012b	
Ulvane lyases							
Ulvane-lyase	PLnc		<i>Nonlabens ulvanivorans</i>	AEN28574.1		Nyvall Collèn <i>et al.</i> 2011	
unsaturated β -glucuronyl hydrolase	GH105	Nu_GH105		AFQ98272.1	4CE7	Nyvall Collèn <i>et al.</i> 2014	
Alginate lyase							
	PL6	AlyGC	<i>Paraglaciicola chathamensis</i> S18K6T	WP_007984897.1	5GKD,Q	Xu <i>et al.</i> 2017	
		+16 autres					
	PL7	ZgAlyA1	<i>Zobellia galactanivorans</i> Dsij	CAZ95239.1	3ZPY	Thomas <i>et al.</i> 2012, 2013	
		ZgAlyA5		CAZ98266.1	4BE3		
		alyPG	<i>Corynebacterium</i> sp. ALY-1	BAA83339.1	1UAI	Osawa <i>et al.</i> 2005	
		+20 autres					
	PL14	ZgAlyA7	<i>Zobellia galactanivorans</i> Dsij	CAZ98462.1		Thomas <i>et al.</i> 2012	
	PL15	OalA	<i>Vibrio splendidus</i> 12B01	EAP93067.1		Badur <i>et al.</i> 2015	
	PL17		<i>Saccharophagus degradans</i>	ABD82539.1	4NEI 4OJZ 4OK2,4	Park <i>et al.</i> 2014a	
		+4 autres					
	PL18	Aly-SJ02	<i>Pseudoalteromonas</i> sp. 272			1J1T	Motoshima <i>et al.</i> 2004 PDB
			<i>Pseudoalteromonas</i> sp. SM0524	ACB87607.2		4Q8K,L	Li <i>et al.</i> 2011; Dong <i>et al.</i> 2014
		+2 autres					

Activité enzymatique	Famille	Protéine	Organisme	Genbank	PDB	Références
Fucoïdanases						
Fucanase	GH107	MfFcnA	<i>Marineflexile fucanivorans</i> SW5	CAI47003.1		Colin <i>et al.</i> 2006
Fucosidase	GH29	TmaFuc	<i>Thermotoga maritima</i> MSB8	AAD35394.1	1HL8,9 1ODU 2WSP 2ZWY,Z 2ZX5-9 2ZXA,B,D	Sulzenbacher <i>et al.</i> 2004 Cobucci-Ponzano <i>et al.</i> 2009 Wu <i>et al.</i> 2009 PDB Wu <i>et al.</i> 2010
			<i>Zobellia galactanivorans</i> Dsij			Matard-Mann <i>et al.</i> , in prep
Laminarinases						
	GH16	ZgLamA	<i>Zobellia galactanivorans</i> Dsij	CAZ96583.1	4BOW 4BPZ 4BQ1	Labourel <i>et al.</i> 2014
		TmLam Lam16	<i>Thermotoga maritima</i> MSB8	AAD35118.1	3AZX,Z	Jeng <i>et al.</i> 2011
		Lam16A	<i>Thermotoga neapolitana</i>	CAA88008.1		Zverlov <i>et al.</i> 1997
		LamR	<i>Rhodothermus marinus</i>	AAC69707.1	3ILN	Bleicher <i>et al.</i> 2011
		pfLamA	<i>Pyrococcus furiosus</i> Archaea	AAC25554.2	2VY0	Ilari <i>et al.</i> 2009
		TpLam	<i>Thermotoga petrophila</i> RKU-1	ABQ46917.1	4DFS	Meza <i>et al.</i> 2012 PDB
			<i>Mycobacterium marinum</i> M Archae	ACC41342.1	4PQ9	Dranow <i>et al.</i> 2014 PDB
		Fjoh_4177	<i>Flavobacterium johnsoniae</i> UW101	ABQ07185.2		McBride <i>et al.</i> 2009
		Orf00540	<i>Formosa algae</i> KMM 3553	AJF35005.1		Kusaykin <i>et al.</i> 2017
		GluB	<i>Pseudomonas</i> sp. PE2	BAC16331.1		Kitamura and Kamei 2006
		GluA		BAC16332.1		

3. Applications des CAZymes en biotechnologie bleue

Au-delà de l'intérêt fondamental à étudier la diversité, les modes d'action et les liens évolutifs des CAZymes marines, elles représentent également de formidables outils d'investigation à des fins de recherche fondamentale mais également dans des perspectives d'application en santé et en nutrition humaine et animale.

3.1 Utilisations en recherche fondamentale

Un des premiers usages des CAZymes de bactéries marines (souvent extraites de glandes digestives de brouteurs marins comme les ormeaux) a été de participer à l'obtention de protoplastes algaux, afin d'étudier les mécanismes de mise en place de la paroi des algues, mais également à des fins de micropropagation et d'amélioration génétique (Boyen *et al.* 1990; Zablackis *et al.* 2009; Bodian *et al.* 2013).

Les CAZymes sont également utilisées pour étudier la composition et la structure chimique détaillée des polysaccharides algaux. La digestion enzymatique des polysaccharides est couplée à différents types d'analyses permettant de caractériser avec précision la structure chimique des oligosaccharides générés : HPLC, spectrométrie de masse, RMN. Une approche de type 'polysaccharide sequencing' a ainsi été menée avec succès sur des oligo-porphyranes grâce au couplage de la chromatographie à haute résolution (UHPLC) avec de la spectrométrie de masse en tandem avec activation XUV (Ropartz *et al.* 2016).

Elles servent aussi à étudier la mise en place des composés pariétaux dans les stades précoces de développement cellulaire algal. Pour cela, la digestion sélective de polysaccharides pariétaux à l'aide d'enzymes dédiées permet de générer la production d'anticorps monoclonaux dirigés contre les produits de digestion enzymatiques. Ces anticorps, couplés à des anticorps secondaires fluorescents, sont ensuite utilisés comme sonde pariétales afin de visualiser la localisation de différents épitopes lors de la mise en place de la paroi (Torode *et al.* 2016).

L'étude des mécanismes de défense chez les macroalgues en réponse au broutage à recours à l'utilisation d'oligoalginates générés par digestion enzymatique (Kupper *et al.* 2001).

Enfin, l'usage des gels d'agarose en biologie moléculaire pour la migration de l'ADN a également vu se développer l'utilisation d'agarases pour l'extraction d'ADN afin d'augmenter les rendements d'extraction finaux (ex. : β -agarase I par NEB, ou Agarase par ThermofischerScientific) (Gold 1992).

3.2 Utilisations en recherche appliquée

Comme on l'a vu au cours de cette introduction, les macroalgues, et en particulier leur paroi, renferment une grande diversité de macromolécules (polysaccharides, protéines, composés volatils,...). Afin d'isoler certains composés d'intérêt comme les polyphénols, les protéines, ou certains polysaccharides (fucoïdanes par exemple) le recours à la dégradation enzymatique des autres polysaccharides algaux représente une solution intéressante (Fleurence *et al.* 1995; Hahn *et al.* 2012). De plus, les oligosaccharides ainsi générés pourront également être valorisés. En effet, de nombreuses études sur les propriétés biologiques des polysaccharides sulfatés ont montré que ce sont les fractions de plus petit poids moléculaire qui présentent une activité maximale. On pourra citer les exemples des propriétés antivirales des oligosaccharides de ι -carraghénane qui vont jouer le rôle de biomimétique des GAGs de la muqueuse et fixer les capsides virales de papillomavirus humains bloquant ainsi l'infection des tissus épithéliaux (Buck *et al.* 2006); mais aussi les propriétés immunostimulantes et/ou anti-tumorales d'oligosaccharides de λ -carraghénane (Zhou *et al.* 2004), de β/κ -carraghénane hybride (Yang *et al.* 2011); ou encore les propriétés antioxydantes d'oligo-porphyrans (Zhao *et al.* 2006). Cependant, du fait de la quasi-absence au niveau commercial de CAZymes actives sur polysaccharides algaux, les études utilisant des enzymes marines pour générer des oligosaccharides sont relativement peu nombreuses (Klarzynski *et al.* 2003; Iwamoto *et al.* 2005; Hatada *et al.* 2006; Fu and Kim 2010; Chauhan and Saxena 2016). D'autres techniques de dépolymérisation comme l'irradiation (micro-ondes, rayons γ) ou la dépolymérisation radicalaire par l' H_2O_2 sont utilisées plus fréquemment (Courtois 2009). Bien que moins agressives que l'hydrolyse acide, ces méthodes présentent le même caractère aléatoire du clivage et une poly-dispersité variable de la taille des fragments générés.

La saccharification des polysaccharides marins (dégradation de polysaccharides complexes libérant des sucres simples) pour la production de bioéthanol est également considérée à l'heure actuelle comme une débouchée innovante (Swain *et al.* 2017). L'utilisation d'enzymes marines se révèle en effet indispensable afin de pouvoir traiter les glucides originaux présents en abondance chez les algues, tels que le 3,6-anhydrogalactose (Kim *et al.* 2012). Le recours à des concepts dérivés de celui des cellulosomes des bactéries terrestres (Smith and Bayer 2013) pourraient représenter une approche intéressante dans ce contexte, mais également de façon plus générale afin de développer les applications industrielles des CAZymes marines (Kang *et al.* 2014).

Cependant, le passage à l'échelle industrielle nécessite de relever plusieurs défis. Tout d'abord, dans le cadre des applications agro-alimentaires et pharmaceutiques, les enzymes doivent être produites dans des organismes classés comme GRAS nécessitant le développement de systèmes d'expression recombinante différents de ceux classiquement utilisés dans les laboratoires de recherche (passage d'*E. coli* à *Pichia pastoris* par exemple) (Ahmad *et al.* 2014). Ensuite, le changement d'échelle entre les conditions de production au laboratoire et les besoins industriels demande un travail de mise au point. Enfin, les hydrolysats enzymatiques doivent être intégrés dans des produits ayant une valeur ajoutée suffisante pour compenser le coût de revient élevé généré par l'utilisation et la production d'enzymes. Actuellement, des programmes de valorisation industrielle tels que Algolife⁴, permettent de poser les premières pierres d'un partenariat français entre des équipes de recherche publiques étudiant ces CAZymes marines et des industriels souhaitant développer leurs innovations dans le domaine de la valorisation de la biomasse algale, via la mise en place de cellules de transfert technologique.

Au terme de cette introduction, il apparaît donc que la connaissance et l'utilisation des enzymes marines sont encore à leurs débuts, mais laissent entrevoir de belles perspectives aussi bien d'un point de vue fondamental qu'appliqué. Tirant profit des avancées technologiques réalisées dans les domaines de la génétique, de la génomique fonctionnelle, de la protéomique et de la métabolomique, les études en cours et à venir promettent des avancées significatives dans la compréhension des mécanismes biochimiques qui sous-tendent les interactions écologiques en milieu marin. C'est dans cette perspective enthousiasmante que ce travail de thèse a été réalisé.

⁴ Algolife est un programme breton de recherche et développement financé par BPI France à hauteur de 10,6 millions d'euros dans le cadre du Programme d'investissements d'avenir. L'objectif est de développer la filière de transformation des macroalgues brunes, rouges et vertes. Les partenaires impliqués dans le projet sont publiques (CNRS, ANSES) et privés (la Sica de Saint-Pol-de-Léon, Spécialités Pet Food, Triballat Noyal et Amadéite).

IV. Problématique scientifique

Grâce au séquençage récent de nombreux génomes de bactéries marines, l'identification de nouvelles enzymes bactériennes actives sur des polysaccharides algaux est désormais possible par génomique comparative. En effet, la comparaison génomique d'un grand nombre d'organismes permet fréquemment de prédire l'appartenance des produits de gènes à différentes familles de CAZymes. Il est aussi possible de prédire la localisation de clusters de gènes potentiellement impliqués dans la dégradation d'un type de polysaccharide précis.

Afin d'améliorer la connaissance des voies du catabolisme des polysaccharides algaux par les bactéries marines, il est donc nécessaire d'étudier en parallèle plusieurs enzymes impliquées dans la dégradation d'un même substrat. De plus, les contingences techniques liées au caractère biologique des enzymes (clonage, expression de protéines recombinantes, cristallogénèse,...) entraînent nécessairement un certain pourcentage de 'perte' tout au long du processus aboutissant à leur caractérisation complète. Ce travail de thèse s'est donc déroulé en gardant à l'esprit ces deux aspects afin d'aborder de façon efficace l'étude des différentes enzymes auxquelles je me suis intéressée.

La première étape, exposée au chapitre 1, a consisté à mettre en place une approche à moyen débit, visant à cloner en parallèle un ensemble de cibles d'intérêt présélectionnées à partir des données génomiques disponibles et pour partie impliquées dans le catabolisme du carraghénane : κ -carraghénase de *Zobellia galactanivorans*, sulfatases et GH du plasmide carraghénolytique de *Pseudoalteromonas carrageenovora*, différentes carraghénases chez *Alteromonas fortis*. Certaines cibles ont également été sélectionnées pour leur originalité de séquence laissant présager de nouvelles activités enzymatiques: GH42-like chez *P. carrageenovora*, famille de GH inconnue dans un PUL fucanolytique chez *Z. galactanivorans*.

Parmi les cibles ayant donné de bons résultats suite aux tests d'expression à petite échelle, quatre GH ont fait l'objet d'une étude plus poussée visant à déterminer leurs caractéristiques biochimiques et structurales : la κ -carraghénase de *Z. galactanivorans* (ZgCgkA), la GH42-like de *P. carrageenovora* (Pcmp0019), une GH29 (Zg3464) et une GHnc (Zg3470) de *Z. galactanivorans*.

Les deux premières, ZgCgkA et Pcmp0019, que nous aborderons respectivement dans les chapitres 2 et 3, sont impliquées dans la dégradation des carraghénanes. Comme nous l'avons vu au paragraphe III.2.3 de cette introduction, la connaissance des enzymes de dégradation du carraghénane est encore parcellaire non seulement au niveau biochimique mais surtout au

niveau structural. L'étude des homologues de celles déjà caractérisées représente donc un enjeu majeur afin de mieux comprendre les spécificités structurales et fonctionnelles des membres d'une même sous-famille de GH. La découverte d'enzymes sans homologues connus est également essentielle afin d'élucider les différentes voies de dégradation des carraghénanes.

L'étude de la κ -carraghénase de *Z. galactanivorans* a ainsi fourni un éclairage nouveau sur la sous-famille des GH16 impliquées dans la dégradation du κ -carraghénane. En effet, grâce aux travaux réalisés précédemment au sein du laboratoire, des données étaient déjà disponibles sur son homologue chez *P. carrageenovora*. Nous avons donc pu mener une étude comparative de ces deux enzymes et discuter de leurs spécificités respectives à la lumière des nouvelles données biochimiques et structurales obtenues pour la κ -carraghénase de *Z. galactanivorans* (Chapitre 2, article soumis à JBC, actuellement sous presse).

Dans le cadre d'une collaboration avec l'équipe danoise de Peter Stougaard, nous nous sommes également intéressés à une GH apparentée à la famille 42 (GH42-like) et présente sur le plasmide carraghénolytique de *P. carrageenovora*. Sa présence au sein d'un cluster de gènes codants trois sulfatases et plusieurs GH actives sur le carraghénane (iota, kappa, bêta) a permis d'identifier son mode d'action, à savoir $\text{exo-}\beta$ -carraghénolytique, et de proposer une voie conjointe de dégradation des ι et κ -carraghénane chez *P. carrageenovora*. Une étude phylogénétique a également permis d'établir que cette enzyme appartient à une nouvelle famille de GH non décrite jusqu'alors et probablement spécifique du milieu marin (Chapitre 3, ébauche d'article pour soumission à FEBS letter après expériences complémentaires).

Les deux dernières enzymes, une GH29 et une GH non classée, font partie d'un cluster de gènes de type PUL (présence notamment du couple SusC/SusD) chez *Z. galactanivorans*. La GH29 est prédite comme appartenant à la catégorie des α -L-fucosidases, et six autres gènes du cluster codent des enzymes terminales du catabolisme du fucose ainsi qu'un symporteur proton-fucose. Ces éléments laissent donc penser que ce PUL pourrait être dédié à la dégradation d'un substrat fucosylé. Nous avons donc cherché à confirmer la spécificité de ces deux GH pour le fucose, d'abord sur substrat synthétique puis sur substrat naturel à partir de différents extraits d'algues brunes réputées pour leur teneur en fucose (Chapitre 4).

Chapitre 1

L'approche à moyen débit permet d'optimiser la découverte de nouvelles enzymes de bactéries marines impliquées dans la dégradation de polysaccharides algaux

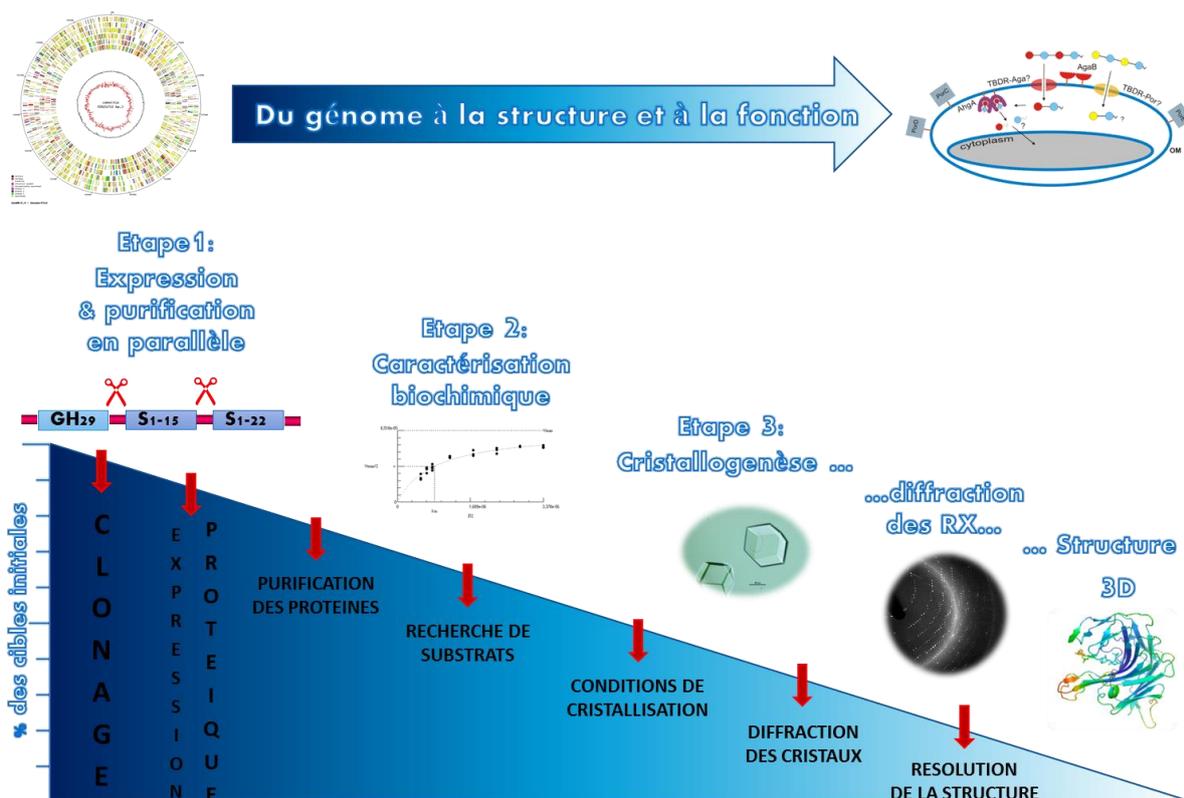


Figure 31 : Représentation schématique des processus expérimentaux permettant de caractériser des protéines d'un point de vue biochimique et structural en suivant une approche à moyen débit

Chapitre 1 : L'approche à moyen débit permet d'optimiser la découverte de nouvelles enzymes de bactéries marines impliquées dans la dégradation de polysaccharides algaux

I. Introduction

A l'air du séquençage haut débit, la facilité d'accès à l'information génétique n'est plus l'étape limitante qu'elle fut par le passé. L'abondance des données obtenues nécessite d'adapter leur traitement en aval par des approches de biologie moléculaire et de biochimie, permettant ainsi de traiter simultanément un grand nombre de cibles d'intérêt (Groissillier *et al.* 2010). Une approche dite « à moyen débit », a été développée dans le contexte de la génomique structurale des années 2000, afin de combler l'écart existant entre l'accumulation des séquences génomiques et la connaissance des structures protéiques associées. Mise au point au sein du laboratoire lors de la thèse de Jérôme Dabin (Dabin 2008), elle a ensuite été appliquée avec succès au cours des thèses de François Thomas ou Aurore Labourel, aboutissant respectivement à l'identification et à la caractérisation du système alginolytique (Thomas *et al.* 2012) ainsi qu'à la découverte d'enzymes du métabolisme de la laminarine (Labourel *et al.* 2014, 2015) chez la bactérie marine *Zobellia galactanivorans*. Le principe de cette approche consiste à croiser toutes les informations disponibles sur le contexte génomique, transcriptomique, ainsi que les données de génomique comparative, afin de sélectionner des cibles d'intérêt thématique majeur, dans notre cas des enzymes potentiellement impliquées dans les voies de dégradation des polysaccharides marins. L'ensemble des cibles présélectionnées sont ensuite clonées et exprimées en parallèle. Cependant, à chaque étape du processus, de la biologie moléculaire jusqu'à l'étude structurale, en passant par la caractérisation biochimique, un certain nombre des cibles de départ va être perdu (Figure 30). Le traitement parallèle d'un grand nombre de cibles permet donc d'optimiser les chances d'obtenir rapidement un nombre suffisant de cibles d'intérêt. C'est cette stratégie à moyen débit qui a été adoptée pour mon projet de thèse et dont je vais à présent détailler les étapes.

II. Analyse *in silico*

Cette approche à moyen débit ne consiste pas à cloner en masse et à l'aveugle, mais bien à sélectionner de façon rationnelle un ensemble de candidats potentiellement intéressants pour une étude biochimique et structurale, et cela en fonction des données à notre disposition. Les

données génomiques de trois bactéries marines récemment séquencées ont servi de point de départ. L'une appartient à la classe des *Flavobacteriia* : *Zobellia galactanivorans* (Barbeyron *et al.* 2016), et les deux autres à la classe des *Gammaproteobacteria* : *Pseudoalteromonas carrageenovora* (Gobet *et al.*, en préparation) et *alteromonas fortis*. Les cibles de clonage ont été choisies selon leur environnement génomique proche, en se focalisant particulièrement sur les groupes de gènes ('clusters') impliqués dans le métabolisme des glucides, principalement des carraghénanes et des fucanes. Comme expliqué en introduction de cette thèse, les bactéries partageant des voies de dégradations similaires pour un substrat donné présentent également une colinéarité dans l'organisation des gènes au sein du cluster en question, témoignant d'une origine évolutive commune. Chez les *Bacteroidetes* auxquels appartient *Z. galactanivorans*, ces regroupements de gènes sont le plus souvent des PULs, généralement identifiables par la présence d'un couple de gènes codants des protéines de type SusD/SusC. Chez les *Gammaproteobacteria*, des équivalents de PULs (des CUTs) existent également, avec la présence de SusC associés à de nombreuses enzymes de type GH et sulfatases.

Une fois ces premiers gènes clés identifiés dans l'un ou l'autre des organismes, des homologues sont recherchés chez les autres. En choisissant ainsi plusieurs homologues du même gène, on augmente les chances de parvenir à la caractérisation complète d'au moins une des protéines homologues codées par ces gènes. Nous nous sommes en particulier focalisés sur des familles de GH, ainsi que des sulfatases et d'autres protéines associées à des GH au sein des clusters de gènes. Des données de transcriptomique obtenues lors de la culture de *Z. galactanivorans* en milieu minimum (Thomas *et al.* 2017) ont également participé à la sélection de nos cibles d'intérêt. Au départ 45 gènes ont donc été sélectionnés dans les trois génomes mentionnés, 8 chez *Z. galactanivorans*, 16 chez *P. carrageenovora* et 21 chez *A. fortis* (Tableau 4), tous impliqués dans des voies de dégradation des polysaccharides algaux, et majoritairement des carraghénanes et des fucanes.

1. Analyse des séquences brutes

Pour chaque cible, la séquence protéique associée a été traitée par une succession d'outils bio-informatiques dans l'ordre suivant, illustrés figure 31 :

- CBS Prediction Servers (<http://www.cbs.dtu.dk/services/>) : utilisation des trois programmes SignalP, TMHMM, LipoP afin de rechercher la présence d'un peptide signal, d'hélices transmembranaires ou de déterminer si la séquence est caractéristique d'une lipoprotéine.

Tableau 4 : Bilan des familles protéiques des différentes cibles du clonage à moyen débit. Sont indiquées les cibles retenues (en bleu dans les colonnes indiquant le nombre de gènes et de constructions) et les cibles écartées (en rouge) à l'issue du choix de la stratégie de clonage.

	Zg	Pc	Af	Nb gènes	%	Nb constructions
Initial	8	16	23	47	100	56
Sélection	Zg	Pc	Af	33	70,2	40
GH	4	8	8	20		27
	GH16, GH29-3c, GHnc-3c, GH97	GH42, GH82, GH16, 2GHnc, 2GH110, GH2	3GH50, 2GH16, GH82, GH82_4c, GH127			
Sulfatases	0	4	2	6		6
	-	S1-8, S1-15, S1-nc, S1-19	S1-nc, S1-19			
Autres	4	0	3	7		7
	AD, Cyp, Ferredox, Ferredox red	-	Cyp, Ferredox, Ferredox red			
Retrait	Zg	Pc	Af	14	29,8	16
			4 GH	4		4
			GH2, GH16, GH29, GH117			
		4 Sulfatases	5 Sulfatases	9		11
		S1-15, 1S1-19, S1-8_3c	3S1-19, S1-4, S1-nc			
			1 Carbohydate estérase	1		1
			CE1			

- BLAST (NCBI) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) : l'utilisation de la suite blastp permet de confronter une (ou plusieurs) séquence(s) cible(s) à différentes bases de données de séquences protéiques. Dans notre cas, nous avons systématiquement utilisé la base de données de séquences protéiques non redondantes, avec l'algorithme blastp, afin de déterminer la modularité éventuelle des séquences. Les informations fournies par Pfam permettent d'identifier des modules et leur fonction potentielle.
- HHpred (Max-Planck Institute) (<https://toolkit.tuebingen.mpg.de/#/tools/hhpred>) (Hildebrand *et al.* 2009; Alva *et al.* 2016) : dans le cas où BLASTp n'a pas permis de trouver des séquences suffisamment proches, la prédiction des structures secondaires par HHpred permet souvent de faciliter la détermination des bornes des modules.
- Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) (Kelly *et al.* 2015): ce serveur permet de prédire la structure tertiaire à partir de la séquence primaire d'une protéine, en se basant sur l'alignement de la séquence d'intérêt avec des séquences de structure connue et sélectionnées sur la base d'un profil de variation évolutive, selon une approche TBM (*Template-Based Modeling*).
- Drawhca (Impmc de Jussieu) (<http://osbornite.impmc.upmc.fr/hca/hca-form.html>): l'HCA (*Hydrophobic Cluster Analysis*) (Henrissat and Callebaut 1995) génère une représentation graphique de la séquence protéique, basée sur les propriétés des acides aminés. Le diagramme HCA permettant d'identifier visuellement des zones typiques de certains éléments de structure secondaire (brin bêta, hélice alpha) et aide ainsi au choix des bornes de clonage pour les protéines modulaires et le retrait des peptides signaux. NB : Le méta-serveur MeDor (<http://www.vazymolo.org/MeDor/>) présente l'avantage de fournir, en plus du graphique HCA, sa traduction en éléments de structure secondaire ainsi que des prédictions sur les zones de désordre de la protéine (Lieutaud *et al.* 2008).

Une fois la (es) présence(s) de peptide signal et/ou de modules détectée(s), on procède à la troncature de la séquence protéique et à celle de la séquence génomique associée, afin de supprimer le peptide signal éventuel et de découper les domaines de façon indépendante. Ainsi, sur les 45 gènes de départ, quatre codaient des protéines modulaires. Il en a résulté 54 constructions à cloner.

Chapitre 1 : Approche à moyen débit et nouvelles enzymes marines

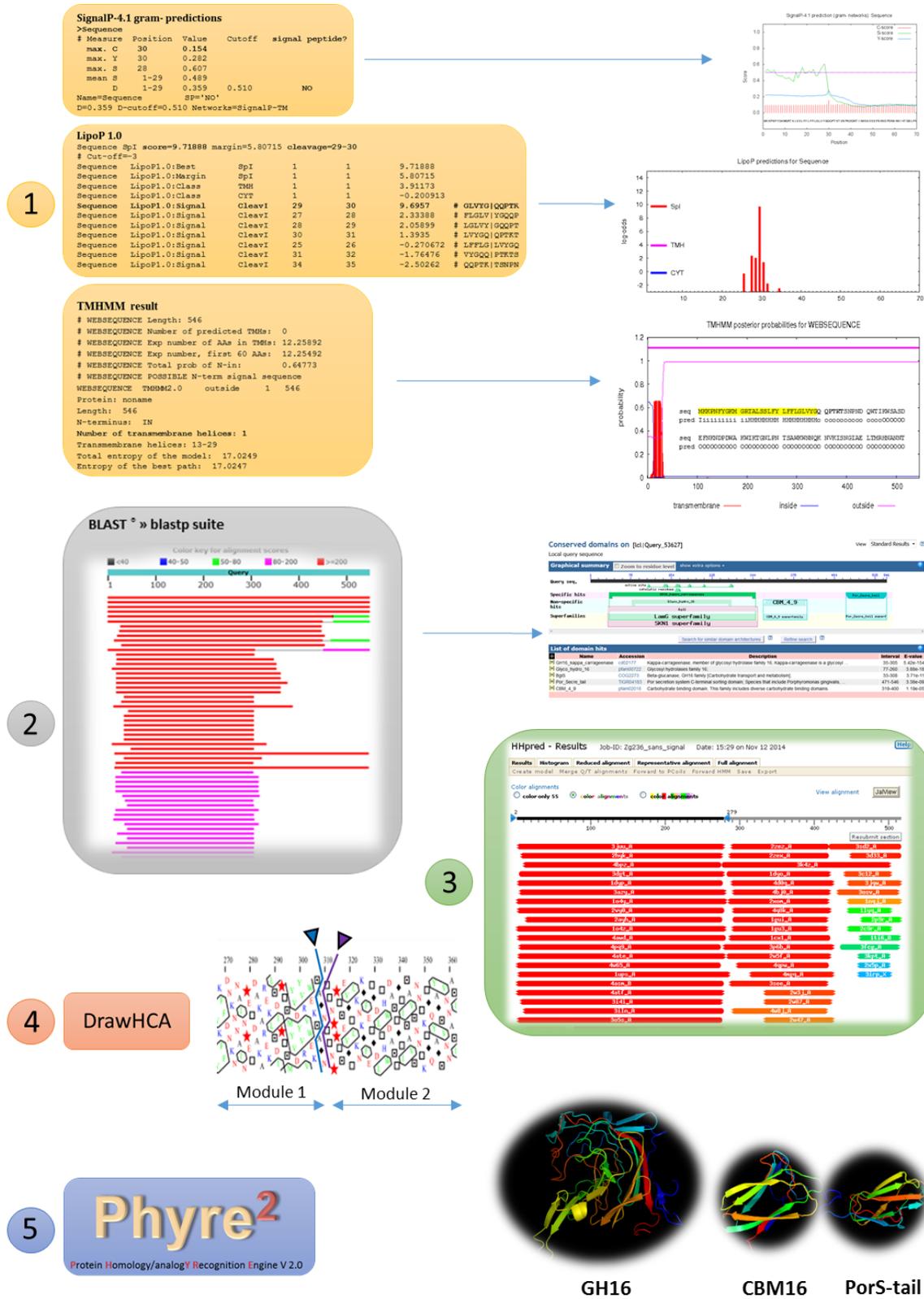


Figure 32 : Succession des outils d'analyse bio-informatique utilisés pour analyser *in silico* une séquence protéique.

2. Recherche de la stratégie de clonage

Le plasmide pFO4, utilisé en routine au sein du laboratoire, a été choisi pour réaliser le clonage. Son site de clonage multiple possède les séquences de reconnaissance uniques des enzymes de restriction BamHI, XhoI, NsiI et EcoRI. Afin d'optimiser le clonage, il faut choisir un nombre minimal de couples d'enzymes de restrictions qui permet d'obtenir un maximum de constructions parmi les 54 envisagées. Afin de maximiser les combinaisons possibles, on ajoute aux 4 enzymes précédentes les 4 enzymes dont les extrémités cohésives générées sont compatibles ce qui nous donne les couples : BamHI/BglII, XhoI/SalI, NsiI/PstI, EcoRI/MunI. Les séquences cibles sont ensuite entrées dans le logiciel CloneManager⁹, qui détecte automatiquement la présence et le nombre de sites de restriction, sachant que pour pouvoir sélectionner un couple d'enzymes, leurs sites de restriction doivent être absents de la séquence cible à cloner. On peut ainsi construire une matrice résumant le nombre de sites de restriction pour chaque séquence et chaque enzyme. Les enzymes présentant les plus grands nombres de sites de clivage sur l'ensemble des séquences cibles sont écartées. C'est le couple BamHI/BglII qui a présenté le nombre minimal d'occurrence parmi nos séquences cibles. On a donc cherché quel autre couple choisir parmi les 3 restants. C'est XhoI/SalI qui s'est révélé le plus intéressant et la combinaison BamHI/XhoI et BglII/SalI celle permettant de cloner le plus de cibles en parallèle dans un seul vecteur pFO4 digéré par BamHI/XhoI. C'est donc 33 gènes qui ont été inclus dans la stratégie de clonage moyen débit, pour un total de 40 constructions. Leur répartition en fonction de l'organisme source et de la famille protéique est résumée dans le tableau 4. Si l'optimisation de la stratégie de clonage (utilisation d'un nombre restreint de couple d'enzymes de restriction) présente l'avantage de faciliter l'étape de réalisation technique du clonage, elle possède néanmoins ses limites. En effet, les cibles qui se retrouvent exclues de la stratégie de clonage peuvent appartenir à des clusters de gènes associés dans un contexte biologique donné. Si l'une des cibles est écartée à cette étape, son absence peut se révéler bloquante à l'étape biochimique de caractérisation des autres enzymes du groupe en question (voir la question des sulfatases aux chapitres 3 et 4).

3. Conception des amorces pour l'amplification des gènes

Afin de faciliter leur clonage, les 40 constructions ont été réparties en deux groupes de 20, l'un compatible avec le couple BamHI/XhoI et l'autre avec BglII/SalI. Pour chaque groupe, les

sites de restrictions respectifs ont été inclus, en amont de la séquence d'appariement du gène et précédé d'une séquence polyA sur l'amorce sens, et en aval suivi d'une séquence polyT sur l'amorce antisens. Un codon stop a également été intercalé entre la fin de la séquence et le site de restriction, dans le cas des cibles modulaires où était absent. Le peptide signal identifié lors de la première analyse des séquences a été systématiquement supprimé, ainsi que le codon d'initiation en l'absence de peptide signal, car apporté par le vecteur d'expression. La taille des amorces a été calculée pour obtenir une température d'appariement comprise entre 70 et 72°C. La liste détaillée des amorces élaborées est donnée en Annexe 3.1.

III. Clonage et tests d'expression

1. Préparation des ADN génomiques

Sur les 3 ADN génomiques nécessaires au clonage, celui de *Z. galactanivorans* et d'*A. fortis* étaient déjà disponibles au laboratoire. Malheureusement les tests d'amplifications à partir de l'ADN d'*A. fortis* n'ont pas fonctionnés. J'ai donc procédé à l'extraction d'ADN « frais » de *P. carrageenovora* et d'*A. fortis* à partir de 100 ml de culture en milieu Zobell, pendant 48h en flasque Erlen de 500 ml. Après 15 min de centrifugation à 3000 g à 20°C, les culots obtenus sont mis à -20°C. Le protocole d'extraction suivi est celui détaillé dans la thèse de Tristan Barbeyron (Barbeyron 1992) et fourni en annexe 2.1. Brièvement, les cellules sont lysées par l'action conjointe du lysozyme, du SDS et de l'EDTA. L'extrait cellulaire est ensuite déprotéinisé par action de la protéinase K et du perchlorate de sodium puis précipité au phénol/chloroforme. L'ADN bactérien est récupéré sur une baguette en verre, lavé à l'éthanol puis séché à l'air libre avant d'être dissout dans une solution de TE. Si la solution présente une contamination protéique importante (couleur opalescente, rapport 260/280 >> 2 au Nanodrop), un passage sur gradient de CsCl peut s'avérer utile (protocole en annexe 2.2).

2. Amplification PCR des inserts

Avant de commencer le clonage des 40 gènes en parallèle, il faut s'assurer de la qualité des deux ADN génomiques ainsi que de la concentration d'ADN génomique nécessaire. Pour ce faire, deux cibles de clonage sont choisies pour chaque ADN génomique, et pour chacune d'elles, 3 dilutions

d'ADN différentes sont testées: 1/10^{ème}, 1/25^{ème}, 1/100^{ème}. L'ADN de *Z. galactanivorans* sert de contrôle positif sur une des cibles, avec une dilution optimale au 2/25^{ème} précédemment établie au laboratoire. L'ADN de *P. carrageenovora* sera utilisé à la dilution 1/25^{ème} et celui d'*A. fortis* à 1/10^{ème}. Une solution stock pour la PCR est préparée en quantité suffisante pour réaliser les amplifications des 35 cibles restantes, plus 1 en contrôle positif de PCR. Sa composition ainsi que le programme PCR sont indiqués en annexe 2.3. Suite à la PCR, la présence et les tailles des 36 fragments amplifiés sont contrôlés sur gel d'agarose (Figure 32). Les inserts sont ensuite digérés avec les couples d'enzymes de restriction approprié, selon le protocole détaillé en annexe 2.4.



Figure 33 : Electrophorèse sur gel d'agarose des produits d'amplification par PCR.

3. Préparation du plasmide pFO4 et clonage

Le plasmide choisi pour réaliser le clonage moyen débit est un plasmide dérivé du pET-15b nommé pFO4 (don de M. Cygler). Il possède la caractéristique d'ajouter une étiquette hexahistidine en amont de l'extrémité N-terminale de l'insert qui sera cloné, et d'être sous le contrôle d'un promoteur fort de type *T7lac*, inductible par le lactose ou son analogue, l'IPTG. La stratégie initialement prévue consistait à digérer le plasmide pFO4 avec le couple d'enzymes BamHI/XhoI avant de procéder à la ligation du plasmide linéarisé avec les inserts digérés par les couples d'enzymes de restriction compatibles. Cependant, j'ai fait face à une difficulté non anticipée lors du choix des enzymes de restriction. Les premiers essais de ligation et de transformation se sont soldés par l'obtention de très peu de colonies (quelques dizaines de cfU) qui se sont révélées être majoritairement des faux positifs lors du criblage par PCR sur colonies (80 % en moyenne)

(protocole en annexe 2.7). Il est possible que ces colonies sont issues de bactéries ayant intégré un plasmide sans insert, soit non digéré, soit s'étant « re-ligué » sur lui-même suite à une digestion incomplète par l'une des deux enzymes de restriction. Il se trouve que lors de sa thèse, François Thomas (Thomas 2011) a rencontré des difficultés similaires avec ce couple d'enzymes. Or après consultation de la documentation technique de New England Biolabs qui commercialise ces enzymes, on peut émettre une hypothèse expliquant le faible taux de succès de clonage associé à l'utilisation de ce couple d'enzymes de restriction. Cela viendrait de la proximité des deux sites de restrictions au sein du site de clonage multiple du plasmide. En effet, comme illustré sur la figure 33, les sites de fixation de BamHI (en vert) et XhoI (en bleu) sur le plasmide sont contigus. Or pour que le clivage par BamHI soit effectif à plus de 90 %, il faut au minimum 2 nucléotides de part et d'autre du site de restriction, et pour XhoI, même avec 3 nucléotides, le clivage n'est que de 75 % après 20h d'incubation (documentation NEB). Or, si l'une ou l'autre enzyme clive le plasmide, il ne restera qu'un seul nucléotide sur l'un des deux brins.



Figure 34 : Séquence du site de clonage multiple du vecteur pFO4 incluant les sites de restriction des enzymes BamHI (en vert) et XhoI (en bleu).

Pour pallier à ce problème, une synthèse *de novo* d'un plasmide pFO4 linéaire a été réalisée par PCR « *inside-out* » du vecteur circulaire. La polymérase haute-fidélité Taq Herculanse II fusion (Agilent) a été utilisée, ainsi qu'un couple d'amorce pFO4-BamHI (Antisens)/pFO4-XhoI (Sens). Après amplification, le vecteur est digéré par le couple d'enzymes de restriction BamHI/XhoI. Le protocole détaillé ainsi que les séquences des amorces sont précisées en annexe 2.5 et 4. Le vecteur et les inserts sont ensuite mis ensemble en présence de ligase et le mélange est transformé dans des cellules compétentes d'*E. coli* NEB5α (annexe 2.6). Après étalement sur boîte de Pétri en milieu LB+ampicilline et incubation à 37°C sur la nuit, on compte le nombre de colonies par boîte et on réalise un criblage PCR de 4 colonies par boîte (annexe 2.7). Si l'amplification PCR a fonctionné, on compare le poids moléculaire apparent de la bande obtenue au poids moléculaire théorique de chaque insert amplifié à l'aide des amorces T7F et T7R.

La stratégie « PCR *inside-out* » s'est révélée efficace avec un nombre de colonies supérieure (75 en moyenne) et un taux de faux positifs plus faible que la double digestion du plasmide

circulaire (54 % au lieu de 80 %). Cependant certains clones se sont montrés très récalcitrants, et ont nécessité le criblage d'une trentaine de colonies avant d'obtenir un clone positif, ceci pouvant souvent s'expliquer par la taille importante de l'insert (ex. : pMM9 et 10 : 2,8 et 2,6 kb respectivement ont nécessité de cribler 39 colonies pour un total de 4 et 1 colonies positives respectivement). La première série de criblage de 4 cfU par boîte a permis d'obtenir 27 clones positifs sur les 40 cibles, puis deux autres séries de criblage ont permis d'obtenir 11 des 13 cibles manquantes

Afin de contrôler la qualité des clones « positifs » obtenus après l'étape de criblage par PCR sur colonie, et ce avant de procéder à un séquençage coûteux, on réalise une deuxième étape de contrôle par digestion des minipreps des clones positifs avec l'enzyme de restriction HincII (annexe 2.8). Le profil de digestion est contrôlé sur gel d'agarose. On aura au préalable calculé le nombre et la taille théorique des fragments à obtenir pour chaque clone en construisant chaque plasmide dans le logiciel CloneManager et en le « digérant » *in silico* par HincII. Cette étape a permis de s'apercevoir que 4 clones apparemment positifs suite à l'étape de criblage sur colonie ne l'étaient pas en réalité. Il a donc fallu revenir parfois à l'étape précédente en criblant de nouvelles colonies pour obtenir un bon profil de digestion HincII (ex : pMM21 : 1cfU+/4 mais profil HincII mauvais → retour au criblage sur colonies : 1 cfU+/13 : profil HincII ok). Le tableau en annexe 3.2 regroupe les résultats de chaque étape du clonage.

4. Tests d'expression à petite échelle

Une fois les clones positifs confirmés par la digestion HincII, on procède à leur transformation dans la souche d'expression BL21(DE3) d'*E. coli* puis à leur mise en pré-culture en milieu LB + ampicilline. A partir de ces pré-cultures on réalise des stocks glycérols (annexe 3.8) puis on ensemence en parallèle des mini-cultures en plaques 24 puits à fond rond (Figure 34) dans 3 ml de milieu auto-inducteur ZYP5052 (Studier 2005) et on incube 3 jours à 20°C, sous agitation à 180 rpm. Après centrifugation des plaques de mini-culture, le surnageant est éliminé et les culots congelés à -20 °C. Le choix du milieu auto-inducteur ZYP5052 a été fait afin d'éviter l'étape d'induction manuelle qui nécessite le suivi de la croissance bactérienne par mesure de la DO_{600nm}, ce qui permet ainsi de traiter en parallèle un grand nombre de cibles en minimisant les manipulations. Ce milieu est également connu pour fournir des rendements protéiques plus élevés

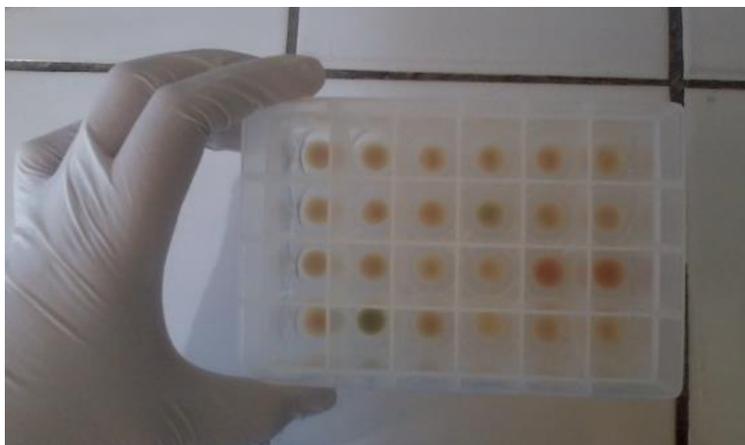


Figure 35 : Plaqué 24 puits utilisée pour les cultures à petite échelle avec culots bactériens après trois jours de culture.

que d'autres milieux plus classiques (LB, YT, SOB, SOC...). Il est donc particulièrement indiqué dans l'optique d'obtenir des quantités de protéines importantes.

Afin de valider la qualité de l'expression protéique, on réalise des mini-purifications sur colonne d'affinité Nickel-sépharose. Brièvement, les culots bactériens sont lysés par sonication, un aliquot de la fraction totale est conservé. Après centrifugation, on prélève un aliquot du surnageant et le reste est purifié sur mini colonne His-SpinTrap (GE Healthcare). On réalise ensuite un gel SDS-PAGE avec 3 pistes par cible protéique : la fraction totale, la fraction soluble et la fraction purifiée. Un exemple de gel est présenté en Figure 35 et la totalité se trouve en annexe 4. La comparaison de l'intensité des bandes entre les trois pistes permet de déduire le type d'expression, si expression il y a. Dans notre cas, 5 protéines n'ont pas du tout été exprimées (pas de bande visible dans aucune des trois pistes, comme pMM9 et 10 par exemple). Parmi les 32 autres qui présentaient une bande à la taille attendue dans la fraction totale, signe d'expression par *E. coli*, 3 sont totalement insolubles (pas de bandes visibles, ni dans les fractions solubles ni après purification, comme pMM7) et 5 autres ne sont que partiellement solubles (perte d'intensité importante de la bande entre la fraction totale et les fractions suivantes, comme pMM8 par exemple). Sur les 37 cibles de départ, 29 protéines sont donc exploitables en purification dont 24 très fortement solubles, condition favorable pour une étude structurale ultérieure nécessitant de grandes quantités de protéines. Il faut cependant noter que pour les protéines partiellement ou totalement insolubles, 8 dans notre cas, tout espoir n'est pas perdu !

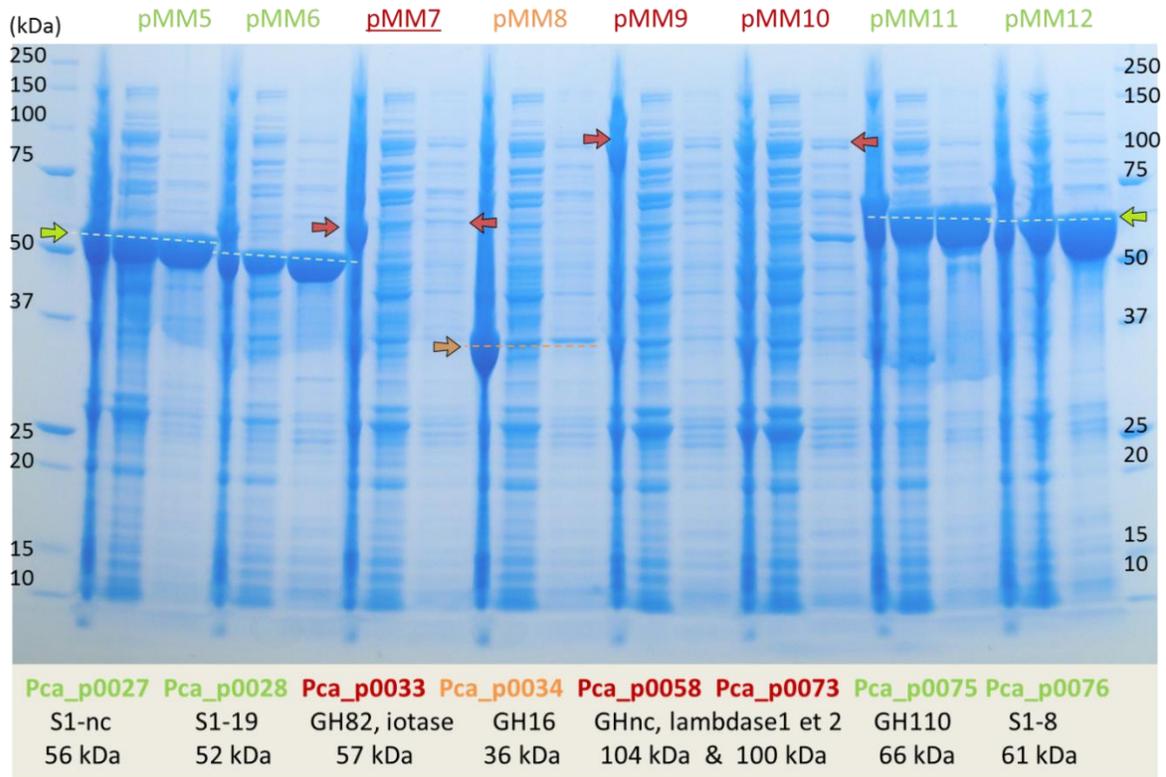


Figure 36 : Exemple de gel SDS-PAGE permettant de visualiser le profil d’expression et de purification des protéines recombinantes suite à leur expression petite échelle. Piste 1 : Fraction totale, Piste 2 : Fraction soluble, Piste 3 : Elution après purification sur colonne His-SpinTrap.

En effet, ces tests d’expression ont été réalisés dans un milieu auto-inducteur à 20°C pour faciliter le traitement simultané d’un grand nombre de protéines. Mais pour certaines protéines, ces conditions d’expression ne sont pas du tout optimales. A titre d’exemple, les conditions idéales pour produire la κ -carraghénase de *P. carrageenovora* s’avèrent être une culture à 12°C en milieu M9 + ampicilline, avec une induction IPTG classique (Michel *et al.* 1999). Pour les protéines partiellement ou totalement insolubles, plusieurs stratégies sont donc envisageables :

1) Le tampon de lyse et de purification ainsi que la méthode de lyse bactérienne peuvent également jouer sur la solubilité des protéines. Ainsi la sonication peut entraîner une précipitation des protéines liée à l’augmentation locale de la température à l’extrémité de la sonde. L’utilisation d’une presse de French ou l’une lyse chimique peuvent représenter des alternatives. De même le choix du pH des tampons de lyse et de purification est très important : il faut être suffisamment éloigné du pI de la protéine pour maximiser sa solubilisation ;

- 2) Tester différentes conditions de cultures, en modifiant la nature du milieu (Graffinity,...), la température d'incubation (12°-15°-17°C), le volume de culture (1L, 5L, ...) et le taux d'oxygénation (en diminuant le volume de culture pour une taille d'erlenmeyer donnée). L'utilisation d'un multi fermenteur serait une solution intéressante pour optimiser ces différents paramètres. Les conditions d'induction peuvent aussi avoir un fort impact (milieu auto-inducteur ou IPTG à différentes concentrations)
- 3) Si aucune de ces solutions ne permet d'obtenir des quantités de protéines suffisantes pour une étude structurale, il faut remonter à l'étape précédente, en changeant de souche d'expression d'*E. coli*. En effet, pour des raisons de préférence de toxicité cellulaire, *E. coli* peut avoir des difficultés à exprimer correctement certaines protéines. Il existe des souches telles que C41 ou 43 (DE3) qui sont plus tolérantes à la toxicité de certaines protéines ;
- 4) Une stratégie visant à co-exprimer la protéine d'intérêt avec une combinaison de protéines chaperonnes est également possible. ex : HSP (*Heat Shock Protein*) : hsp60, hsp70.
- 5) Le choix de l'emplacement C-terminal ou N-terminal de l'étiquette peut aussi avoir de l'importance ; de même que le changement de l'étiquette, telle que le GST (glutathion S-transferase), *via* le clonage dans un vecteur pGEX;
- 6) Retourner à l'analyse de la séquence primaire et concevoir différentes constructions protéiques en faisant varier les bornes de l'insert à cloner. Cela peut avoir un impact déterminant. A titre d'exemple, pour la κ -carraghénase de *Z. galactanivorans*, les bornes établies dans un premier temps et délimitées par des traits violets sur le diagramme HCA de la figure 31 n'ont permis d'obtenir qu'une expression partiellement soluble du module catalytique de l'enzyme (pistes correspondant à pMM1 en annexe 4). Une deuxième construction, ne différant de la première que par l'ajout d'un acide aminé en N-ter et le retrait de trois autres en C-ter, a permis d'obtenir une protéine totalement soluble avec un très fort taux d'expression.

Dans le cas des protéines non exprimées (ou non détectable sur gel SDS), on peut également envisager les solutions présentées en 5) et 6) ou encore en 2) et 3) avec d'autres souches spécifiques à des problèmes liés aux biais de codon.

5. Contrôle qualité par séquençage

En parallèle de l'étape précédente d'expression à petite échelle, les 37 cibles ont été séquencées par la méthode de Sanger, sur la plateforme Génomer de la Station Biologique. Les protocoles de réaction de séquence, à l'aide du kit BigDye® XTerminator v3.1, ainsi que de purification des réactions de séquence sont fournis par la plateforme. Chacune des 37 séquences a subi deux opérations de séquençage (ou '*read*'), une sur le brin sens et une sur le brin antisens afin de permettre un recouvrement total de la séquence. En effet, la taille moyenne des inserts étant de 1500 pb, avec des tailles allant de 330 à 2799 pb, et les réactions de séquence ayant lieu en moyenne sur 700 à 800 pb sans générer d'erreur, dans la majorité des cas 2 '*reads*' suffisent. Cependant, pour 9 des cibles de grande taille, un couple d'amorces interne (fourni en annexe3.3) a dû être conçu en plus du couple standard T7F/T7R afin d'obtenir un recouvrement total de la séquence. Ce séquençage a validé la qualité des 37 clones identifiés comme positifs suite à la digestion par HincII. Cette étape de contrôle par HincII a donc permis d'économiser au minimum 10 réactions de séquence, correspondant aux 5 cibles présentant un profil HincII non conforme.

C'est à cette étape que s'achève la partie « moyen débit » à proprement parler. Les parties de biochimie et de biologie structurale n'ont pas été abordées dans l'optique de mener en parallèle un grand nombre de cibles mais ont plutôt tiré profit de l'approche moyen débit dans la possibilité de choisir, parmi les cibles obtenues, les plus prometteuses en termes de rendement protéique.

IV. Purification, cristallogénèse et cristallographie aux rayons X

L'obtention de la structure cristallographique d'une protéine nécessite de franchir trois étapes qui sont autant de points critiques :

- 1) L'obtention de la protéine soluble, purifiée et concentrée en quantité suffisante
- 2) La détermination de paramètres de cristallisation permettant d'obtenir des cristaux stables et qui diffractent
- 3) La capacité à résoudre le problème des phases permettant d'expliquer en intégralité la carte de densité électronique du cristal et de pouvoir ainsi reconstruire la structure 3D de la protéine

1. Production et purification des protéines recombinantes

Une fois les cibles d'intérêt sélectionnées pour poursuivre leur étude biochimique et structurale, des cultures sont réalisées en plus gros volumes. Classiquement pour des protéines solubles qui exprime fortement, 1 l de culture en milieu ZYP5052 dans un erlenmeyer de 5 l permet de produire suffisamment de protéines pour réaliser les étapes ultérieures. Après une incubation de trois jours à 20 °C sous agitation à 180 rpm, les cultures sont centrifugées et les culots congelés à -20°C pour leur traitement ultérieur.

Pour chaque protéine à purifier, une analyse de la séquence protéique recombinante (avec étiquette histidine) grâce à l'outil ProtParam disponible sur le serveur ExPASy (<http://web.expasy.org/protparam/>) permet d'obtenir les valeurs théoriques du poids moléculaire, du coefficient d'extinction molaire et du pI de la protéine. Les deux premiers paramètres serviront à doser la concentration de la protéine après sa purification. Quant au pI (point isoélectrique), c'est en fonction de sa valeur que le pH de la solution tampon A, dans laquelle le culot bactérien va être lysé puis purifié, est mis au point. La solubilité d'une protéine étant minimale lorsque le pH environnant est égal à son pI, le pH choisit pour le tampon A est généralement distant de 1 à 2 unités de la valeur du pI. Le tampon A se constitue donc d'une solution au pouvoir tampon adaptée à la valeur de pH choisie (pour un pH compris entre 7 et 9 on utilisera classiquement du Tris à 50 mM), de NaCl pour assurer une force ionique suffisante à la solution (classiquement de 250 à 500 mM), et une faible concentration d'imidazole (10-50 mM) sur laquelle je reviendrai plus loin.

Au cours de ma thèse, j'ai eu recours exclusivement à la lyse bactérienne mécanique par utilisation d'une presse de French. Brièvement, le culot bactérien est resuspendu à froid dans la solution tampon A additionnée de lysozyme, de DNase et d'un mélange d'antiprotéases EDTA-free. Le passage dans la presse permet de faire éclater les cellules bactériennes. Une fois le lysat récupéré, il est centrifugé 30 min à 10 000 g puis filtré sur 0,2 µm avant d'être purifié par chromatographie d'affinité sur une colonne Nickel-sépharose. Grâce à la présence du tag hexahistidine la protéine recombinante va s'adsorber sur la résine par interaction de charges avec les ions nickel. Une fois l'échantillon injecté (en 1 sur la figure 36), le rinçage de la colonne avec la solution tampon A permet d'éliminer les protéines d'*E. coli* fixées de façon aspécifique (en 2 et 3), et ce grâce en partie à la présence d'imidazole en faible concentration dans le tampon A. En effet,

l'imidazole est constitué d'un noyau imide, comme l'histidine, qui lui confère le même caractère basique que celle-ci. Il va donc se fixer aux ions Ni^{2+} de la colonne, prenant la place des protéines non taguées. Après l'étape de lavage, la protéine d'intérêt est éluée au moyen d'un gradient d'imidazole (en 4), en augmentant progressivement le pourcentage de tampon B passant sur la colonne. La composition du tampon B est identique à celle du tampon A à l'exception de la concentration d'imidazole, généralement comprise entre 500 mM et 1 M. Lorsque la concentration d'imidazole atteint un seuil critique, différent pour chaque protéine, celle-ci est éluée de la colonne, générant un pic sur la courbe de la DO_{280nm} (en 5).

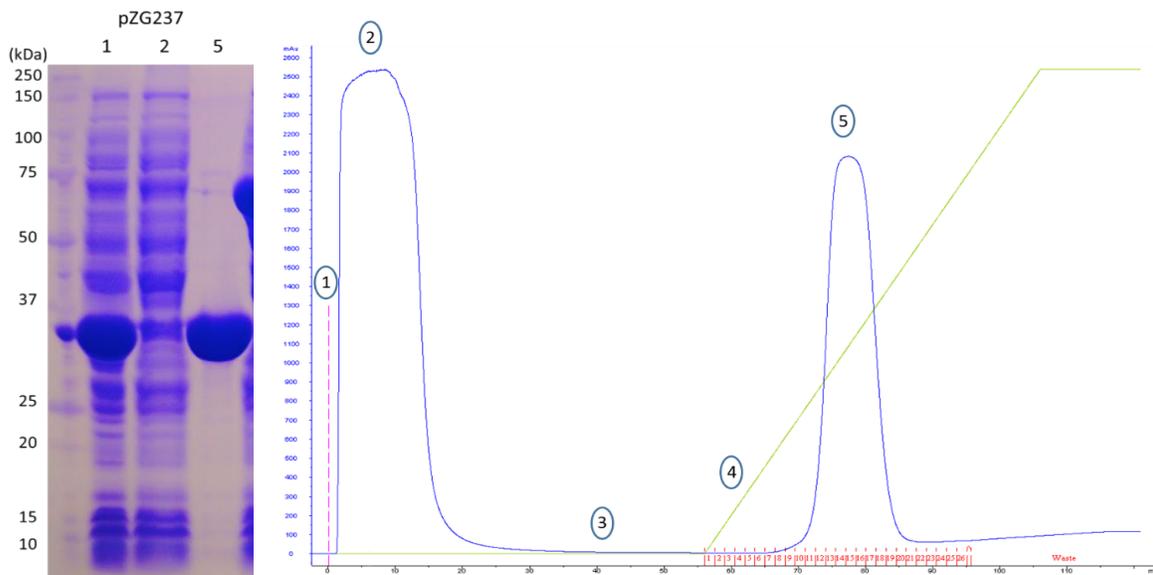


Figure 37 : Exemple de gel SDS-PAGE réalisé au cours d'une chromatographie d'affinité dont les étapes sont numérotées sur le chromatogramme associé.

Piste 1 : échantillon injecté en 1 ; Piste 2 : partie non retenue sur la colonne durant l'injection en 2 ; Piste 3 : fractions 11 à 19 correspondant au pic en 5

Les fractions correspondant au pic sont ensuite collectées et déposées sur gel SDS-PAGE pour contrôler la qualité de la purification et l'état de notre protéine (poids moléculaire apparent comparé au théorique). Les fractions du pic sont ensuite réunies et concentrées avant de subir une deuxième étape de purification, par chromatographie d'exclusion stérique, sur une colonne de type Superdex 75 ou 200 selon la taille de la protéine.

Cette fois-ci, la séparation des protéines se fait en fonction de leur rayon hydrodynamique. Les protéines de grande taille vont être éluées en premier car exclues de la porosité du gel. Les plus

petites vont mettre le temps maximal pour sortir de la colonne car elles auront parcouru le chemin le plus long à l'intérieur du gel.

Plus la gamme de taille des protéines contaminantes de l'échantillon est éloignée de celle de la protéine à purifier, meilleure sera la qualité de la purification. Le débit joue également un rôle important sur la résolution des pics, ainsi que le volume et la concentration de l'échantillon injecté, mais également le volume de résine parcouru par l'échantillon.

C'est lors de cette deuxième étape de purification que la protéine est dessalée, en équilibrant la colonne Superdex dans une solution tampon C sans imidazole, avec une concentration atténuée en NaCl, et en changeant la nature de la molécule tampon si nécessaire (Tris, HEPES,...). Une fois le pic majoritaire correspondant à la protéine d'intérêt identifié, les fractions correspondantes sont rassemblées et concentrées une nouvelle fois, au moyen d'un système d'ultrafiltration de type Centricon, avec une valeur de cut-off adaptée au poids de la protéine. La concentration de la protéine est mesurée au Nanodrop® à 280 nm. Pour initier la cristallogénèse, on cherche généralement à atteindre une concentration protéique inférieure à 50 % de la limite de solubilité de la protéine.

2. Cristallogénèse, diffraction des rayons X et résolution structurale

Durant cette thèse, la technique de cristallogénèse des protéines basée sur la diffusion de vapeur a été utilisée. Cette technique consiste à permet à une goutte contenant un mélange 1 : 1 ou 2 : 1 de protéine purifiée et d'une solution de cristallisation (composée d'agent précipitant, d'additifs et d'un tampon) d'établir un équilibre avec un réservoir beaucoup plus grand qui contient uniquement la solution de cristallisation. Au départ, la goutte de protéine ne contient pas une concentration suffisante en agent précipitant pour permettre la cristallisation. Le gradient de diffusion de vapeur qui s'établie entre la goutte et le réservoir entraîne des échanges de solvant volatile aboutissant à la concentration de la protéine et de l'agent précipitant dans la goutte quand il s'agit d'un solvant aqueux. C'est au cours de ce processus que des événements de nucléation de cristaux peuvent se produire. La recherche des meilleures conditions de cristallisation se fait en deux temps. Tout d'abord un grand nombre de conditions sont criblées, par la méthode de la goutte assise en plaque Corning 96 puits remplie par un robot HoneyBee 963 (Cartesian). Les kits commerciaux testés en premier sont le PACT et le JCSG+ (Quiagen). Si aucun des deux ne donne de résultats, d'autres peuvent être tentés. Les meilleures conditions de cristallisation obtenues sont

ensuite affinées manuellement en plaque 24 puits par la technique de la goutte suspendue. Le but est d'obtenir des cristaux suffisamment gros (de l'ordre de $100\ \mu\text{m} \times 100\ \mu\text{m} \times 100\ \mu\text{m}$) et avec une bonne qualité de diffraction des RX. La qualité de diffraction peut être évaluée avant de se déplacer au synchrotron par utilisation du diffractomètre de laboratoire disponible à la Station (X8 Proteum, Bruker). La collecte des données a ensuite été réalisée suivant les cas, à l'ESRF de Grenoble sur la ligne ID29, ou au synchrotron Soleil (ligne PX1 pour GH29 et GH42-like) de Paris-Saclay.

Le principe de la diffraction des RX par un cristal protéique repose sur les propriétés intrinsèques de diffusion élastique des photons par les atomes au sein de l'arrangement cristallin. Le faisceau de RX incident projeté sur le cristal va être diffusé différemment par chaque atome qui croisera son trajet optique. A cela s'ajoute le phénomène de diffraction par un réseau régulier constitué par l'agencement des protéines au sein du cristal. La convolution de ces deux phénomènes génère un profil de diffraction constitué de taches discrètes où la distance entre les taches est inversement proportionnelle aux distances du réseau régulier, et où l'intensité est liée au pouvoir diffusant de l'ensemble des atomes. Cette carte de diffraction est en quelque sorte une projection en 3D dans l'espace dit réciproque, de la structure 3D de la protéine au sein du cristal, dans l'espace dit réel ou direct. Les coordonnées (x, y, z) des atomes dans l'espace réel sont donc liées à ceux dans l'espace réciproque (h, k, l) par le facteur de structure du cristal qui s'écrit pour chaque réflexion hkl :

$$F(\vec{H}) = \int_{x=0}^1 \int_{y=0}^1 \int_{z=0}^1 \rho(x, y, z) e^{2i\pi(hx+ky+lz)} dx dy dz = \sum_{j=1}^n f_j e^{2i\pi\vec{H}\cdot\vec{r}_j} = |F(\vec{H})| e^{i\varphi(\vec{H})}$$

où \vec{H} est le vecteur de diffraction de la réflexion, ρ est la densité électronique de la maille, \vec{r}_j est le vecteur position et f_j le facteur de diffusion atomique de l'atome dans la maille, et $\varphi(\vec{H})$ est la phase de la réflexion. Or la phase n'est pas accessible expérimentalement. Seules les intensités peuvent être mesurées, l'information sur les phases des réflexions est donc perdue. C'est ce qui constitue le « problème des phases » en cristallographie. L'indexation et l'intégration des taches de diffraction sont d'abord réalisées à l'aide d'un logiciel tel que Mosflm ou XDS, puis les données sont mises à l'échelle à l'aide d'un logiciel de type SCALA ou XDSCALE. Pour résoudre le problème des phases, différentes méthodes existent, basées sur l'utilisation de la fonction de

Patterson, qui fait appel à la transformée de Fourier du carré des intensités mesurées expérimentalement.

Dans le cas où des structures de protéines homologues sont disponibles, à savoir au cours de cette thèse pour *ZgCgkA_{GH16}* et *ZgAlfA_{GH29}*, la résolution par remplacement moléculaire est envisageable. Le principe de cette méthode consiste à utiliser comme modèle structural une protéine homologue (à partir de 30% d'identité) dont la structure a déjà été résolue. En effet, le positionnement correct d'une structure homologue par rapport au groupe d'espace de la structure à déterminer permettra de faire une première estimation des phases. Le programme MolRep calcule ensuite, en entrant notre jeu de données traité précédemment ainsi que les coordonnées de la protéine homologue, la corrélation entre le modèle et les données expérimentales pour en déduire une possible solution structurale.

Dans le cas où la séquence est trop originale pour trouver un homologue suffisamment proche d'une structure connue, comme pour la GH42-like de *P. carrageenovora*, une résolution par mesure de la diffusion anormale d'une protéine séléno-méthionylée à plusieurs longueurs d'angles (MAD) est envisagée. Pour cela il faut faire exprimer la protéine recombinante dans un milieu de croissance où la méthionine est remplacée par de la séléno-méthionine, ce qui garantit un marquage homogène des protéines (Hendrickson *et al.* 1990). Ensuite les cristaux sont soumis à trois longueurs d'ondes correspondant aux seuils d'excitation des différentes couches électroniques de l'atome de sélénium. Ces trois jeux de données permettent de résoudre le problème de la phase et ainsi de reconstruire la carte de densité électronique de la protéine.

V. Détermination des substrats des glycosides hydrolases

Afin de déterminer le ou les substrats privilégié(s) d'une enzyme donnée, il est nécessaire d'avoir à notre disposition un panel suffisamment large de substrats, naturels comme artificiels, à cribler. Le tableau 5 montre la liste des substrats utilisés au cours de cette thèse. Pour les substrats naturels, j'ai constitué une « mini-banque » de substrats en collectant des algues fraîches, en les broyant et en leur faisant subir une extraction à l'alcool. Les produits obtenus sont nommés AIR (*Alcohol Insoluble Residue*). Le protocole détaillé est indiqué en annexe 5.2. Certains AIR m'ont été gracieusement donnés par des collègues. J'ai également rassemblé divers polysaccharides

commerciaux disponibles au laboratoire et leur ai fait subir une étape de dialyse afin d'éliminer les éventuels produits de dégradation oligosaccharidiques présents. Le recours aux substrats de synthèse peut également s'avérer très utile. En effet, comme évoqué au cours de l'introduction, la paroi des algues est d'une telle complexité que la stratégie de dégradation par les bactéries marines implique généralement l'action simultanée de plusieurs enzymes, de spécificités légèrement différentes, et dont les effets seront synergiques. Si l'on se contente de tester l'activité de nos enzymes de fonction inconnue sur des extraits naturels, il peut arriver (comme présenté au chapitre 4) qu'on n'obtienne aucun résultat. Dans ce cas de figure, l'usage de substrats synthétique, de structure standardisée et relativement simple, permet souvent d'obtenir une première idée de la spécificité des enzymes apparemment inactives sur substrat naturel.

1	<i>Ulva</i> (UF30kDa)
2	Ulvanes
3	<i>Solieria chordalis</i>
4	Porphyrane
5	κ -carrageenan- <i>Euchema cottonii</i>
6	ι -carrageenan- <i>Euchema denticulatum</i>
7	κ/ι -carrageenan- <i>Chondrus crispus</i> (gametophyte)
8	λ -carrageenan- <i>Chondrus crispus</i> (sporophyte)
9	ξ -carrageenan- <i>Chondracanthus acicularis</i>
10	<i>Palmaria palmata</i> , mix-linked-xylans (AIR)
11	<i>Pelvetia canaliculata</i> fucoïdanes (AIR)
12	<i>Fucus serratus</i> fucoïdanes (AIR)
13	Fucoïdanes (fraction S28 from <i>P. canaliculata</i> - UF30kDa)
14	<i>Himanthalia elongata</i> (AIR)
15	<i>Laminaria digitata</i> (AIR)
16	<i>Pylaiella littoralis</i> (AIR)
17	<i>Ascophyllum nodosum</i> (AIR)
18	<i>Laminaria hyperborea</i> (AIR)
19	pNP-Glucose
20	pNP-Galactose
21	pNP-Fucose
22	3-Fucosyllactose
23	2-Fucosyllactose
24	α -3-galactosyl-3 fucosyllactose
25	lacto-N-fucanopentaose I (LNFP I)

Tableau 5 : Liste des substrats utilisés au cours de la thèse. Les numéros des substrats issus des algues vertes sont colorés en vert, ceux des algues rouges en rouge, des algues brunes en marron, ceux des oligosaccharides fucosylés sont en bleu clair, et ceux des substrats synthétiques en jaune.

Afin de déterminer le(s) substrat(s) d'une nouvelle enzyme, la marche à suivre dépend du niveau de connaissance que l'on a de l'enzyme. Dans certains cas, sa séquence a permis de lui

attribuer une famille de glycoside hydrolase, voire même parfois une sous-famille spécifique d'un substrat précis. Dans ce dernier cas, comme par exemple la GH29 de *Z. galactanivorans* qui est prédite comme étant une α -L-fucosidase, il reste à démontrer la véracité de la fonction prédite. Cela facilite grandement les choses puisqu'on peut ainsi adapter la technique de détection au type d'enzyme. En effet, les enzymes de type exolytique génèrent uniquement des oligosaccharides de très petite taille voir des monosaccharides, et parfois dans des quantités moindres que des enzymes endolytiques. Il faut donc avoir recours à des techniques de révélation de type TLC (*Thin Layer Chromatography*) afin de pouvoir visualiser les monosaccharides ou aux électrophorèses pour saccharides de type FACE (*Fluorophore Assisted Carbohydrate Electrophoresis*) ou C-PAGE (*Carbohydrate-Polyacrylamide Gel Electrophoresis*) ou encore par des techniques d'HPLC (*High Pressure Liquide Chromatography*) pour visualiser des oligosaccharides de petite taille. Dans le cas des enzymes de type endolytique, leur activité génère une série d'oligosaccharides de tailles variables, qui est visualisable par C-PAGE, FACE ou HPLC.

Dans le cas d'enzymes de séquence trop originale pour pouvoir être associée à une famille pré-existante de GH, une première approche consiste à regarder son environnement génomique afin de tenter de restreindre la gamme de substrats à tester. Ainsi la GHnc qui sera présentée au chapitre 4, se trouve dans une structure de type PUL, aux côtés de la GH29 précédemment citée. La GHnc devrait donc agir sur le même substrat naturel que la GH29, mais avoir un autre site de reconnaissance que cette dernière. Si les substrats testés ne donnent pas de résultats positifs, cela peut vouloir dire que notre enzyme fonctionne dans une étape avale du traitement du substrat, après que d'autres enzymes aient rendu accessible le site de reconnaissance de notre enzyme. Par exemple après désulfatation du substrat. Cette action synergique des enzymes sur un substrat naturel complexe peut parfois empêcher de déterminer le substrat d'une nouvelle enzyme tant que toutes les enzymes nécessaires à la dégradation du substrat n'ont pas été identifiées. Comme cela sera discuté au chapitre 4, ce cas de figure doit amener à une stratégie de clonage systématique du groupe d'enzymes entourant l'enzyme de fonction inconnue afin de maximiser les chances de découvrir le substrat naturel.

Enfin, dans le cas où le substrat naturel n'a pu être trouvé, le recours à des substrats synthétiques de type pNP (para-nitrophényl) ou des oligosaccharides issus du lait peuvent être testés. Les pNP-oses sont des oses sur lesquels a été greffée chimiquement une fonction para-

nitrophényl. Tant que celle-ci est branchée à l'ose, le couple est incolore en solution. Si une enzyme est capable de reconnaître l'ose et de cliver la liaison chimique le reliant au pNP, le pNP libéré en solution absorbe dans le visible à 410 nm. Cela permet de cribler par spectrométrie la spécificité de plusieurs enzymes en parallèle et même de réaliser des cinétiques d'activité enzymatique.

Dans le cas des oligosaccharides issus du lait, on n'est pas à proprement parler sur des substrats artificiels, mais dans ce cas leur source ne correspond bien-sûr pas au substrat naturel de nos enzymes marines. Cependant, les motifs saccharidiques (branchements, enchainement et types de liaisons) se retrouvant aussi bien dans les parois des algues que dans le lait humain, ils représentent des outils précieux pour déterminer la spécificité d'une enzyme.

VI. Caractérisation biochimique des glycosides hydrolases

Comme vu au paragraphe III de l'introduction, les glycosides hydrolases clivent la liaison glycosidique entre deux oses, générant ainsi une extrémité réductrice à chaque cycle catalytique. On peut donc suivre l'activité des enzymes et déterminer leurs paramètres cinétiques par mesure de la variation du pouvoir réducteur du mélange réactionnel contenant l'enzyme, le substrat et les produits de clivage. Pour cela, on réalise un dosage colorimétrique par une méthode adaptée de Kidby et Davidson pour le dosage des « sucres réducteurs » (Kidby and Davidson 1973). Contrairement aux enzymes de type glycoside lyases où la formation d'une insaturation absorbant à 235 nm peut être suivie au cours d'une cinétique en temps réel dans un spectromètre, le suivi cinétique des glycosides hydrolases ne peut être fait en temps réel, et nécessite de réaliser une réaction colorimétrique secondaire. Ainsi, le dosage des sucres réducteurs par la méthode du ferricyanure, qui a été utilisée durant cette thèse, nécessite de chauffer le mélange réactionnel en présence d'un réactif colorimétrique (le ferricyanure de potassium), dont la décoloration sera proportionnelle à la concentration en extrémité réductrice présente dans le mélange. Pour obtenir des points de cinétique, il faut donc arrêter manuellement la réaction à chaque temps voulu pour les données de cinétiques. Pour les réactions sur pNP-oses, la réaction peut être bloquée manuellement par ajout de carbonate de sodium ou suivie en cinétique en temps réel dans un spectrophotomètre, mais ne nécessite pas de chauffage.

Dans tous les cas, la caractérisation biochimique d'une enzyme se passe en plusieurs étapes

afin de déterminer successivement :

- 1) la stabilité thermique : à court terme en utilisant une mesure par DLS (*Dynamic Light Scattering*) et au cours du temps, lors en suivant l'impact du temps et des conditions de stockage sur l'activité de l'enzyme
- 2) la température optimale d'activité de l'enzyme
- 3) le pH optimal
- 4) les paramètres cinétiques de l'enzyme

1. Stabilité thermique

La mesure par DLS permet de déterminer les tailles et la dispersion des tailles de particules ou de polymères en solution. Cette technique utilise le principe de diffusion de la lumière qui sera différente en fonction de la taille des particules et de leur coefficient de diffusion. L'intensité de diffusion qui est enregistrée au cours du temps est ensuite traitée mathématiquement par un logiciel pour calculer le rayon hydrodynamique des particules en solution et leur distribution de taille. Le rayon hydrodynamique correspond au rayon d'une sphère théorique qui aurait le même coefficient de diffusion que la particule isolée. Au cours du test de stabilité thermique, l'échantillon est soumis à des cycles de montées en température successifs et sa diffusion est enregistrée à chaque fois, et le rayon hydrodynamique moyen est calculé. Quand la température où l'échantillon commence à se dénaturer est atteinte, les valeurs du rayon hydrodynamique augmentent rapidement. Cela permet de déterminer à quelle température maximale réaliser les réactions enzymatiques.

Pour la stabilité au cours du stockage de l'enzyme, on effectue une mesure cinétique de référence de l'enzyme fraîchement purifiée, dans les conditions optimales déterminées lors des étapes 2 et 3, et on répète cette mesure dans les mêmes conditions après différents temps de stockage (1 semaine, 3 semaines, 1 mois, 2 mois,...) et pour différentes conditions de stockage (température de la pièce, +4°C, -20°C). Cela permet de déterminer les conditions optimales de stockage de l'enzyme ainsi que sa sensibilité à la congélation. Dans certains cas, l'adjonction d'un pourcentage de glycérol permet de préserver l'activité de certaines enzymes particulièrement sensibles à la congélation/décongélation. Cependant, le glycérol est aussi connu pour être un inhibiteur des glycosides hydrolases. Si son emploi est nécessaire, il faudra au préalable réaliser

des tests d'activité afin d'établir son effet sur le profil cinétique de l'enzyme.

2. Température optimale de l'enzyme

Suite à la mesure DLS, la gamme de températures à tester encadre généralement celle à laquelle l'enzyme commence à se dénaturer. En effet, une enzyme plus flexible est souvent plus efficace, mais cela au détriment de sa stabilité dans le temps. Cependant, pour des réactions relativement courtes dans le temps, on aura tendance à privilégier une activité maximale et donc une température plus élevée afin de réduire les temps de réaction. Des cinétiques sont réalisées aux différentes valeurs de températures choisies. Le profil d'activité ainsi obtenu en fonction de la température possède un maximum relatif. Afin d'affiner la valeur de température optimale, des cinétiques sont effectuées pour une seconde série de températures, aux alentours du maximum relatif obtenu lors de la première série de mesures.

3. pH optimal

Une solution tampon universelle Teorell permet de déterminer le pH optimal des enzymes (Östling and Virtama 1946a). Grâce à un mélange de différents sels, elle permet de couvrir toute la gamme des pH de 2 à 12. On commence donc par réaliser des cinétiques à différents points de pH assez espacés. De même que pour la température, la courbe d'activité tracée présente un maximum relatif, qui permet ensuite de restreindre la gamme de pH à tester afin d'affiner la valeur de l'optimum. Une fois la valeur optimale en tampon Teorell déterminée, de nouvelles cinétiques doivent être réalisées dans des solutions tampons standards, efficaces dans une gamme compatible avec la valeur de pH optimal déterminée. Par exemple, pour un pH optimal de 6, on peut tester les tampons MES, PIPES, Na-phosphate. Il faut cependant faire attention, car certains tampons, notamment le Bis-Tris et l'HEPES, sont incompatibles avec le dosage des sucres réducteurs.

4. Les paramètres cinétiques

Il s'agit ici de se placer dans des conditions de cinétique Michaelienne afin de pouvoir déterminer le couple de constantes cinétiques V_{\max} et K_m et surtout la constante de spécificité k_{cat}/K_m qui caractérise chaque enzyme, k_{cat} correspondant au rapport $V_{\max}/\text{concentration d'enzyme}$.

En effet, pour que soit vérifiée l'équation de Michaelis-Menten,
$$V_i = V_{max} \frac{[S]}{[S] + K_m}$$

il faut que la réaction enzymatique soit réalisée dans des conditions d'état stationnaire et que le réactif soit en large excès par rapport à l'enzyme.

Il faut donc commencer par déterminer la concentration d'enzyme adaptée à la gamme de concentration de substrat envisagée, qui permette d'obtenir une cinétique linéaire sur la durée de la cinétique. La vitesse initiale de la réaction (V_i) correspondra ainsi au coefficient directeur de la droite obtenue par la jonction des points de cinétique. Une fois la concentration d'enzyme déterminée, on calcule les valeurs de V_i pour différentes concentrations de substrat, et ce en triplicata a minima. Ces valeurs, de même que les concentrations de substrat associées, sont ensuite entrées dans le logiciel Hyper32, développé par John Easterby à l'Université de Liverpool (<http://hyper32.software.informer.com/download/>). Ce programme réalise des représentations graphiques du jeu de données et calcule les valeurs de V_{max} et K_m associées selon quatre modélisations mathématiques, dont la représentation en double inverse de Lineweaver-Burk et un calcul de régression hyperbolique. La représentation par régression hyperbolique possède au moins deux avantages, comparée à celle de Lineweaver-Burk : 1) les valeurs de V_i obtenues pour les concentrations en substrat les plus faibles n'ont pas d'impact plus important sur la représentation graphique que celles obtenues aux concentrations les plus élevées ; 2) le calcul par régression hyperbolique fournit une approximation de l'erreur entre le modèle mathématique et les valeurs mesurées.

Cette stratégie à moyen débit aura permis le clonage d'un grand nombre de gènes en parallèle, augmentant ainsi le nombre de possibilités pour parvenir jusqu'à la résolution structurale et à la caractérisation biochimique de protéines au cours de cette thèse. Des points critiques ont également pu être identifiés, comme le choix des enzymes de restriction ou les bornes des séquences à cloner. Les étapes suivantes, aboutissant à la caractérisation de nouvelles enzymes, nécessitent quant à elles un traitement au cas par cas de chaque enzyme. Et bien que certaines étapes soient automatisables (cribles de cristallogénèse par exemple), l'essentiel du travail reste manuel. Ceci explique l'écart croissant entre le nombre de séquences disponibles dans les bases de données génétiques et le nombre d'enzymes caractérisées d'un point de vue biochimique et *a fortiori* structural.

Chapitre 2

La comparaison de nouvelles structures cristallines de kappa-carraghénases révèle des informations clés sur les liens entre la structure et le mode d'action des enzymes de cette sous-famille des GH16

Chapitre 2 : La comparaison de nouvelles structures cristallines de kappa-carraghénases révèle des informations clés sur les liens entre la structure et le mode d'action des enzymes de cette sous-famille des GH16.

Le κ -carraghénane est un polysaccharide marin sulfaté spécifique de la paroi de certaines macroalgues rouges. Ses propriétés gélifiantes en font un produit très recherché par l'industrie des hydrocolloïdes pour des applications principalement alimentaires. Cependant, de nombreuses études ont montré que, *via* ses motifs sulfatés, le κ -carraghénane possède des propriétés biomimétiques très intéressantes dans le cadre des phénomènes de reconnaissance et d'interactions cellulaires (infections bactériennes et virales, régulation de la prolifération cellulaire, ...). Ces effets biologiques sont souvent liés à la taille mais aussi aux motifs de sulfatation des fragments de carraghénane. Pour pouvoir générer des fragments de taille contrôlée, les enzymes de type glycosides hydrolases sont des outils moléculaires d'une grande spécificité. Or, jusqu'à présent, une seule κ -carraghénase était caractérisée d'un point de vue biochimique et structural, à savoir *PcCgkA_{GH16}* de *Pseudoalteromonas carrageenovora*.

Au cours de cette thèse, nous avons résolu la seconde structure d'une κ -carraghénase, *ZgCgkA_{GH16}* de *Zobellia galactanovorans*. Sa comparaison avec la structure précédemment obtenue de la première κ -carraghénase en complexe avec un néo- κ -carratétraose a permis de définir des éléments structuraux caractéristiques représentatifs d'une partie de la phylogénie des κ -carraghénases. Cette comparaison structurale a également permis d'expliquer les différences d'activité observées au niveau biochimique entre ces deux enzymes. En effet, la κ -carraghénase de *Z. galactanivorans* présente une efficacité catalytique très supérieure à celle de *P. carrageenovora* sur un substrat soluble, et elle génère principalement des hexa et tétrasaccharides, alors que *P. carrageenovora* génère principalement des tétra et disaccharides. Ces différences de tailles mais aussi d'efficacité peuvent s'expliquer au niveau structural par une modification du nombre et de la répartition des sous-sites catalytiques ainsi que de la flexibilité de l'enzyme. Ces deux facteurs sont eux-mêmes liés aux différences de répartition des éléments structuraux qui constituent le tunnel catalytique, caractéristique du repliement de ces enzymes. Une étude biochimique comparée sur substrat soluble et sur microgels de κ -carraghénane a également permis de démontrer le caractère

processif de ZgCgkA_{GH16} sur substrat gélifié, caractéristique qui n'avait été démontrée à ce jour que chez PcCgkA_{GH16}. De plus, une autre originalité de la κ -carraghénase de *Z. galactanivorans* est de posséder un module CBM à la suite du module catalytique ZgCgkA_{GH16}. L'étude comparée en solution et en microgel a permis de montrer que ce module participe fortement à la processivité de l'enzyme, en gel comme en solution. Enfin, une étude par mutagenèse dirigée sur PcCgkA_{GH16} a permis de mieux comprendre le rôle de certains résidus clés dans l'interaction avec un substrat hétérogène comme le carraghénane, en particulier l'importance des résidus chargés, comme les arginines et des résidus aromatiques.

L'ensemble de ces résultats a fait l'objet d'un article soumis et accepté dans la revue scientifique JBC. La version finale de l'article se trouve dans les pages suivantes.

Structural insights into marine carbohydrate degradation by family GH16 kappa-carrageenases

Maria Matard-Mann^{1,4}, Thomas Bernard², Cédric Leroux³, Tristan Barbeyron¹, Robert Larocque¹, Aurélie Préchoux¹, Alexandra Jeudy³, Murielle Jam¹, Pi Nyvall Collén⁴, Gurvan Michel¹ and Mirjam Czjzek^{1*}

¹Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, Roscoff, Bretagne, France.

²Architecture et Fonction des Macromolécules Biologiques, Unité Mixed de Recherche 6098, Centre National de Recherche Scientifique, Universités Aix-Marseille I & II, Case 932, 163 Avenue de Luminy, 13288 Marseille cedex 9, France

³Sorbonne Universités, UPMC Univ Paris 06, CNRS, FR 2424, Station Biologique de Roscoff, F-29682, Roscoff, Bretagne, France.

⁴Amadéite SAS, « Pôle biotechnologique » du Haut du Bois, 56580 Bréhan, France

Running title: Structure and function of GH16 kappa-carrageenase enzymes

*Corresponding author: E-mail: czjzek@sb-roscoff.fr Tel +(33) 298 29 23 75, Fax +(33) 298 29 23 24

Keywords: kappa-carrageenan, enzyme/substrate complex, carbohydrate recognition, sulfated polysaccharides, crystal structure

ABSTRACT

Carrageenans are sulfated α -1,3- β -1,4-galactans found in the cell wall of some red algae that are practically valuable for their gelation and biomimetic properties but also serve as a potential carbon source for marine bacteria. Carbohydrate degradation has been studied extensively for terrestrial plant/bacterial systems, but sulfation is not present in these cases, meaning the marine enzymes used to degrade carrageenans must possess unique features to recognize these modifications. To gain insights into these features, we have focused on κ -carrageenases from two distant bacterial phyla, which belong to glycoside hydrolase family 16 (GH16) and cleave the β -1,4 linkage of κ -carrageenan. We have solved the crystal structure of the catalytic module of ZgCgkA from *Zobellia galactanivorans* at 1.66 Å

resolution, and compared it with the only other structure available, that of PcCgkA from *Pseudoalteromonas carrageenovora* 9^T (ATCC 43555^T). We also describe the first substrate complex in the inactivated mutant form of PcCgkA at 1.7 Å resolution. The structural and biochemical comparison of these enzymes suggests key determinants that underlie the functional properties of this subfamily. In particular, we identify several arginine residues that interact with the polyanionic substrate, and confirmed the functional relevance of these amino acids using a targeted mutagenesis strategy. These results give new insight in the diversity of the κ -carrageenase subfamily. The phylogenetic analyses show the presence of several distinct clades of enzymes that relate to differences in modes of action or subtle differences within the same substrate specificity, matching the

hybrid character of the κ -carrageenan polymer.

INTRODUCTION

Marine ecosystems house highly complex trophic networks. At the basis of these networks, photosynthetic organisms play a key role. As primary producers, they initiate and maintain the biogeochemical carbon cycle in the ocean and constitute the first link in the marine food chain. Among them are found macroalgae, which grow in the photic zone. Macroalgae are polyphyletic and belong to three distinct phyla, namely red, green and brown algae. In addition to their role in climate regulation (O'Dowd *et al.* 2002) and as shelter for many animal and microbial marine species (Duggins *et al.* 1989), they are a major source of carbon for many communities. Particularly, they establish close interactions with marine bacteria living freely in the water column or entrapped in biofilm on the surface of macroalgae (Martin *et al.* 2014). Increasing interest has arisen for the study of the interactions between microbial communities and their associated host, especially in the case of terrestrial association in human gut (Larsbrink *et al.* 2014).

Heterotrophic bacteria have developed sophisticated enzymatic tools, in order to extract the carbon from plant cell walls. Notably, the ecological position of a bacterial strain in the microbial communities will influence the degradation strategy deployed. Some possess a wide set of enzymes that make them generalist organisms, like *Bacteroidetes*, others being more specialized in the degradation of certain families of polysaccharides (Grondin *et al.* 2017). As an example of an integrative strategy, *Bacteroidetes* species have been shown to usually possess all the genes involved in the

degradation, recognition and transport of a specific polysaccharide organized in clusters called polysaccharide utilization loci (PUL) (Bjursell *et al.* 2006; Martens *et al.* 2008).

Yet, the majority of studies are done on terrestrial bacteria, producing enzymes against land plant polysaccharides. In the case of marine heterotrophic bacteria, their arsenal of carbohydrate active enzymes are still largely unexplored (Michel and Czjzek 2013). Indeed, they are adapted to the diversity of macroalgal polysaccharides and many of these marine polysaccharides have no equivalent in land plants. One of their peculiarities is the abundance of sulfation motifs, which are absent in the land plant polysaccharides, but also present in polysaccharides of the mammalian extra cellular matrix (Senni *et al.* 2011).

One of the most well-known families of algal sulfated polysaccharides is carrageenan. This family is composed of linear chains of α -1,3- β -1,4-galactans with an increasing number of sulfate groups per disaccharide, unit depending on the categories of carrageenans. Accordingly, the three main structures are called kappa- (κ -, one sulfate group), iota- (ι -, two), and lambda- (λ -, three) carrageenans (Figure 1) (Knutsen *et al.* 1994). The carrageenases, which cleave the β -glycosidic linkages of carrageenans, belong to different families of glycoside hydrolases (GH), based on their amino acid sequences: GH16 for κ -carrageenases (Barbeyron *et al.* 1998; Michel *et al.* 2001a), GH82 for ι -carrageenases (Michel *et al.* 2003; Rebuffet *et al.* 2010), and a new family of GH for λ -carrageenases (Potin *et al.* 1991; Guibet *et al.* 2007). In this study, we will focus on the family of κ -carrageenases.

During the past few years, many new κ -carrageenases have been identified in recently sequenced bacterial genomes (244 entries in UniProtKB), belonging mainly to the three

phyla *Bacteroidetes*, *Proteobacteria* and *Planctomycetes*. However, only seven have been cloned and biochemically studied, namely from *P. carrageenovora* 9^T (ATCC 43555^T) (Barbeyron *et al.* 1994), *Zobellia galactanivorans* Dsij^T (Barbeyron *et al.* 1998), *P. porphyrae* LL1 (Liu *et al.* 2011), *P. tetraodonis* JAM-K142 (Kobayashi *et al.* 2012), *Zobellia* sp. ZM-2 (Liu *et al.* 2013), *Shewanella* sp. P1-14-1 (Kz7) (Wang *et al.* 2015) and *Pseudoalteromonas* sp. QY203 (Xu *et al.* 2015). Up to now, only one κ -carrageenase has been biochemically and structurally studied at 1.54 Å resolution, namely *PcCgkA_{GH16}* from the gammaproteobacterium *Pseudoalteromonas carrageenovora* 9^T (Pc) (Michel *et al.* 2001b) that was isolated from the marine water column and algae at the Atlantic coast of Canada (Yaphe and Baxter 1955; Weigl and Yaphe 1966).

We report here the second crystal structure of a catalytic module of a GH16 κ -carrageenase, *ZgCgkA_{GH16}* from *Zobellia galactanivorans* Dsij^T, a marine flavobacterium model for the bioconversion of algal polysaccharides that was isolated from the surface of the red macro-algae *Delesseria sanguinea* (Potin *et al.* 1991). In contrast to *P. carrageenova*, *Z. galactanivorans* grows perfectly well with κ -carrageenan as sole carbon source (Barbeyron *et al.* 2016). Both enzymes do not require any desulfation of natural κ -carrageenan substrate to be active. Despite the similarity of sequences between *ZgCgkA_{GH16}* and *PcCgkA_{GH16}* (40 % identity, 50% similarity), several areas of divergence were observed, corresponding to the deletion or insertion of sequence stretches (Figure 2A). The biochemical and structural comparison of both enzymes, along with a phylogenetic analysis of κ -carrageenases, give a first insight into the diversity existing within this

subfamily. This allowed us to propose some structural determinants that underlie the biochemical properties of these enzymes, active on polyanionic substrate. In addition, the mutagenesis and analysis of kinetic properties of *PcCgkA_{GH16}* point mutants allowed proposing potential targets of interest for enzymatic engineering of more efficient κ -carrageenases.

RESULTS

Phylogenetic analysis of κ -carrageenases - A Blast search with *PcCgkA_{GH16}* as query identified 76 putative κ -carrageenases in GenBank (identity ranging from 30 to 97 %). All these sequences have conserved several key residues for the recognition of κ -carrageenan (R196, R260 in *PcCgkA_{GH16}*). A phylogenetic tree was built (Figure 3) based on the multiple alignment of the catalytic module sequences of these 76 κ -carrageenases and of 3 GH16 β -agarases chosen as outgroup (Figure S1).

This phylogenetic tree can be divided into at least ten major clades, based on the highest values of bootstrap of deepest nodes. Four major clades can be delineated. Clade A, which contains the κ -carrageenase from *P. carrageenovora*, and clade D correlate well with bacterial phyla from *Proteobacteria*, with a bootstrap value of 100 % for both. Clade B is composed of two subgroups of sequences belonging to *Bacteroidetes* and *Planctomycetes*, with a bootstrap value of 98 %. Clade C contains sequences from *Bacteroidetes* only, including *ZgCgkA_{GH16}*, with a bootstrap value of 100 %.

It is noteworthy that some bacterial strains contain several κ -carrageenase sequences that spread out throughout the whole phylogenetic tree, as exemplified by *Algibacter* sp. SK-16 which displays 6 sequences (highlighted by arrows in Figure 3; specific alignment provided in Figure S2) or *Flammeovirga* sp. OC4 with 3 sequences.

Interestingly, when looking at the specificities of the two clades A and C, we can notice that the presence or absence of sequence-stretches in the alignment correlates with the phylogenetic clustering. Specific alignment of both *PcCgkA_{GH16}* and *ZgCgkA_{GH16}* allowed the definition of these sequence-stretches more precisely as “fingers”, numbered from F1 to F6. They are boxed on the sequence alignment of Figure 2A. Their delimitation is based on analysis of secondary structure elements. These ‘fingers’ are generally constituted by a succession of a [β -strand - tight turn/loop or helix - β -strand] motive and were identified as regions containing critical divergence in the amino acid sequences of both κ -carrageenases studied here (green boxes in Figure 2A).

All the enzymes belonging to clade A possess a “finger F2-type” motif which is absent in clade C and all the enzymes from clade C possess a “finger F4-type” motif, absent in clade A enzymes. Conversely, clade B and D enzymes possess both types of motifs, clade D containing the most conserved sequences for both fingers, and clade B showing modifications of the nature and of the length of ‘typical’ sequences.

The remaining sequences of the phylogenetic tree possess bootstrap values that are too low to constitute a monophyletic clade and also for allowing the establishment of clear correlations between the sequences and structural features.

Another interesting feature observed in the phylogenetic tree of κ -carrageenases is that some enzymes possess a modular architecture (e.g. *Z. galactanivorans*), while others consist of catalytic modules only (e.g. *Z. uliginosa*, *P. carrageenovora*). Notably, the modular nature is not correlated with the clades, except for the clade B *Planctomycetes* sequences, which are all only composed of a catalytic module.

Biochemical characterization of ZgCgkA_{GH16} and ZgCgkA_{GH16}-CBM16-PorSS and comparison with PcCgkA_{GH16} - The κ -carrageenase from *Z. galactanivorans*, *ZgCgkA*, is a modular enzyme displaying a N-terminal GH16 catalytic module, a carbohydrate binding module of the family 16 (CBM16) and a C-terminal Type-IX secretion module (PorSS) (Barbeyron *et al.* 1998, 2016). Two forms of the enzyme were cloned from genomic DNA, the catalytic module alone (*ZgCgkA_{GH16}*) and the entire form of the enzyme (*ZgCgkA_{GH16}-CBM16-PorSS*). *PcCgkA* is composed of a GH16 catalytic domain and an Ig-like module of unknown function. Only the catalytic module (*PcCgkA_{GH16}*) was previously cloned (Barbeyron *et al.* 1994) and has been used in this study. The three proteins were successfully expressed in soluble form and purified to homogeneity.

The two forms of *ZgCgkA*, with and without CBM16, were studied in parallel to the recombinant enzyme *PcCgkA_{GH16}*. At first, we determined the optimal conditions for pH, buffer and temperature for both forms of *ZgCgkA*. Efficiency was highest at pH 6 in MES buffer for both constructions, and analysis by Dynamic Light Scattering (DLS) revealed that both forms are stable up to 40 °C (data not shown). The comparison of the kinetic parameters of the three enzymes on diluted solutions of κ -carrageenan shows that *ZgCgkA_{GH16}-CBM16-PorSS* and *ZgCgkA_{GH16}* are more efficient than *PcCgkA_{GH16}* by 3 and 5 times, respectively (Table I). The absence of CBM in the construct of *ZgCgkA_{GH16}* causes a 40 % increase of efficiency in solution, compared to the entire enzyme.

Degradation kinetics on a longer time course (5 h) gave a similar result in solution (Figure 4A). The concentration of product released after 30 minutes of incubation was increased by 26 and 143 times for *ZgCgkA_{GH16}-CBM16-PorSS* and *ZgCgkA_{GH16}*, respectively, as

compared to *PcCgkA*_{GH16}. Thus *ZgCgkA*_{GH16} is 5.5 times more efficient in solution than *ZgCgkA*_{GH16-CBM16-PorSS}. However on microgels (Figure 4B), differences of released products between the three enzymes are far less important than reactions in solution. Under these conditions, subtle differences can be observed: *ZgCgkA*_{GH16-CBM16-PorSS} releases 2 times more product than *PcCgkA*_{GH16} and 1.5 more than *ZgCgkA*_{GH16} after 90 minutes of incubation.

HPAEC analyses of reaction end products produced by the recombinant enzymes confirms the previously reported difference of product patterns observed for the native enzymes (Potin *et al.* 1991). Namely, *ZgCgkA*_{GH16} mainly produces oligosaccharides with a degree of polymerization (DP) 4 and 6 (Figure S3), while *PcCgkA*_{GH16} produces oligosaccharides with DP2 and DP4 (data not shown). Moreover, for both constructs of *Z. galactanivorans* κ -carrageenase, the profile of degradation products is different in solution and in microgels, with a ratio DP4/DP6 always higher in microgels compared to that in solution (Table II). But the presence of the CBM modifies the values of this ratio. In solution (Figure 4C-E), the ratio becomes greater than 1 more quickly for *ZgCgkA*_{GH16-CBM16-PorSS} than *ZgCgkA*_{GH16} (after 45 min and 90 min respectively), and on microgels (Figure 4D-F) it reaches highest levels with *ZgCgkA*_{GH16-CBM16-PorSS} after one week of extensive digestion. This means that the construction with CBM (*ZgCgkA*_{GH16-CBM16-PorSS}) favors the release of DP4, mainly in microgels, but also in solution, to a higher extent than the construction without CBM (*ZgCgkA*_{GH16}).

Crystal structure of the catalytic module of Z. galactanivorans κ -carrageenase (*ZgCgkA*_{GH16}) - *ZgCgkA*_{GH16} was crystallized and the three-dimensional structure determined at 1.66 Å resolution (Figure 2B) by molecular replacement, using the native

*PcCgkA*_{GH16} structure as a model (PDB code: 1DYP) (Michel *et al.* 2001a). The asymmetric unit contains four copies of the protein, covering the same sequence from Q30 to S307. The RMSD values, calculated for C-alpha atoms, between the four copies are low, ranging from 0.21 and 0.31 Å, and throughout the following, we will thus refer only to molecule A, except if otherwise specified. B factors of the overall chain (C- α and lateral chains) have been calculated for the different fingers defined upward as the extensions of β -strands and/or loops that flank the catalytic tunnel (Table III). All further data statistics are summarized in Table IV.

Crystal structure of the inactivated catalytic module of the κ -carrageenase from P. carrageenovora (PcCgkA_{GH16-E168D}) in complex with a κ -neocarratetraose - In order to identify the key residues involved in the interaction with the sulfated substrate, we designed an inactive mutant of the κ -carrageenase from *P. carrageenovora* (*PcCgkA*_{GH16-E168D}) in order to trap a stable complex enzyme-substrate. As a member of GH16 family, *PcCgkA* possesses the characteristic catalytic triade ExDxxE responsible for the double displacement mechanism which results in the cleavage of the β -1,4 linkage with retention of the anomeric configuration (Koshland 1953; Zechel *et al.* 2000; Planas 2000). E163 is the nucleophile residue which binds the D-galactose-4-sulfate (G4S) unit at C1 to form the glycosyl-enzyme intermediate, and E168 is the acid/base catalyst. D165 is also involved in the mechanism by accelerating the deglycosylation step (Michel *et al.* 2001a). Site directed mutation of the catalytic E168 to aspartate resulted in a significant loss of enzymatic activity, although not abolished totally (data not shown). This allowed the co-crystallization of the recombinant enzyme in presence of a mixture of κ -carrageenan oligosaccharides, purified from the digestion

of κ -carrageenan polymer with *PcCgkAGH16*. The crystals obtained allowed to solve the three-dimensional structure at 1.7 Å resolution (Figure 2C) by molecular replacement again using the native *PcCgkAGH16* as a model. The asymmetric unit contains two copies of the protein, covering residues from M28 to V297. The Fo-Fc difference density for the complex crystal that we obtained unambiguously showed a tetrasaccharide spanning the active site cleft in subsites from -4 to -1 (subsite numbering according to Davies *et al.* (Davies *et al.* 1997)), and the oligosaccharide could readily be built into the density (Figure 5A). Nineteen residues have been identified as involved in substrate binding (Figure 5B). Bonding distances are listed in Table V. 3,6-anhydro-D-galactose (DA) moieties in subsites -4 and -2 interact respectively with 3 and 5 lateral chains of the catalytic tunnel residues. Comparatively, D-galactose-4-sulfate (G4S) moieties in subsite -3 and -1 show the highest number of interactions with the enzyme, 12 and 13 respectively. Remarkably, R260 is involved in 5 interactions in -1 subsite, and out of the 19 interacting residues, the 3 arginines are responsible for 12 out of the 32 total interactions between the enzyme and its substrate. It can also be underlined that the two sulfate functions of the substrate are particularly implicated in the network of hydrogen bonding, as underlined in figure 5B.

By comparison with the available structure of the enzyme without substrate (PDB code: 1DYP), we can notice that the general fold of the enzyme displays no significant difference (RMSD= 0.48 and 0.46 Å for molecules A and B respectively). However, some side chains show translational movements, between 1.2 and 1.6 Å, and others show marked rotations, between 90° and 180° (Figure 5D). The residues affected by these movements are mainly located in finger F2 above subsite -4 of the catalytic channel for the translational movements, namely W95,

Q100 and Q102 (straight arrows in figure 5D), and in fingers F5-F6, which constitute the closed part of the tunnel between the -2 and -1 subsites, for the rotational movements (curved arrows). In particular, R196 and N269 adopt a “closed” conformation in the presence of the substrate, with their functional groups pointing into the channel, with an overall displacement of 3.6 Å when compared to the structure without substrate. Interestingly, R151 adopts two alternative conformations when the substrate is present as compared to the apo-structure of the enzyme, one pointing towards the -4 subsite, the second pointing towards the -3 subsite (Figure 5C). The B factor of this arginine (20.0 Å²) is also higher than the overall mean B factor.

Structural comparison of ZgCgkAGH16 and PcCgkAGH16-E168D - As a member of the GH16 family, *ZgCgkAGH16* adopts a β -jelly roll fold. Similar to *PcCgkAGH16-E168D*, the catalytic channel of *ZgCgkAGH16* is partially closed forming a tunnel, constituted by the junction of fingers F5 and F6 (Figure 6A) through R199 and N272, the equivalent of R196 and N269 in *PcCgkAGH16-E168D* (Figure S4). Despite the conservation of the overall fold, four main structural differences can be underlined when superposing their structures (Figure 6B).

At first, substantial differences concern the extremity of the catalytic channel containing the negative subsites (-4 and -3 subsites). In *PcCgkAGH16-E168D*, the top finger F2 establishes a direct link to the substrate molecule through three amino acids, W95, Q100 and Q102. Residues of this finger also take part, together with finger F1, in binding a dense network of 34 water molecules that wraps around the oligosaccharide bound in the catalytic channel from subsites -4 to +2 (Figure S5). Finger F2 is absent in *ZgCgkAGH16*, as well as C98 and C268, which form the disulfide bond between F2 and F6 in *PcCgkAGH16-E168D*. The dense water molecule

network, described above, is also far from being as extended as that in *ZgCgkAGH16*, being confined to only 7 water molecules located between subsites -1 to +2. It can also be noticed that the key basic residue R151 in *PcCgkAGH16-E168D*, located on finger F3 and that adopts two alternative conformations, is replaced by an aromatic residue (Y148) in *ZgCgkAGH16*.

The second main difference between both enzymes is the presence of an additional finger, F4, at the extremity of the tunnel forming positive subsites for *ZgCgkAGH16*, in the place of a key residue, K172, identified in *PcCgkAGH16-E168D* (Michel *et al.* 2001a). This finger F4 is composed of the loop F168-D177 and contains a combination of aromatic and charged residues (Figure 2A). Notably, W170 forms a hydrophobic platform that extends the end of the binding cleft (Figure S6). The distance of 2.8 Å between D169 in finger F4 and K201 in finger F5 is compatible with a hydrogen bond. Altogether these features partially obstruct the base level of the tunnel exit, where the positive subsites are located in *PcCgkAGH16-E168D*, forcing the substrate into an orientation at 45° with respect to the substrate chain orientation on the side of negative subsites (Figure 6C).

A third difference is the presence of a helix ($\alpha 3$ in Figure 2A) on the very top of *ZgCgkAGH16*, in the prolongation of finger F6, just above the positive binding subsites of the catalytic cleft (Figure 6B). Notably two lysine residues, K284 and K288, contribute to a basic patch, together with K264, located just between this helix and finger F4. Altogether, finger F4, the top helix and the charged side chains, create a unique basic environment located at the positive subsites of *ZgCgkAGH16* (Figure S6).

A fourth major difference between the two catalytic modules is the increased values of B factors of fingers F1 and F6 in *ZgCgkAGH16* compared to those of *PcCgkAGH16-E168D* (Table III, Figure S4). For finger F1, the

relative value of the B factor with respect to the overall value changes by 49 % in *PcCgkAGH16-E168D* (81 % in *PcCgkAGH16*) and up to 119 % in *ZgCgkAGH16*. This is associated with an extension of finger F1 by two residues in *ZgCgkAGH16*, resulting in a spatial displacement with respect to the position in *PcCgkAGH16-E168D*. Notably, N70 is oriented toward the -3 subsite with a rotation of 45° in contrast to the equivalent N63 in *PcCgkAGH16-E168D*, which points in the opposite direction (Figure S4). Concerning finger F6, the relative B factor value changes by 89 % in *PcCgkAGH16-E168D* (88 % in *PcCgkAGH16*) and by 190 % in *ZgCgkAGH16* with respect to the overall B value. This increased mobility is especially marked for residue N272, which is equivalent to N269 of *PcCgkAGH16-E168D* that was previously identified for its implication in forming the tunnel above the catalytic site. We note that the side chain of this residue in *ZgCgkAGH16* is in the “open” position, as compared to what is observed in the apo-structure of *PcCgkAGH16* (PDB: 1DYP), and that the associated R199 (the equivalent of R196 in *PcCgkAGH16*) adopts two alternative conformations in the crystal structure of *ZgCgkAGH16*, one in the “open” and one in the “closed” position.

Site-directed mutagenesis of PcCgkAGH16 - The structure of *PcCgkAGH16-E168D* in complex with a κ -neocarratetraose allowed identifying nineteen amino acids directly involved in the interaction with the substrate (listed in Table V) in addition to some other residues more distant, and involved in the network of water molecules. Six of these residues have been successfully mutated individually into alanine (Figure 5C). The recombinant mutated proteins were purified and biochemically characterized in solution on κ -carrageenan substrate. Details of the determination of kinetic parameters are provided in Table SI.

Depending on the position of the mutated amino acid, the effect on the catalytic efficiency is very different (Table I). Indeed, R196A, R260A and W266A cause a dramatic loss of catalytic efficiency, between 95 and 99%, as compared to the wild type enzyme. For R196 and R260 the loss of activity can be explained by their key role in binding to the charged sulfate groups of G4S in the -1 subsite. The importance of both these residues and especially R260 in substrate binding was already detected by Michel *et al.*, based on a structural model of a docked substrate molecule in the catalytic active site of the native κ -carrageenase (Michel *et al.* 2001a). It is noteworthy that these basic residues are strictly conserved in all the sequences of κ -carrageenases, with a few exceptions, where R196 is replaced by a lysine, and thus must play an important role in the specificity of recognition of the substrate in this GH16 sub-family. In the case of our enzyme/substrate complex *PcCgkA*_{GH16-E168D}, we show that the importance of R196 is connected to its involvement in shaping the tunnel above the catalytic active site, and its concerted interaction with N269 in the binding of the substrate in subsites -1 and -2 (Table V, Figure 5B). We also confirm the crucial importance of R260 in stabilizing the substrate in subsite -1, notably through the interactions with the sulfate group, which is the hallmark of κ -neocarrabiose (Table V). As a consequence, the remaining activity of the protein with R260 mutated to alanine was less than 1 % of the wild type enzyme, and kinetic parameters were not determined. By contrast, W266 is positioned in the catalytic groove where the positive substrate binding sites must be, and consequently does not interact with the substrate molecule in the crystal structure of this substrate/enzyme complex. Nevertheless, the drastic loss of activity when mutated to alanine clearly indicates that it plays an important role in the

interaction with the saccharide units that bind to the positive sub-binding sites.

Surprisingly, R151A and Q171A have an opposite effect, causing an increase of efficiency by 1.4 to 1.8 times, respectively (Table I). R151 is located in between subsites -4 and -3 and establishes strong interactions with the substrate at the entrance of the catalytic groove, through a salt bridge and direct hydrogen bonds. On the other hand, Q171 is located in the putative +1 sub-binding site, where it could establish hydrogen bonds with the leaving group G4S. This residue is highly conserved in κ -carrageenases, so it may be linked with the thermodynamic equilibrium of enzymatic reaction. When mutated into alanine, the interaction with the leaving group is strongly decreased, allowing a more rapid product release into the medium, and thus accelerating the enzymatic turnover.

R92A is the only mutated residue not directly in contact with the substrate, but involved in binding water molecules at the entrance of the channel (Figure S5). The absence of any effect on catalytic efficiency when mutating R92 shows that this arginine does not seem to play a major role in substrate binding or mode of action, at least on soluble substrates.

DISCUSSION

Although family GH16 κ -carrageenases are frequently identified in macroalgae-associated bacteria (Martin *et al.* 2014, 2015), only seven enzymes have actually been produced (Barbeyron *et al.* 1998, 1994; Liu *et al.* 2011; Kobayashi *et al.* 2012; Liu *et al.* 2013; Wang *et al.* 2015; Xu *et al.* 2015) and only one structurally characterized to date (Michel *et al.* 2001a). In polyspecific GH families, such as the GH16 family, the subfamilies based on phylogenetic analyses usually correspond to distinct substrate specificities (Barbeyron *et al.* 1998; Davies *et al.* 1997; Aspeborg *et al.* 2012). Within a specific subfamily, the large diversity of sequences and the phylogenetic distribution

in clades is rather related to differences in modes of action or subtle differences within the same substrate specificity (Stam *et al.* 2006). It is important to mention that carrageenans occur as hybrid polymers within the cell walls of red algae, containing numerous variant motifs beyond the frequent repeating units κ - and ι -carrabioses (e.g. the precursors units μ - and η -carrabioses, the desulfated units β - (Knutser and Grasdalen 1987) and α -carrabiose (Falshaw *et al.* 1996; Usov 2011)). These complex structures are further modified by the addition of methyl or pyruvate groups. In this context, one can expect significant differences between κ -carrageenases and the results presented here are the first example of exploring this biochemical diversity.

A first obvious difference between *PcCgkA* and *ZgCgkA* is the modular architecture of these κ -carrageenases. Even if both are secreted enzymes, only *ZgCgkA* possesses a CBM (Potin *et al.* 1995; Barbeyron *et al.* 1998). The presence of CBMs appended to catalytic GH modules is in general related to the capacity of the adjacent enzyme to tackle recalcitrant, complex substrates in the solid form, as encountered in the context of plant cell wall (Tomme *et al.* 1988; Din *et al.* 1991; Black *et al.* 1996). Four main roles are attributed to CBMs: 1) to allow proximity between the catalytic domain and its substrate, 2) to target the catalytic module to specific parts of the cell-wall, 3) to help disrupt the substrate organization, 4) to anchor the enzyme to the bacterial cell-wall (Boraston *et al.* 2004; Gilbert *et al.* 2013). The isolated CBM16 module of *ZgCgkA* was cloned and expressed as a recombinant protein in an independent study, and was indeed shown to bind κ -carrageenan in ELISA tests (Salmean *et al.*, in preparation), indicating that this secreted enzyme most probably degrades semi-crystalline κ -carrageenans within the algal cell wall. Moreover, contrary to *PcCgkA*, DP6 is not hydrolyzed by *ZgCgkA*, which confirms what

Potin *et al.* (Potin *et al.* 1991) suggested, that is to say DP8 is the minimal size of oligosaccharide that can be degraded by *ZgCgkA*. According to the displacement scheme proposed by Lemoine *et al.* (Lemoine *et al.* 2009) for *PcCgkA*, DP6 is exclusively produced by random processing, whereas DP4 is produced by random and processive modes of action. Similarly we propose to use the DP4/DP6 ratio as a marker of processivity for *ZgCgkA*. In this study, comparison of behavior of recombinant *ZgCgkA* with and without CBM, and also on microgels compared to κ -carrageenan solutions, suggests that *ZgCgkA* is an endo-processive enzyme, mainly on microgel, and that its appended CBM16 contributes to the processivity, even on soluble substrate, probably by favoring and maintaining the proximity between catalytic module and substrate. The difference of processivity also explains the difference of catalytic efficiency observed with and without CBM16, owing to the fact that the selective advantage of processivity on insoluble substrate is in balance with the loss of catalytic efficiency on soluble substrate (Horn *et al.* 2006a). This phenomenon has already been observed for processive endoglucanases, where deletion of the CBM associated to the catalytic module leads to an increased activity on soluble substrate (Gilad *et al.* 2003).

Dissection of the molecular details of protein-carbohydrate interactions within the catalytic active site is crucial to understand the enzymatic efficiency, the mode of action and the substrate specificity of a given enzyme class. From our structural and mutagenesis study, we can establish several key parameters in the κ -carrageenase sub-family:

- The importance of arginine residues: first underlined in the 70's, with the study of carboxypeptidase A (Riordan and Hayashida 1970), the positive charge and basic character of arginines allow them to establish salt bridges and/or multiple hydrogen bonds with negatively charged substrates like phosphate

groups (Riordan *et al.* 1977) or sulfated polysaccharides, as observed in the case of chondroitin lyase (Huang *et al.* 2001).

- The importance of enzymatic flexibility and substrate binding in catalytic efficiency: loss of finger F2 and the stabilizing disulfide bond between F2 and F6, as well as the associated water network, result in a more flexible F6 finger that correlates well with the fivefold increase of enzymatic efficiency of Zg CgkAGH16 in solution compared with PcCgkAGH16. Interestingly, mobility of fingers also seem to be inversed between both enzymes, since finger F6 at the top of the groove in ZgCgkAGH16 has 2-times higher B-factors than overall, while fingers F3 and F5 of PcCgkAGH16 display two-times higher than average values (Table III). Moreover, the units at the reducing end of the oligosaccharide might be more tightly bound in ZgCgkAGH16 (at least four positive sub-binding sites: F4-5-6) than in PcCgkAGH16 (two positive sub-binding sites: F5-6), and less tightly bound at the non-reducing end, with only three negative sub-binding sites (F1-3) instead of four (F1-2-3). These findings, along with the experiments on microgels, lead us to assume that ZgCgkAGH16 is an endoprocessive enzyme with at least seven sub-binding sites, moving from the reducing end towards the non-reducing end of a κ -carrageenan chain, movement opposite to that proposed for PcCgkAGH16 (Figure 6D)(Michel *et al.* 2001a).

- Processivity on charged substrates necessitates a compromise between affinity and mobility, which can be fulfilled by a combination of basic and mobile residues, such as arginines and lysines, associated with aromatic residues (R151/W95 for PcCgkAGH16; K201/W170 for ZgCgkAGH16). The alternative conformations of the side chain of R151 observed in presence of the

ligand, together with the relative high B factor of this residue lead us to assume it might assist the establishment of a first contact to the substrate, ‘grabbing’ a sulfate group, but also assist in sliding the substrate further, after a catalytic cleavage. We can further assume that the mutation of R151 into an alanine induces a weakening of the substrate binding at the non-reducing end of the polysaccharide chain, possibly hindering the processivity and thus leading to an increase of the catalytic efficiency on a soluble substrate. The effect of this point-mutation can thus be compared to what is observed for the processive cellobiohydrolase of *T. reesei*, where the deletion of one of the active-site loops leads to an increased efficiency of the enzyme on amorphous cellulose (Von Ossowski *et al.* 2003).

- The importance of subsites associated with the leaving group: a balance is needed between the necessity of maintaining the stability of the enzyme-substrate complex before cleavage and facilitating the release of the product. In PcCgkAGH16 Q171 is possibly involved in a process defined as “product inhibition” by binding the product tightly to the enzyme, a feature which is frequently observed in processive enzymes active on crystalline substrates, such as cellulases or chitinases (Watanabe *et al.* 2003; Horn *et al.* 2006a). We can assume a similar function for N70 in ZgCgkAGH16. Conversely, W266 might facilitate the sliding of the released product, thus leading to a loss of enzymatic efficiency when mutated into alanine. Indeed, hydrophobic amino acids and tryptophanes in particular, are well known for their role in facilitating the sliding of oligosaccharide substrates in glycoside hydrolases (Divne *et al.* 1998). In ZgCgkAGH16, an equivalent residue could be F146.

To summarize, our data indicate that despite sharing a global common ancestor, *PcCgkA_{GH16}* and *ZgCgkA_{GH16}* have evolved unique structural features, which shape distinct tunnel topologies and likely result in opposite directions of processivity. Interestingly, in the point of view of industrial applications, the enzyme from *Z. galactanivorans* is 5 times more efficient in solution. The two GH16 enzymes described here are representatives of different clades of κ -carrageenases (clades A and C, respectively, Figure 3). While clade C enzymes clearly originate from marine *Bacteroidetes* and clades A and D from *Proteobacteria*, we cannot currently determine whether the common ancestor of clade B κ -carrageenases belonged to the *Bacteroidetes* or *Planctomycetes* phyla. The exact taxonomic nature of the common ancestor of the entire κ -carrageenase subfamily is even more difficult to establish, considering the low bootstrap values of the deep nodes of the remaining sequence clusters. Remarkably, some marine bacteria have multigenic families of κ -carrageenases, as for example *Algibacter* sp SK-16 that possesses 3 clade B sequences, 1 clade C sequence and two other sequences (including the current deepest sequence of the κ -carrageenase subfamily). This diversity of enzymes active on κ -carrageenan within one organism most likely results from various evolutionary events, such as horizontal gene transfers or sequence duplications followed by diverging evolution. The presence of this multigenic family might as well give an adaptive advantage to this species to face the chemical complexity of carrageenans, perhaps also indicating that this organism potentially adopts an even different strategy than *Z. galactanivorans* and *P. carragenovora* for the degradation of this abundant marine polysaccharide.

EXPERIMENTAL PROCEDURES

Bioinformatics analysis - Homologous sequences have been extracted with the blastp suite from NCBI server (NCBI Resource Coordinators 2017) (<https://blast.ncbi.nlm.nih.gov>) from the non-redundant protein sequences database, manually curated before being aligned with MAFFT program available at: <http://www.ebi.ac.uk/Tools/msa/mafft/> and edited with BioEdit free software (Hall *et al.* 2011) available at: <http://www.mbio.ncsu.edu/BioEdit/bioedit.html> in order to eliminate all the insertion areas before calculation of the phylogenetic tree with MEGA software (<http://www.megasoftware.net/>). Bootstrap values inferior to 50 % do not appear on the tree.

Figures of multiple alignment were done using ESPript 3.0 server (Robert and Gouet 2014) available at this address: <http://esprict.ibcp.fr/ESPript/ESPript/index.php> from the MAFFT alignment, edited with BioEdit when necessary to fit the structural data.

*Cloning of the two forms of κ -carrageenases of *Z. galactanivorans* in *E. coli** - Two different forms of the gene *cgkA* (*zobellia_236*) encoding the κ -carrageenase from *Z. galactanivorans* have been cloned. Briefly, primers (sequences provided in supplementary data, Table SII) were designed to amplify by PCR, from *Z. galactanivorans* genomic DNA, the coding regions corresponding to the complete sequence of the κ -carrageenase without the signal peptide sequence in N-terminal position (*ZgCgkA_{GH16}-CBM16- α PorSS*) and to the catalytic module alone (*ZgCgkA_{GH16}*). After digestion with the restriction enzymes BamHI and PstI, the purified PCR products were ligated using T4 DNA ligase into the expression vector pFO4 predigested by a compatible couple of restriction enzymes. The recombinant proteins encoded in the plasmids pZG237 and

pZG238 correspond respectively to the peptide sequences Q30 - K306 (ZgCgkA_{GH16}) and Q30 - E546 (ZgCgkA_{GH16-CBM16-PorSS}) with an added N-terminal hexahistidine tag. The plasmids were transformed into *E. coli* BL21 (DE3) strain for protein expression.

Mutagenesis of κ -carrageenase from P. carrageenovora - The plasmid pCGK that was obtained by cloning the native κ -carrageenase gene into the C-terminal His-tag encoding expression plasmid pET20b (Invitrogen) was subsequently used as a template for mutagenesis using the “QuickChange Mutagenesis Kit” (Stratagene). The primers used to do the amplification of the different mutants are summarized in the Table SII in supplementary data. The plasmids obtained were used to transform *E. coli* C43 (DE3) strain for protein expression.

Expression and purification of the recombinant enzymes - A single colony containing the desired plasmid (pCGK or pZG) was used to inoculate LB medium (supplemented with 100 μ g/ml ampicillin). This pre-culture was incubated at 37 °C and was then used to inoculate ZIP5052 auto inducing medium for cell growth at 17 °C. The culture was stopped after 2 days when A₆₀₀ had reached 15, and pelleted (30 min). If needed, the pellet was frozen and stored at -20 °C before use. The bacterial pellet was lysed by means of a French-press after resuspension in buffer A (Tris, HCl 10 mM, pH=7.2 for *P. carrageenovora* enzymes or 50 mM pH=7.5 for *Z. galactanivorans* enzymes, NaCl 500 mM, imidazole 40 mM), with an anti-protease mixture (Complete EDTA-free; Roche Applied Science), DNase (0.1 mg/ml) and addition of lysozyme (0.2 mg/ml). The supernatant after lysis containing the soluble proteins was separated from the pellet by centrifugation (20,000 \times g for 30 min at 4 °C) and then filtered prior to the two-step purification by nickel affinity chromatography (equilibrated with A buffer) and size exclusion chromatography on a Sephadex column (Amersham Biosciences).

Elution on IMAC was done with a gradient of buffer B that differs from buffer A only in the concentration of imidazole, which are 250 mM for *P. carrageenovora* enzymes and 500 mM for *Z. galactanivorans* enzymes. Size exclusion chromatography was done with a flow rate at 1 ml.min⁻¹ in buffer C (Tris, HCl 50 mM, pH7.2, NaCl 150 mM for *P. carrageenovora* enzymes and Tris, HCl 10 mM pH7.2, NaCl 250 mM for *Z. galactanivorans* enzymes). The different fractions were analyzed by SDS-PAGE. Those containing the protein purified to homogeneity were pooled and concentrated by ultrafiltration on a styrene acrylonitrile membrane (10-kDa or 30 kDa cutoff) (Millipore) to the desired concentration for biochemistry analysis, and to 6.9 mg/ml and 5 mg/ml for crystallogenesis of PcCgkA_{GH16-E168D} and of ZgCgkA_{GH16}, respectively.

Crystallization of PcCgkA_{GH16-E168D} and ZgCgkA_{GH16}, Structure Determination and Refinement - In a first step crystallization conditions were screened for using the commercial kits PACT and JCSG+ (Qiagen), dispensed by a robot (200 nl of protein solution mixed with 100 nl of reservoir solution). For both proteins initial crystallization conditions identified from these screens were then manually optimized. Single crystals of suitable size were obtained for PcCgkA_{GH16-E168D} in complex with substrate as follows: 200 μ l of enzyme at 6.9 mg.ml⁻¹ were supplemented with 2 mg of a purified mix of κ -carrabiose, tetraose and hexaose (weight ratio: 30:60:10), obtained from the digestion of κ -carrageenan by PcCgkA_{GH16}. Of this solution, 2 μ l were mixed with 2 μ l of reservoir solution containing 1.0 M sodium citrate and 100 mM cacodylate at pH6.5, in hanging drops equilibrated against 500 μ l of reservoir solution at 19°C. Single crystals of ZgCgkA_{GH16} were obtained from 2 μ l of the enzyme solution (5.0 mg.ml⁻¹) that were added to 1 μ l of reservoir solution containing 28-29 % of PEG 3350, 100 mM 2-(N-

morpholino)ethanesulfonic acid (MES) buffer at pH6.5 and 0.3 M NaNO₃, in hanging drops equilibrated against 250 μ l of reservoir solution at 19°C. Prior to flash-freezing in a nitrogen stream at 100K, single crystals were quickly soaked in their respective crystallization solution supplemented with 20 % glycerol. Diffraction data for *PcCgkA_{GH16-E168D}* complex crystals were collected on beamline ID14-1 and those for *ZgCgkA_{GH16}* on beamline ID29 (ESRF, Grenoble, France). X-ray diffraction data were integrated using Mosflm and scaled with SCALA (Winn *et al.* 2011). Both structures of *ZgCgkA_{GH16}* and *PcCgkA_{GH16-E168D}* in complex with κ -carratetraose were determined by molecular replacement with MolRep (Vagin and Teplyakov 1997) using chain A of the *PcCgkA_{GH16}* native structure (PDB code 1DYP) as starting model. The structure of *ZgCgkA_{GH16}* was then manually corrected and built using COOT (Emsley and Cowtan 2004). For both structures, the initial structural models were refined with the program REFMAC5 (Vagin *et al.* 2004), alternating with cycles of further manual rebuilding using COOT. A subset of 5 % randomly selected reflections was excluded from computational refinement to calculate the R_{free} factors throughout the refinement. The addition of the ligand sugar units for the complex structure was performed manually using COOT. Water molecules were added automatically with REFMAC-ARP/wARP and visually verified. All data collection and refinement statistics are presented in Table 1. The structure obtained for the ligand tetrasaccharide has also been checked with the software Privateer in CCP4 (Agirre *et al.* 2015).

Enzymatic activity assays on κ -carrageenan by reducing sugar analysis - The processive character of an enzyme can only be clearly observed on a non-soluble substrate (Lemoine *et al.* 2009). However, the determination of kinetic parameters implies a total availability of the substrate to fulfill

Mickaelis-Menten conditions (Johnson and Goody 2011). Thus, we have decided to do the biochemical experiments in diluted solutions of κ -carrageenan to have an estimate of the apparent efficiency under similar conditions for all studied enzymes and mutants, although not entirely reflecting their behavior in 'natural' conditions (i.e. on solid algal cell walls).

A stock solution of 0.5 % (w/v) κ -carrageenan (Danisco) was prepared in 50 mM MOPS, pH 7, 150 mM NaCl for *PcCgkA_{GH16}* and its mutants; in Teorell buffer (Östling and Virtama 1946b) to determine the best range of pH for both construction of *ZgCgkA_{GH16}* and then in 50 mM MES pH 6, 300 mM NaCl for the determination of kinetic constants. Aliquots (14 μ l) of enzyme at the appropriate concentration were incubated in triplicate in the presence of κ -carrageenan solution (126 μ l) at a final concentration of 0.125 % (w/v) at 40 °C. The amount of reducing sugars released was assayed using a ferricyanide method adapted from that of Kidby and Davidson (Kidby and Davidson 1973). The ferricyanide reagent consisted of 300 mg of potassium hexacyanoferrate III and of 28 g of hydrated Na₂CO₃ dissolved in 1 l of distilled water, to which 1 ml of 1 M aqueous NaOH was added. Aliquots (20 μ l) of the incubation medium were mixed with 180 μ l of ferricyanide reagent, and absorbance was recorded at 420 nm. Reducing sugar concentrations were calibrated using glucose as standard.

The initial reaction rates of κ -carrageenan hydrolysis were measured at least in triplicate by the ferricyanide reagent method for ten κ -carrageenan concentrations (1.25, 1.05, 0.85, 0.65, 0.45, 0.30, 0.25, 0.20, 0.15, 0.13 g/l). The Michaelis constant (K_m) and the reaction rate at infinite substrate concentration (V_m) were calculated with the software Hyper, with the hyperbolic regression model applying weighting. Note: for *ZgCgkA_{GH16}*, parameters have been calculated without data

at 0.13 g/l and for *PcCgkA_{GH16-R196A}* and *PcCgkA_{GH16-W266A}* without data at 1.25 and 1.05 g/l. For *PcCgkA_{GH16-R260A}*, the concentration of enzyme needed to get comparable kinetic curves at 1.25 g/l of substrate was 100 time higher than for *PcCgkA_{GH16}*, thus preventing us to calculate kinetic parameters, owing to the fact that we were no more in conditions respecting the Micaelis-Menten assumption of $[S] \gg [E]$. For long time course kinetics, digestions were conducted at 30 °C, in a final volume of 250 μ l of 0.2 % (w/v) κ -carrageenan solutions, in 10 mM Tris buffer pH 7.2, 150 mM NaCl. 3.5 nM of each enzyme were dissolved in the κ -carrageenan solution, or 14 nM in microgels, obtained from the previous solution at 0.2 % κ -carrageenan supplemented with 60 mM KCl. Aliquots (20 μ l) of the incubation medium were pipetted after 15, 30, 45, 60, 90, 180 and 300 min, and mixed with 180 μ l of 2.5X ferricyanide reagent and treated the same way as above. Experiments were done

in triplicate and the standard deviation of the mean has been drawn for each spot on the graphics.

HPAEC analysis - Analyses of degradation products generated by the different enzymes were conducted by HPAEC on an ICS-5000 system (*Dionex*). 20 μ l of samples were injected (AS50 Automatic Autosampler, *Dionex*) onto an AS11 analytical anion-exchange column (4 \times 250 mm, IonPac®, *Dionex*) equipped with an AG11 pre-column (4 \times 50 mm, IonPac®, *Dionex*). Elution was performed at a flow rate of 0.5 ml/min (Gradient Pump SP-5, *Dionex*) by NaOH multi-step gradients from 8 to 280 mM (40 min). Oligosaccharides were detected in conductivity mode (Analytical DC-5000 detector, *Dionex*) with a suppressor (AERS 500, 4 mm, *Dionex*) delivering a current from 50 to 300 mA. Acquisition of chromatograms was realized by the "Chromeleon" Software version 6.8 (*Dionex*).

Acknowledgments: We are indebted to local contacts for their support during data collection at beamlines ID14-1 and ID29, ESRF (European synchrotron, Grenoble, France).

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article

Author contributions: MMM, MC and PNC designed the study and wrote the paper. RL, ThB and MMM did the molecular biology experiments. AJ, MJ, ThB, and MC performed crystallizations and ThB, CL MC and MMM did the crystal structure determinations. MMM did the biochemical characterization of the recombinant proteins. MJ and AP provided technical assistance for chromatography and HPAEC analysis of the samples. GM and TrB did the phylogenetic analysis. All authors analyzed the results and approved the final version of the manuscript.

REFERENCES

1. O'Dowd, C., Jimenez, J. L., Bahreini, R., Flagan, R. C., Seinfeld, J. H., Hameri, K., Pirjola, L., Kulmala, M., Jennings, S. G., and Hoffmann, T. (2002) Marine aerosol formation from biogenic iodine emissions. *Nature*. **417**, 1–5
2. Duggins, D. O., Simenstad, C. A., and Estes, J. A. (1989) Magnification of Secondary Production by Kelp Detritus in Coastal Marine Ecosystems. *Science*. **245**, 170–173
3. Martin, M., Portetelle, D., Michel, G., and Vandenberg, M. (2014) Microorganisms living on macroalgae: Diversity, interactions, and biotechnological applications. *Appl. Microbiol. Biotechnol.* **98**, 2917–2935

4. Larsbrink, J., Rogers, T. E., Hemsworth, G. R., McKee, L. S., Tauzin, A. S., Spadiut, O., Klintner, S., Pudlo, N. A., Urs, K., Koropatkin, N. M., Creagh, A. L., Haynes, C. A., Kelly, A. G., Cederholm, S. N., Davies, G. J., Martens, E. C., and Brumer, H. (2014) A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature*. **506**, 498–502
5. Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W., and Brumer, H. (2017) Polysaccharide Utilization Loci: Fuelling microbial communities. *J. Bacteriol.* **199**. pii: e00860-16.
6. Bjursell, M. K., Martens, E. C., and Gordon, J. I. (2006) Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J. Biol. Chem.* **281**, 36269–36279
7. 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
8. Michel, G., and Czjzek, M. (2013) Polysaccharide-degrading enzymes from marine bacteria. in *Marine enzymes for biocatalysis: sources, biocatalytic characteristic and bioprocesses of marine enzymes*, pp. 429–464
9. Senni, K., Pereira, J., Gueniche, F., Delbarre-Ladrat, C., Siquin, C., Ratiskol, J., Godeau, G., Fischer, A. M., Helley, D., and Collic-Jouault, S. (2011) Marine polysaccharides: A source of bioactive molecules for cell therapy and tissue engineering. *Mar. Drugs*. **9**, 1664–1681
10. Knutsen, S. H., Myslabodski, D. E., Larsen, B., and Usov, A. I. (1994) A Modified System of Nomenclature for Red Algal Galactans. *Bot. Mar.* **37**, 163–170
11. Barbeyron, T., Gerard, A., Potin, P., Henrissat, B., and Kloareg, B. (1998) The Kappa-carrageenase of the marine bacterium *Cytophaga drobachiensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.* **15**, 528–537
12. Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) The κ -carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure*. **9**, 513–525
13. Michel, G., Helbert, W., Kahn, R., Dideberg, O., and Kloareg, B. (2003) The Structural Bases of the Processive Degradation of ι -Carrageenan, a Main Cell Wall Polysaccharide of Red Algae. *J. Mol. Biol.* **334**, 421–433
14. Rebuffet, E., Barbeyron, T., Jeudy, A., Jam, M., Czjzek, M., and Michel, G. (2010) Identification of catalytic residues and mechanistic analysis of family GH82 ι -carrageenases. *Biochemistry*. **49**, 7590–7599
15. Potin, P., Sanseau, A., Le Gall, Y., Rochas, C., and Kloareg, B. (1991) Purification and characterization of a new κ -carrageenase from a marine *Cytophaga*-like bacterium. *Eur. J. Biochem.* **201**, 241–247
16. Guibet, M., Colin, S., Barbeyron, T., Genicot, S., Kloareg, B., Michel, G., and Helbert, W. (2007) Degradation of λ -carrageenan by *Pseudoalteromonas carrageenovora* λ -carrageenase: a new family of glycoside hydrolases unrelated to κ - and ι -carrageenases. *Biochem. J.* **404**, 105–114
17. Barbeyron, T., Henrissat, B., and Kloareg, B. (1994) The gene encoding the kappa-carrageenase of *Alteromonas carrageenovora* is related to β -1,3-1,4-glucanases. *Gene*. **139**, 105–109

18. Liu, G. L., Li, Y., Chi, Z., and Chi, Z. M. (2011) Purification and characterization of kappa-carrageenase from the marine bacterium *Pseudoalteromonas porphyrae* for hydrolysis of kappa-carrageenan. *Process Biochem.* **46**, 265–271
19. Kobayashi, T., Uchimura, K., Koide, O., Deguchi, S., and Horikoshi, K. (2012) Genetic and Biochemical Characterization of the *Pseudoalteromonas tetraodonis* Alkaline κ -Carrageenase. *Biosci. Biotechnol. Biochem.* **76**, 506–511
20. Liu, Z., Li, G., Mo, Z., and Mou, H. (2013) Molecular cloning, characterization, and heterologous expression of a new κ -carrageenase gene from marine bacterium *Zobellia* sp. ZM-2. *Appl. Microbiol. Biotechnol.* **97**, 10057–10067
21. Wang, L., Li, S., Zhang, S., Li, J., Yu, W., and Gong, Q. (2015) A new κ -carrageenase CgkS from marine bacterium *Shewanella* sp. Kz7. *J. Ocean Univ. China.* **14**, 759–763
22. Xu, X., Li, S., Yang, X., Yu, W., and Han, F. (2015) Cloning and characterization of a new κ -carrageenase gene from marine bacterium *Pseudoalteromonas* sp. QY203. *J. Ocean Univ. China.* **14**, 1082–1086
23. Yaphe, W., and Baxter, B. (1955) The enzymic hydrolysis of carrageenin. *Appl. Microbiol.* **3**, 380–383
24. Weigl, J., and Yaphe, W. (1966) The enzymatic hydrolysis of carrageenan by *Pseudomonas carrageenovora*: purification of a κ -carrageenase. *Can. J. Microbiol.* **12**, 939–947
25. Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., Goesmann, A., Leblanc, C., Oliver Glöckner, F., Czjzek, M., Amann, R., and Michel, G. (2016) Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* DsijT. *Environ. Microbiol.* **18**, 4610–4627
26. Koshland, D. E. (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev. Camb. Philos. Soc.* **28**, 416–436
27. Zechel, D. L., Withers, S. G., And, D. L. Z., and Withers, S. G. (2000) Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **33**, 11–18
28. Planas, A. (2000) Bacterial 1, 3-1, 4- L -glucanases : structure , function and protein engineering
29. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Glycosyl Hydrolases. *Biochem. J. Lett.* **321**, 557–559
30. Martin, M., Barbeyron, T., Martin, R., Portetelle, D., Michel, G., and Vandenbol, M. (2015) The cultivable surface microbiota of the brown alga *Ascophyllum nodosum* is enriched in macroalgal-polysaccharide-degrading bacteria. *Front. Microbiol.* **6**, 1–14
31. Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H., and Henrissat, B. (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol. Biol.* **12**, 186
32. Stam, M. R., Danchin, E. G. J., Rancurel, C., Coutinho, P. M., and Henrissat, B. (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: Towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng. Des. Sel.* **19**, 555–562
33. Knutser, S. H., and Grasdalen, H. (1987) Characterization of Water-extractable Polysaccharides from Norwegian *Furcellaria lumbricalis* (Huds.) Lamour. (Gigartinales, Rhodophyceae) by IR and NMR Spectroscopy. *Bot. Mar.* **30**, 497–506
34. Falshaw, R., Furneaux, R. H., Wong, H., Liao, M., Bacic, A., and Chandkrachang, S. (1996) Structural analysis of carrageenans from Burmese and Thai samples of *Catenella*

- nipae* Zanardini. *Carbohydr. Res.* **285**, 81–98
35. Usov, A. I. (2011) *Polysaccharides of the red algae*, 1st Ed., Elsevier Inc.
 36. Potin, P., Richard, C., Barbeyron, T., Henrissat, B., Gey, C., Petillot, Y., Forest, E., Dideberg, O., Rochas, C., and Kloareg, B. (1995) Processing and Hydrolytic Mechanism of the *cgkA*-Encoded κ -Carrageenase of *Alteromonas carrageenovora*. *Eur. J. Biochem.* **228**, 971–975
 37. Tomme, P., van Tilbeurgh, H., Pettersson, G., van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeysens, M. (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. *Eur. J. Biochem.* **170**, 575–581
 38. Din, N., Gilkes, N. R., Tekant, B., Miller, R. C., Warren, R. A. J., and Kilburn, D. G. (1991) Non hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. *Nat. Biotechnol.* **9**, 1096–1099
 39. Black, G. W., Rixon, J. E., Clarke, J. H., Hazlewood, G. P., Theodorou, M. K., Morris, P., and Gilbert, H. J. (1996) Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. *Biochem. J.* **319**, 515–520
 40. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781
 41. Gilbert, H. J., Knox, J. P., and Boraston, A. B. (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr. Opin. Struct. Biol.* **23**, 669–677
 42. Lemoine, M., Nyvall Collén, P., Helbert, W., Nyvall Collén, P., and Helbert, W. (2009) Physical state of κ -carrageenan modulates the mode of action of κ -carrageenase from *Pseudoalteromonas carrageenovora*. *Biochem. J.* **419**, 545–553
 43. Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sorlie, M., Synstad, B., Vriend, G., Varum, K. M., Eijsink, V. G. H., Sørli, M., Synstad, B., Vriend, G., Vårum, K. M., and Eijsink, V. G. H. (2006) Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc. Natl. Acad. Sci.* **103**, 18089–18094
 44. Gilad, R., Rabinovich, L., Yaron, S., Bayer, E. A., Lamed, R., Gilbert, H. J., and Shoham, Y. (2003) Cell, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *J. Bacteriol.* **185**, 391–398
 45. Riordan, J. F., and Hayashida, H. (1970) Chemical evidence for a functional carboxyl group in carboxypeptidase A. *Biochem. Biophys. Res. Commun.* **41**, 122–127
 46. Riordan, J. F., McElvany, K. D., and Borders, C. L. (1977) Arginyl residues: anion recognition sites in enzymes. *Science.* **195**, 884–886
 47. Huang, W., Boju, L., Tkalec, L., Su, H., Yang, H. O., Gunay, N. S., Linhardt, R. J., Yeong, S. K., Matte, A., and Cygler, M. (2001) Active site of chondroitin AC lyase revealed by the structure of enzyme - Oligosaccharide complexes and mutagenesis. *Biochemistry.* **40**, 2359–2372
 48. Von Ossowski, I., Ståhlberg, J., Koivula, A., Piens, K., Becker, D., Boer, H., Harle, R., Harris, M., Divne, C., Mahdi, S., Zhao, Y., Driguez, H., Claeysens, M., Sinnott, M. L., and Teeri, T. T. (2003) Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D. *J. Mol. Biol.* **333**, 817–829
 49. Watanabe, T., Ariga, Y., Sato, U., Toratani, T., Hashimoto, M., Nikaidou, N., Kezuka, Y., Nonaka, T., and Sugiyama, J. (2003) Aromatic residues within the substrate-binding cleft of *Bacillus circulans* chitinase A1 are essential for hydrolysis of crystalline chitin.

- Biochem. J.* **376**, 237–244
50. Divne, C., Ståhlberg, J., Teeri, T. T., and Jones, T. (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *J. Mol. Biol.* **275**, 309–325
 51. NCBI Resource Coordinators (2017) Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **45**, D12–D17
 52. Hall, T., Biosciences, I., and Carlsbad, C. (2011) BioEdit: An important software for molecular biology. *GERF Bull. Biosci.* **2**, 60–61
 53. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, 320–324
 54. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242
 55. Vagin, A., and Teplyakov, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst.* **30**, 1022–1025
 56. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132
 57. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2184–2195
 58. Agirre, J., Iglesias-Fernández, J., Rovira, C., Davies, G. J., Wilson, K. S., and Cowtan, K. D. (2015) Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* **22**, 833–834
 59. Johnson, K. A., and Goody, R. S. (2011) The original Michaelis constant: Translation of the 1913 Michaelis-Menten Paper. *Biochemistry.* **50**, 8264–8269
 60. Östling, S., and Virtama, P. (1946) A Modified Preparation of the Universal Buffer Described by Teorell and Stenhagen. *Acta Physiol. Scand.* **11**, 289–293
 61. Kidby, D., and Davidson, D. (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.* **55**, 321–325

FOOTNOTES

The authors gratefully acknowledge the financial help of ANRT and Olmix group for their funding of a PhD grant (CIFRE-684/2014) for M. Matard-Mann. This work was also financially supported by the French National Research Agency with regards to the investment expenditure program IDEALG (<http://www.idealg.ueb.eu/>, grant agreement No. ANR-10-BTBR-04) and supported by the “Region Bretagne”.

The abbreviations used are: GH, glycoside hydrolase; PUL, polysaccharide utilizing locus; CBM, carbohydrate binding module; PDB, protein databank; DP, degree of polymerization; DA, 3,6-anhydro-D-galactose; G4S, D-galactose-4-sulfate.

TABLES

TABLE I. Kinetic parameters of the different recombinant and point-mutant enzymes in presence of liquid substrate. Calculations have been done from at least 3 replicates and treated with Hyper software.

	kcat (s ⁻¹)	Km (M)	kcat/Km (M ⁻¹ . s ⁻¹)	% vs WT
<i>PcCgkAGH16</i> (WT)	135.84	6.83x10⁻⁷	1.99x10⁸	100
<i>ZgCgkAGH16-CBM16-PorSS</i>	660.91	1.04x10⁻⁶	6.35x10 ⁸	319
<i>ZgCgkAGH16</i>	914.65	8.22x10⁻⁷	1.11x10⁹	560
<i>PcCgkAGH16-R92A</i>	183.20	8.99x10⁻⁷	2.04x10 ⁸	102
<i>PcCgkAGH16-R151A</i>	478.79	1.21x10⁻⁶	3.97x10 ⁸	199
<i>PcCgkAGH16-Q171A</i>	120.17	3.85x10⁻⁷	3.12x10 ⁸	157
<i>PcCgkAGH16-R196A</i>	6.13	1.15x10⁻⁶	5.31x10 ⁶	3
<i>PcCgkAGH16-R260A</i>	nd	nd	nd	<1
<i>PcCgkAGH16-W266A</i>	27.36	2.63x10⁻⁶	1.04x10 ⁷	5

TABLE II. Area ratio of chromatographic peaks corresponding to DP4 and DP6 in Figure 4. $\Delta M/S$ corresponds to the ratio of DP4/DP6 obtained in microgels versus solutions.

	<i>ZgCgkAGH16</i>			<i>ZgCgkAGH16-CBM16-PorSS</i>		
	DP4/DP6		$\Delta M/S$	DP4/DP6		$\Delta M/S$
	solution	microgel		solution	microgel	
15 min	0.68	1.53	2.24	0.90	1.50	1.66
45 min	0.85	1.79	2.12	1.04	1.90	1.83
90 min	1.19	2.26	1.89	1.13	2.33	2.06
24 h	1.37	2.68	1.95	1.49	4.17	2.80
1 week	1.54	2.90	1.88	2.25	11.12	4.94

TABLE III. Delimitation of structural “fingers” of the two catalytic modules of kappa-carrageenases from *Z. galactanivorans* (5OCR) and *P. carrageenovora*, mutant E168D in complex with κ -neocarratetraose (5OCQ) and also the apo-enzyme (1DYP), and values of B factor associated

		B factor		B factor		B factor
	ZgCgkAGH16	26.0	PcCgkAGH16-E168D	12.6	PcCgkAGH16	10.3
Finger F1	W62-W76	30.8	W57-W69	6.2	W57-W69	8.3
Finger F2	-	-	R92-Y106	11.3	R92-Y106	12.4
Finger F3	D145-V156	24.9	T148-Q160	22.4	T148-Q160	12.8
Finger F4	F168-D177	32.3	-	-	-	-
Finger F5	K190-K201	28.5	V187-G198	25.3	V187-G198	14.6
Finger F6	K264-A276	49.6	A261-P273	11.2	A261-P273	9.1

TABLE IV. Data collection and refinement statistics for *PcCgkAGH16-E168D* mutant and *ZgCgkAGH16* crystals

Data statistics	<i>PcCgkAGH16-E168D</i>	<i>ZgCgkAGH16</i>
ESRF Beam line	ID14-1	ID29
Wavelength (Å)	0.99	0.97
Space group	P2 ₁ 2 ₁ 2 ₁	P1
Unit cell parameters (Å)	a=62.63 b=67.39 c=159.44 a=b=g=90,00°	a=40.96 b=83.04 c=85.71 $\alpha=72.60^\circ$ $\beta=88.30^\circ$ $\gamma=89.50^\circ$
V _m	2.69	2.22
% solvent	54.4	44.7
Resolution range ^a	79.0 – 1.710 (1.76 – 1.70)	81.7 – 1.61 (1.70 – 1.61)
No. of observations	268064	261640
No. of unique reflections	74710	132172
Rmerge (%)	7.1 (15.5)	0.052 (0.54)
Mean I/s(I)	14.5 (6.2)	11.97 (1.7)
Redundancy	3.6 (3.4)	1.98
Completeness (%)	99.6 (99.8)	93.9 (91.0)
Molecular replacement	Native κ -carrageenase	
Rfactor (%)	37.0	58.2
Correlation	60.0	38.8
molecules/asymmetric unit	2	4
Refinement		
Resolution range	79.06 - 1.70	81.74 - 1.66
Resolution	1.70	1.66
No. of unique reflections	70358	114681
Rfactor (%)	14.6	16.9
Rfree (%)	17.7	20.5
No. of protein atoms per molecule A; B; C; D (mean B factor in Å ²)	2252 ; 2311 (12.6) (9.8)	2411; 2399; 2395; 2437 (26.0) (29.5) (30.6) (27.3)
No. of solvent atoms (mean B factor: Solvent in Å ²)	585 (24.40)	793 (35.70)
Mean B factor: Ligand in Å ²	16.70; 11.30	-
RMSD bond length (Å)	0.012	0.019
RMSD torsion angle (°)	1.3	2.0
Mean overall B factor (Å ²)	12.84	28.90
Ramachan plot, most favoured (%)	97.06	95.5
Ramachan plot, additional allowed (%)	2.76	4.21
Ramachan plot, outliers (%)	0.18	0.29
PDB code	5OCQ	5OCR

^aValues in parenthesis are for the high resolution shell

TABLE V. Close contacts ($<3.6\text{\AA}$) between κ -neocarratetraose and amino acid residues or water molecules in the crystal structure of *PcCgkA*_{GH16-E168D}

Oligosaccharide		Interacting atoms	Distance (\AA)	
Residues	Atoms			
DA -4	O6	Gln102 ^{NE2}	3.1	
	O5/O2	Arg151 ^{NH1}	2.6/3.2	
	Hydrophobic	Trp95	-	
G4S -3	O4S1O	Trp95 ^{NE2} /H ₂ O344	2.9/2.8	
		Gln100 ^{NE2} /H ₂ O344	3.5/2.8	
		Gln270 ^{OE1} /H ₂ O116/H ₂ O344	2.7/2.71/2.8	
	O4S3O	Cys268 ^{OC1} /H ₂ O116/H ₂ O344	2.9/2.71/2.8	
		Trp194 ^{NE1}	3.0	
		Arg151 ^{NH1/NH2}	2.9/2.9	
	O3	Arg151 ^{NH1}	3.0	
	O4S2O	Asn269 ^{OD1}	3.2	
	O2	Glu62 ^{OE1} /H ₂ O343/H ₂ O449	2.6/2.6/2.8	
		Asn63 ^{OC1} /H ₂ O478	3.5/3.1	
		Arg196 ^{NH1} /H ₂ O322	2.9/3.0	
	O6	Asn269 ^{OC1}	2.6	
		DA -2	Glu62 ^{OE1} /H ₂ Ot343	2.6/2.8
			Hydrophobic clamp	Tyr64 ^{OH} /H ₂ O14/Tyr146 ^{OH}
	O5		Arg196 ^{NH1} /H ₂ O322	2.9/3.0
O1	Tyr161 ^{OH}	3.5		
G4S -1	O2	Arg196 ^{NH2}	3.0	
		Glu163 ^{OE1}	2.9	
	O4S3O	Arg196 ^{NH2} /H ₂ O322/H ₂ O95	2.9/3.0/2.9	
	O4S2O	Arg260 ^{NH1}	2.9	
	O4S1O	Arg260 ^{NH2} /H ₂ O77/H ₂ O109	3.0/2.8/2.8	
	O4	Arg260 ^{NH2} /H ₂ O331	2.5/2.4	
	O5/O6	Arg260 ^{NH2} /H ₂ O331	2.5/2.5-2.5	
	O6	Arg260 ^{NH2} /H ₂ O77	3.0/2.9	
		Asp168 ^{OD1}	2.6	
	O5	Asp168 ^{OD1} /H ₂ O373	2.3/2.5	
	O1	Asp168 ^{OD1} /H ₂ O373	2.3/2.6	
		His183 ^{NE2}	3.1	
	Hydrophobic	Trp144	-	

FIGURES LEGENDS

FIGURE 1. Schematic representations of the chemical structure of kappa-carrageenan tetraose (κ -neocarratetraose). At the top, the chemical representations of the units show the standard chair conformation. At the bottom, the simplified scheme follows the same order of monosaccharides and glycosidic linkages as above.

DA: 3,6-anhydro-D-galactose, G4S: D-galactose-4-sulfate, NR: non-reducing end, R: reducing end

FIGURE 2. Comparison of the catalytic modules of *PcCgkAGH16* and *ZgCgkAGH16*

A. Multiple sequence alignment of the catalytic modules of both enzymes. The alignment was produced with MultAlin, and the figure using ESPript 3.x software; Secondary structure elements are symbolized with arrows for beta-sheets, curls for alpha-helix and T for turns; the “finger” regions surrounding the catalytic channel are boxed and numbered from F1 to F6; the sequence divergences that have structural consequences are boxed in green; the numbers ‘1’ in green refer to the cysteine residues involved in the single disulfide bond of *PcCgkAGH16*; the black stars above the sequence highlight the amino acids that were mutated into alanine for *PcCgkAGH16* in this study. **B.** Cartoon representation of the structural fold of *ZgCgkAGH16*. The color code is graduated from blue (N-terminus) to red (C-terminus) along the polypeptide chain. Catalytic residues E159-D161-E164 are shown as sticks and colored in red. **C.** Cartoon representation of the structural fold of *PcCgkAGH16-E168D* in complex with a κ -neocarratetraose, shown in stick representation (carbon-atoms are white, oxygen red and sulfur-atoms orange). The color code of the polypeptide chain is the same as in **B**. Catalytic residues E163-D165-E168D are shown as sticks and colored in red.

FIGURE 3. Phylogenetic tree of the currently sequenced κ -carrageenases

The multiple alignments were done with 37 sequences of catalytic modules of κ -carrageenases. Clade A to D are defined according to the bootstrap values. Structural common features are identified for these clades and are represented in the circled medals; in purple: presence of finger F2 and absence of finger F4; in blue: presence of finger F4 and absence of finger F2; in black: presence of both fingers F2 and F4, subgroup I being the closer in sequence similarity for F2 and F4 motives and subgroup II showing the higher divergence. Blue arrows identify the six κ -carrageenases from *Algibacter sp.* SK-16.

FIGURE 4. Kinetics of degradation of κ -carrageenan

A, B. Reducing sugar analysis in solution (A) and microgels (B) of kappa-carrageenan for *PcCgkAGH16* (white circles), *ZgCgkAGH16* (light grey squares) and *ZgCgkAGH16-CBM16-PorSS* (dark grey triangles). Standard deviation of triplicates is represented as a bar.

C, D. HPAEC analyses of degradation products of *ZgCgkAGH16* in solution (C) and in microgel (D).

E, F. HPAEC analyses of degradation products of *ZgCgkAGH16-CBM16-PorSS* in solution (E) and in microgel (F). For **C, D, E, F**, different curves correspond to kinetics at $t_0=0$, $t_1=15$ min, $t_2=45$ min, $t_3=24$ h, $t_4=1$ week.

FIGURE 5. Structural representations of the substrate/enzyme interactions in *PcCgkAGH16-E168D*

A. Fourier difference map (Fo-Fc) displaying the electron density (blue mesh) in the active site at a level of 3.5 σ that indicates the presence of the DP4 oligo-saccharide (stick representation with

the same color code as in Figure 2). The catalytic sub-binding sites are numbered from -4 to -1. The polypeptide chain is represented as purple cartoon. **B.** Schematic representation of the enzyme/substrate complex (produced with ChemDraw). All residues that interact with the substrate molecule are displayed. Possible hydrogen bonds are represented in dashes and those established with sulfate groups are underlined in yellow, as well as the labels of the corresponding residues; water molecules are drawn as spheres; hydrophobic interactions are symbolized with “arcs”; the mutation E168D is preceded by an asterisk; the three arginine residues that were mutated into alanines are identified by a star; the three catalytic residues are labeled in red, and red waves symbolize the hydrogen bonds that are subsequently formed during the catalytic cleavage of the glycosidic linkage. **C.** Cartoon representation of *PcCgkA_{GH16-E168D}*, showing the positions of the amino acids mutated into alanine, with respect to the catalytic sub-binding sites. The positions of residues W266 and Q171 are on either side of the subsite +1. **D.** Superposition of *PcCgkA_{GH16}* (1DYP) in grey and *PcCgkA_{GH16-E168D}* in purple, in complex with a neocarratetraose in light grey (DA-G4S-DA-G4S) (5OCQ).

FIGURE 6. Structural comparison of the catalytic modules of *PcCgkA_{GH16-E168D}* and *ZgCgkA_{GH16}* and their interactions with κ -neocarratetraose

A. Schematic representation of the overall structure with the six identified fingers numbered from F1 to F6; the catalytic triad is identified with red letters; the substrate is symbolized by the orange arrow oriented from its non-reducing (NR) to its reducing (R) end; the specific fingers F2 and F4 are identified with a purple and a blue ring, respectively. **B.** Ribbon representation of the two superposed structures with the oligosaccharide in orange, the catalytic triad in red, and the specific structural features of both enzymes in purple for *PcCgkA_{GH16-E168D}*, and blue for *ZgCgkA_{GH16}*. **C.** *ZgCgkA_{GH16}* surface representation with its specific structural features in blue (top) superimposed with the oligosaccharide trapped in *PcCgkA_{GH16-E168D}* in orange, and *PcCgkA_{GH16-E168D}* surface representation with its specific structural features in purple (bottom) and the oligosaccharide in orange. The tilt of the positive substrate binding sites by 45° is highlighted in orange. (+) and (-) indicate the regions of the respective subsites **D.** Schematic representation of the interactions between the catalytic subsites of *ZgCgkA_{GH16}* (blue) and *PcCgkA_{GH16-E168D}* (purple) with a DP8 κ -carrageenan; the red dashed arrow symbolizes the cleavage site; white arrows symbolize the proposed direction of enzymatic processivity.

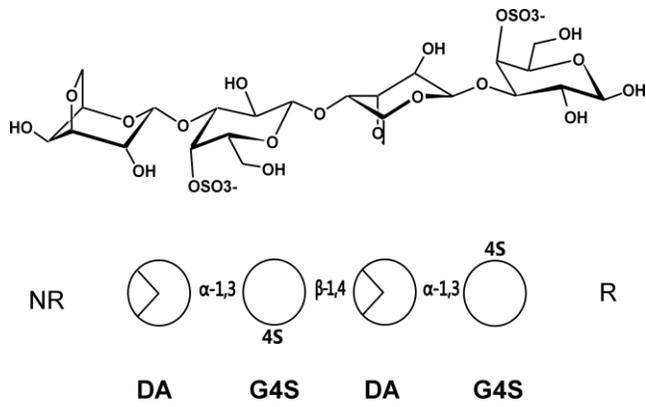


Figure 1

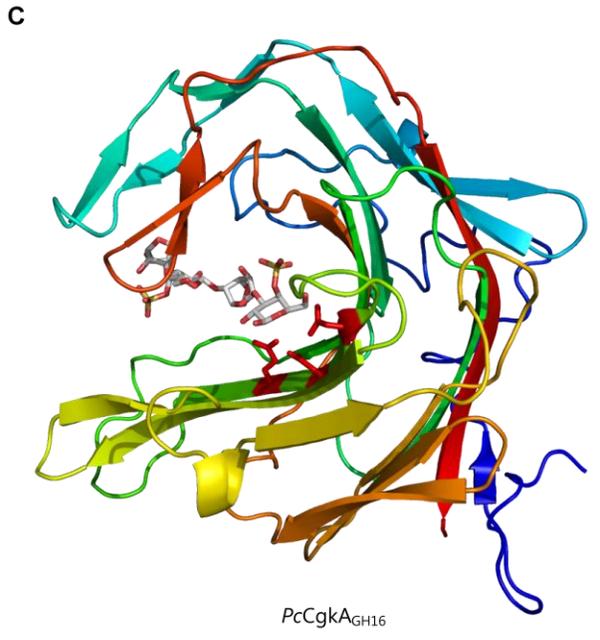
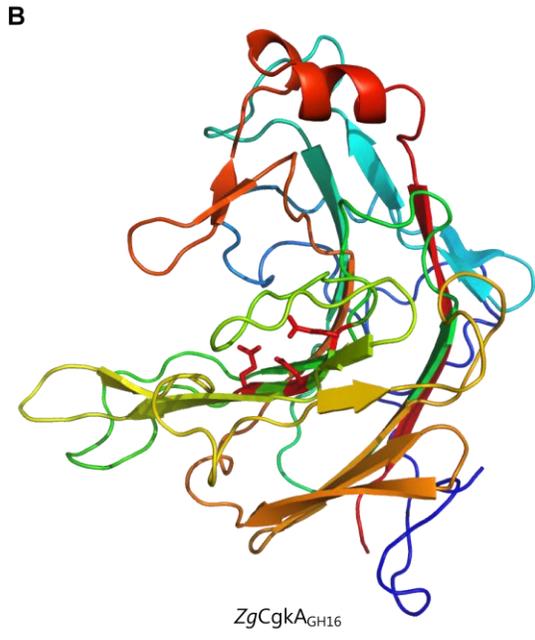
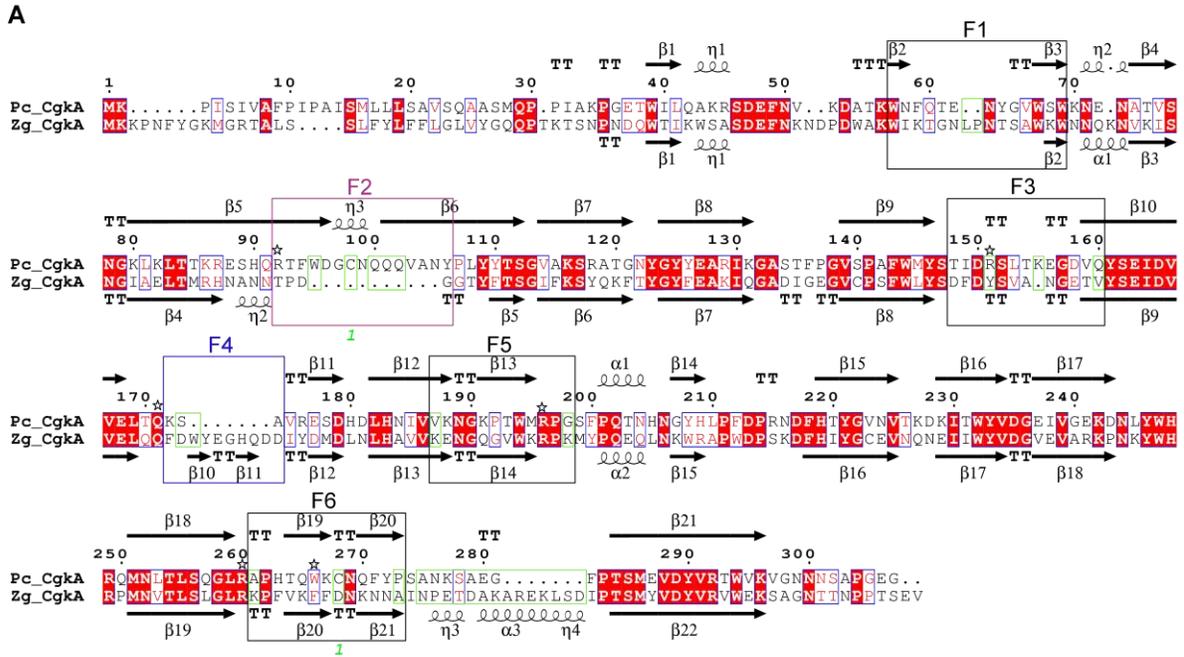


Figure 2

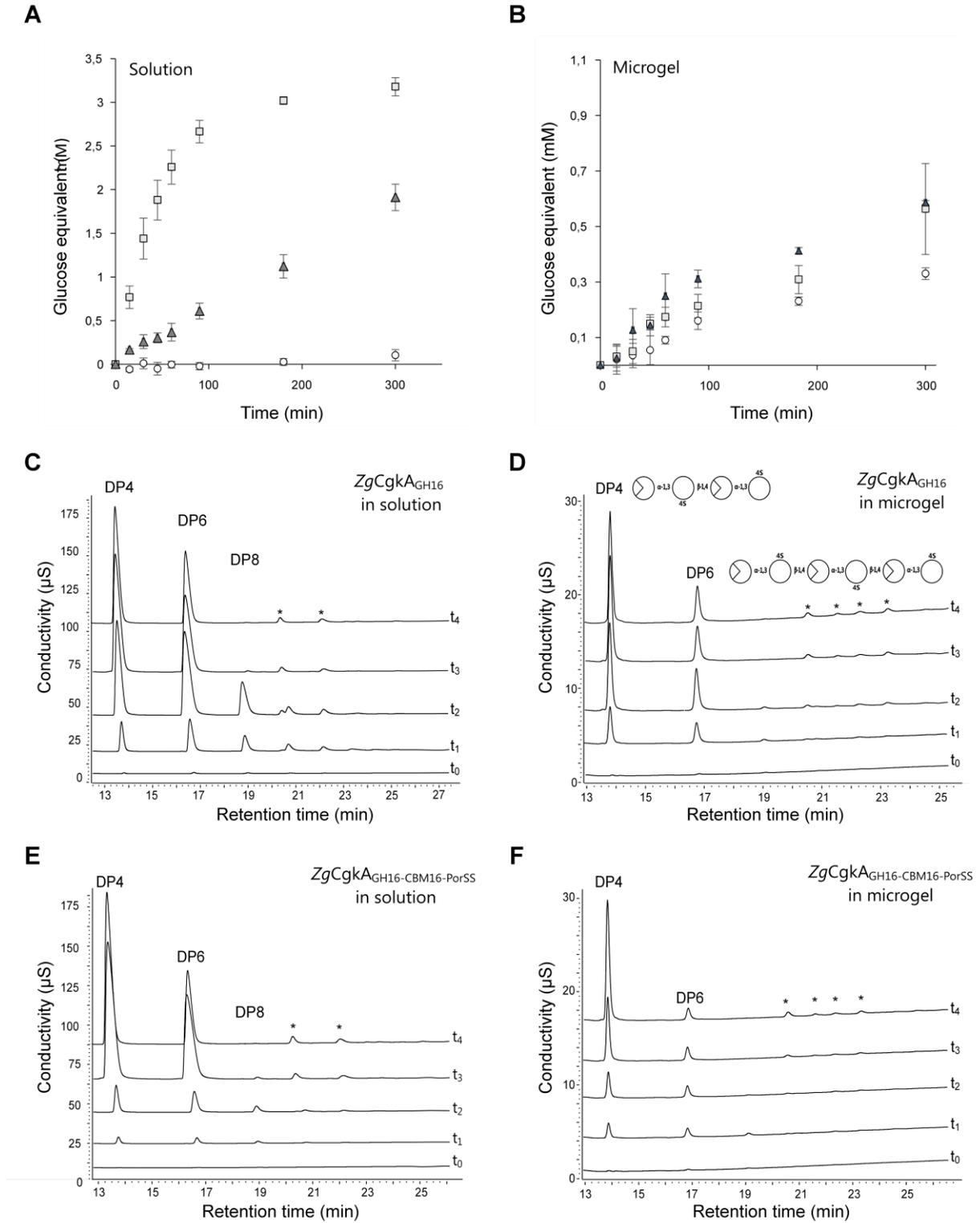
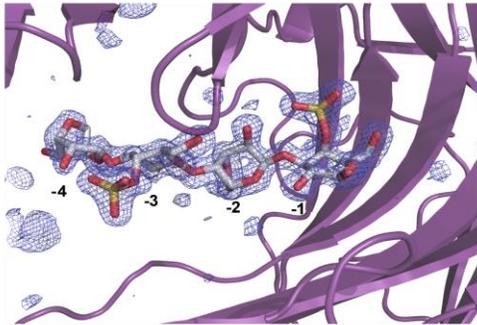
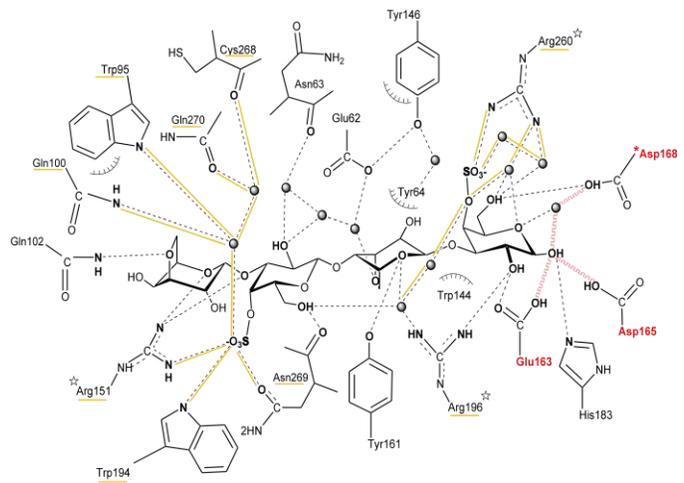


Figure 4

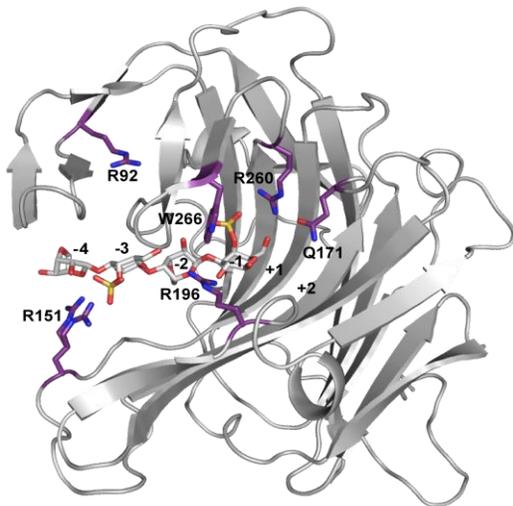
A



B



C



D

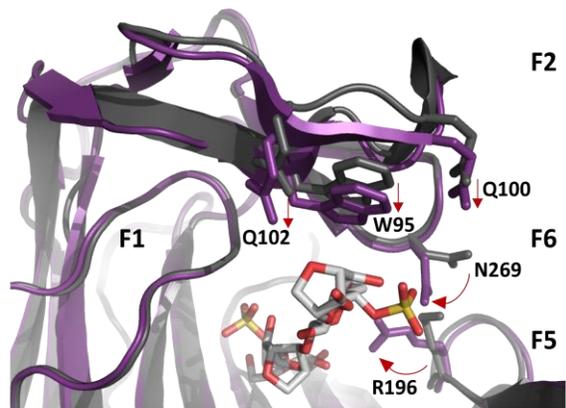
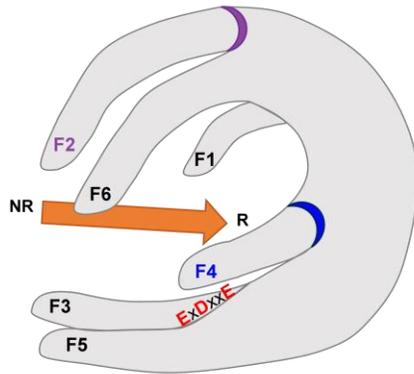
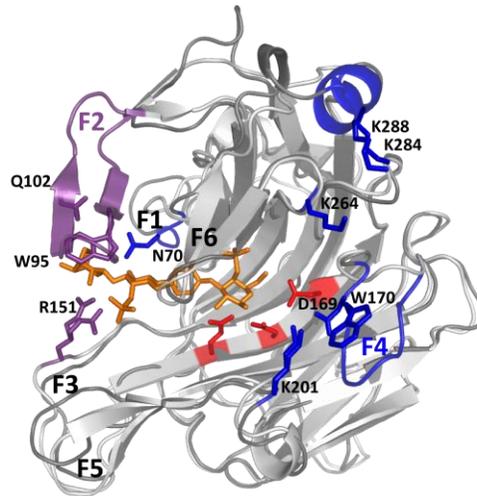


Figure 5

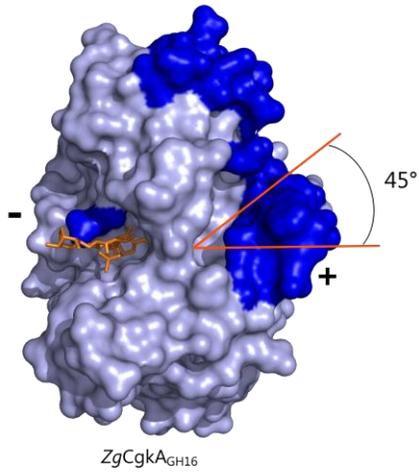
A



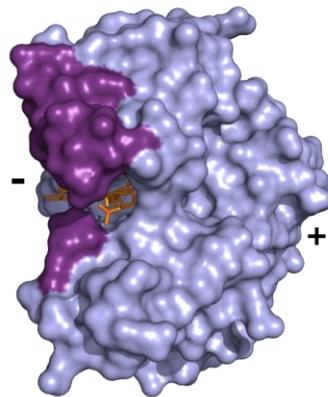
B



C

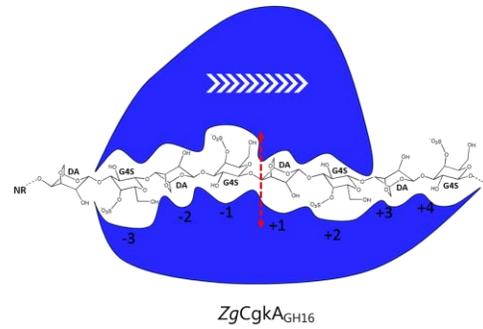


ZgCgkA_{GH16}

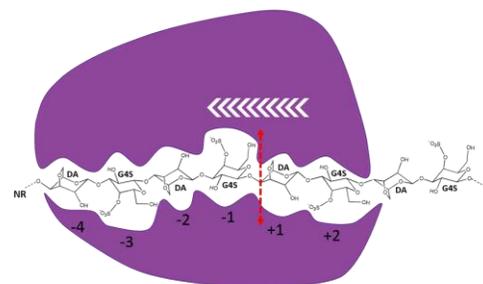


PcCgkA_{GH16}

D



ZgCgkA_{GH16}



PcCgkA_{GH16}

Figure 6

SUPPLEMENTARY DATA

Table SI: Kinetic parameters and data plot calculated with hyper32 software, with hyperbolic regression model applying weighting

Enzyme name (concentration, nM)	Vmax (M.s ⁻¹)	Km (M)	Hyperbolic regression plot
<i>PcCgk</i> _{AGH16} (36.9)	5.006×10^{-6} $\pm 2.409 \times 10^{-7}$	6.830×10^{-7} $\pm 7.486 \times 10^{-8}$	
<i>ZgCgk</i> _{AGH16-CBM16} (10.1)	6.650×10^{-6} $\pm 4.665 \times 10^{-7}$	1.041×10^{-6} $\pm 1.343 \times 10^{-7}$	
<i>ZgCgk</i> _{AGH16} (5.87)	5.371×10^{-6} $\pm 2.683 \times 10^{-7}$	8.218×10^{-7} $\pm 8.741 \times 10^{-8}$	
<i>PcCgk</i> _{AGH16-R92A} (37.0)	6.775×10^{-6} $\pm 3.784 \times 10^{-7}$	8.992×10^{-7} $\pm 1.011 \times 10^{-7}$	

Chapitre 2 : Etude comparée de deux κ -carraghénases

<p><i>PcCgkA</i>_{GH16-R151A} (15.4)</p>	<p>7.371×10^{-6} $\pm 4.932 \times 10^{-7}$</p>	<p>1.207×10^{-6} $\pm 1.468 \times 10^{-7}$</p>	
<p><i>PcCgkA</i>_{GH16-Q171A} (30.8)</p>	<p>3.700×10^{-6} $\pm 1.846 \times 10^{-7}$</p>	<p>3.847×10^{-7} $\pm 5.877 \times 10^{-8}$</p>	
<p><i>PcCgkA</i>_{GH16-R196A} (169.3)</p>	<p>1.038×10^{-6} $\pm 7.204 \times 10^{-8}$</p>	<p>1.154×10^{-6} $\pm 1.404 \times 10^{-7}$</p>	
<p><i>PcCgkA</i>_{GH16-W266A} (92.5)</p>	<p>2.530×10^{-6} $\pm 3.480 \times 10^{-7}$</p>	<p>2.632×10^{-6} $\pm 4.964 \times 10^{-7}$</p>	

TABLE SII. Primers sequences used for the cloning of the different κ -carrageenases in the study. Mutated codon are underlined.

	Forward	Reverse
<i>Pc CgkA</i> _{GH16-E168D}	5' GAAATAGATGTAGTT <u>ATT</u> CTACAACAATTCGATG 3'	5' CATCGAATTGTTGTAGAA <u>ATA</u> ACTACATCTATTTC 3'
<i>Pc CgkA</i> _{GH16-R92A}	5' CTAAGCGAGAATCTCATCAAG <u>GCT</u> ACATTCTGGGATGGCTGT 3'	5' ACAGCCATCCCAGAAATGT <u>AGC</u> TTGATGAGATTCTCGCTTAG 3'
<i>Pc CgkA</i> _{GH16-R151A}	5' TGGATGTATAGCACCA TTGAT <u>GCA</u> TCA TTAACGAAAAGAAGGGGA 3'	5' TCCCCTTCTTTCGTTAATGAT <u>TCC</u> ATCAATGGTGCTATACATCCA 3'
<i>Pc CgkA</i> _{GH16-Q171A}	5' GCGAAATAGACGTAGTGGAGTTAACT <u>GCA</u> AAAAAGTGCA GTGAGAGAGT 3'	5' ACTCTCTCACTGCAC TTTT <u>TGC</u> AGTTAACTCCA CTACGTCATTTTCGC 3'
<i>Pc CgkA</i> _{GH16-R196A}	5' ATGGAAAACCAACATGGATG <u>CCG</u> CCA GGGTCTTTCCGCAG 3'	5' CTGCGGAAAAGACCC TG <u>CCG</u> CCATCCATGTTGGTTTCCAT 3'
<i>Pc CgkA</i> _{GH16-R260A}	5' CTCACATTATCACAAGGCTTA <u>GCG</u> GCCGCCATACACAATGG 3'	5' CCATTGTGTATGCGGCGC <u>CGC</u> TAAGCCTTGTGATAATGTGAG 3'
<i>Pc CgkA</i> _{GH16-W266A}	5' GCGCGCCGCATACACAAG <u>GCG</u> AAATGTAATCAATTTTACCCA 3'	5' TGGGTA AAAATTGATTACATTT <u>CGC</u> TTGTGTATGCGGCGCGC 3'
<i>Zg CgkA</i> _{GH16}	5' AAAAAAGGATCCCAACAACCTACGAAGACTTCAAATC 3'	5' TTTTTCTGCAGTATGATTTTTCCAGACCCGAACGTA 3'
<i>Zg CgkA</i> _{GH16-CBM16-PofSS}	5' AAAAAAGGATCCCAACAACCTACGAAGACTTCAAATC 3'	5' TTTTTCTGCAGTACTCCACGATATCTTTTTTGAAC 3'

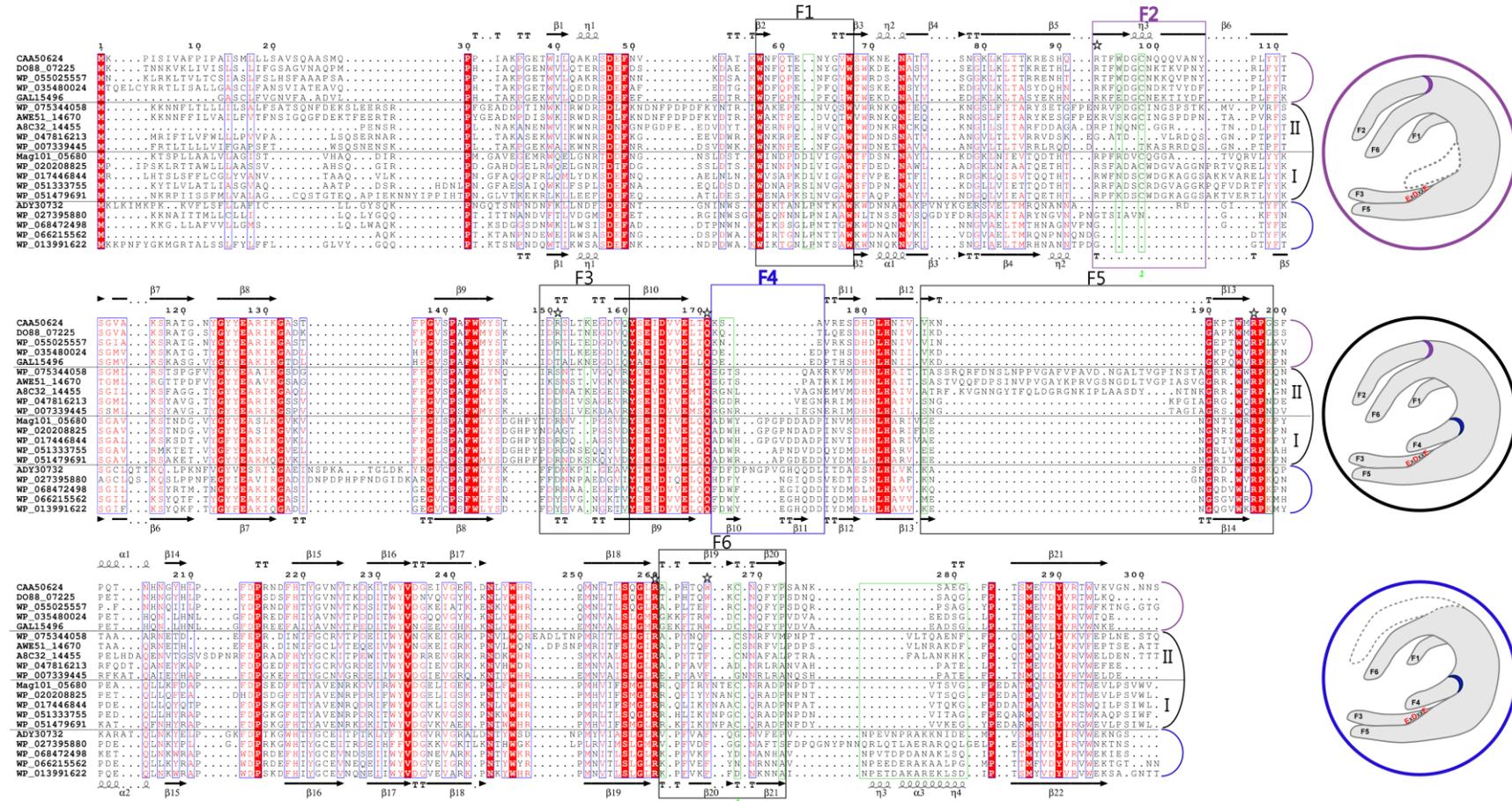


FIGURE S1. Alignment of 20 sequences of κ -carragenases from clades A-B-C-D with the one of *P. carrageenovora* as first sequence and *Z. galactanivorans* as last sequence. Annotation code is the same as in Figure 2A, and color and identification codes the same as Figure 3. Five sequences have been chosen for each clade. Horizontal lines delimitate the clades and the corresponding structural schemes are drawn on the right.

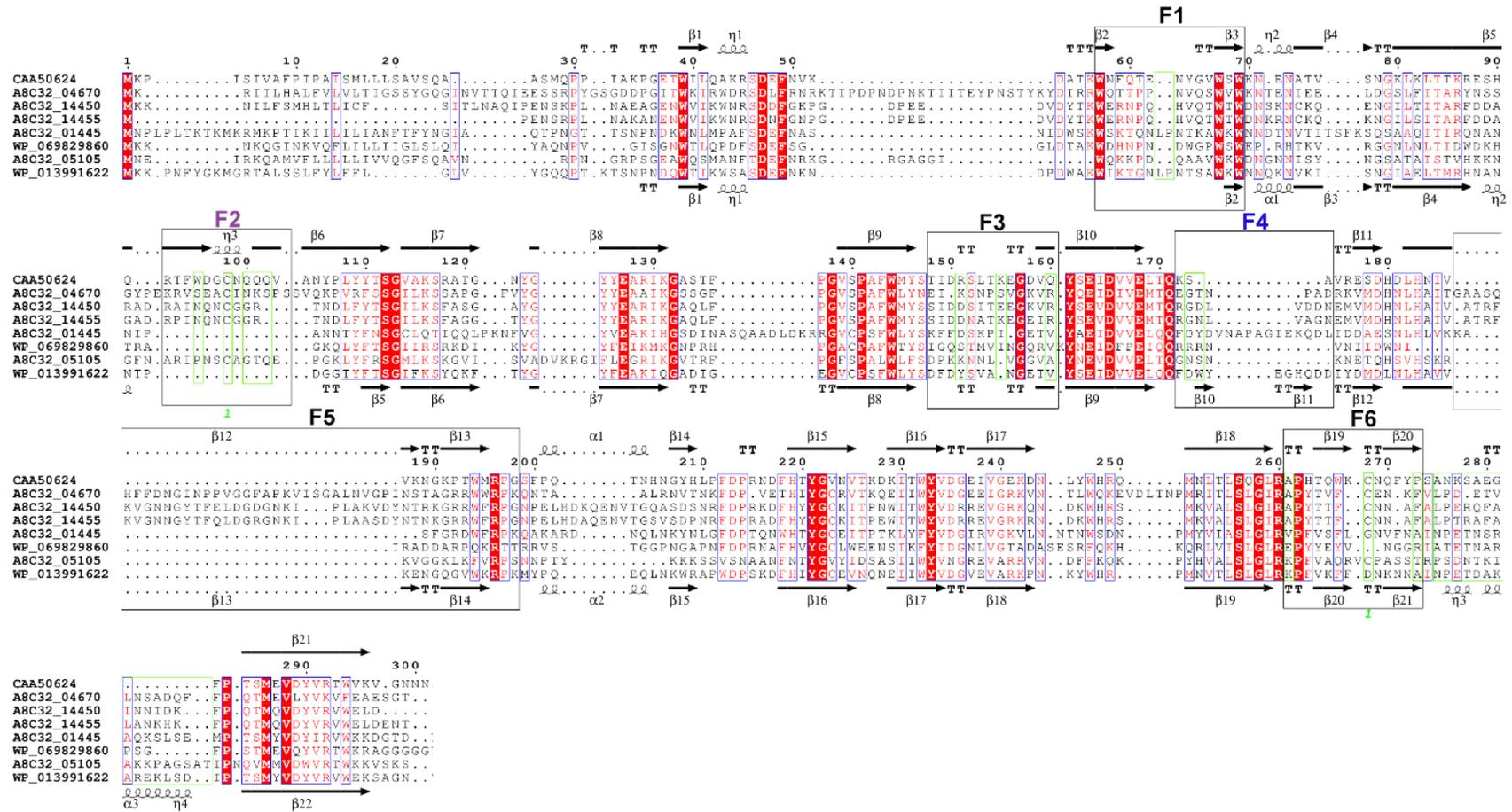


FIGURE S2. Alignment of the six κ -carragenases of *Algibacter* sp. SK-16 with those of *P. carrageenovora* (first sequence) and *Z. galactanivorans* (last sequence). Annotation code is the same as in Figure 2A.

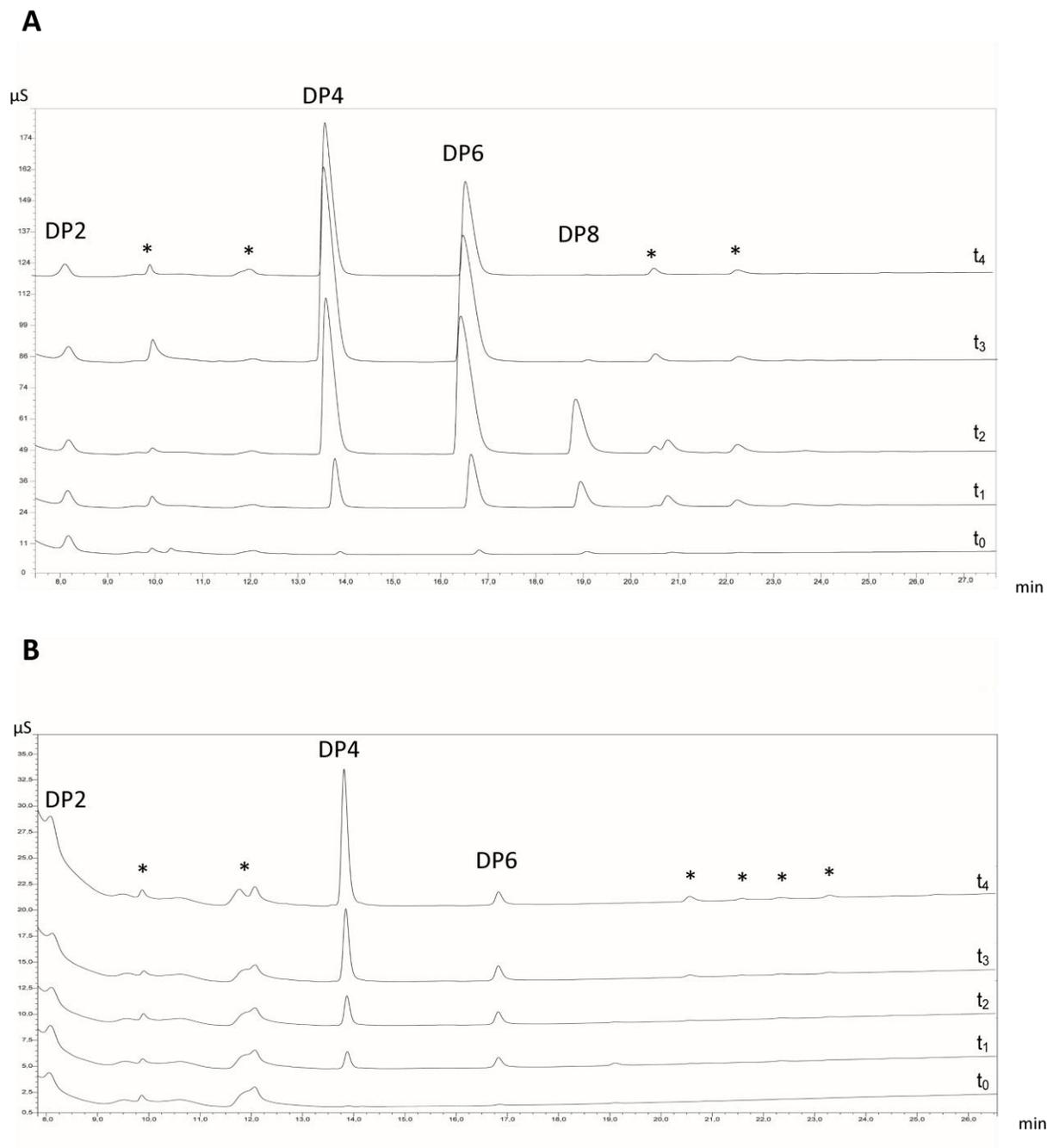


FIGURE S3 Examples of chromatograms obtained for the degradation products of **ZgCgkAGH16 in solution (A)** and **ZgCgkAGH16-CBM16-PorSS in microgel (B)**. In solution we can notice the presence of a peak corresponding to DP2 the height of which remains constant throughout the kinetic experiment. Conversely, in microgel DP2 co-elutes with the salt peak of KCl obstructing a clear interpretation. Asterisks symbolize uncharacterized contaminant peak. Curves correspond to kinetics at $t_0=0$, $t_1=15$ min, $t_2=45$ min, $t_3=24$ h, $t_4=1$ week

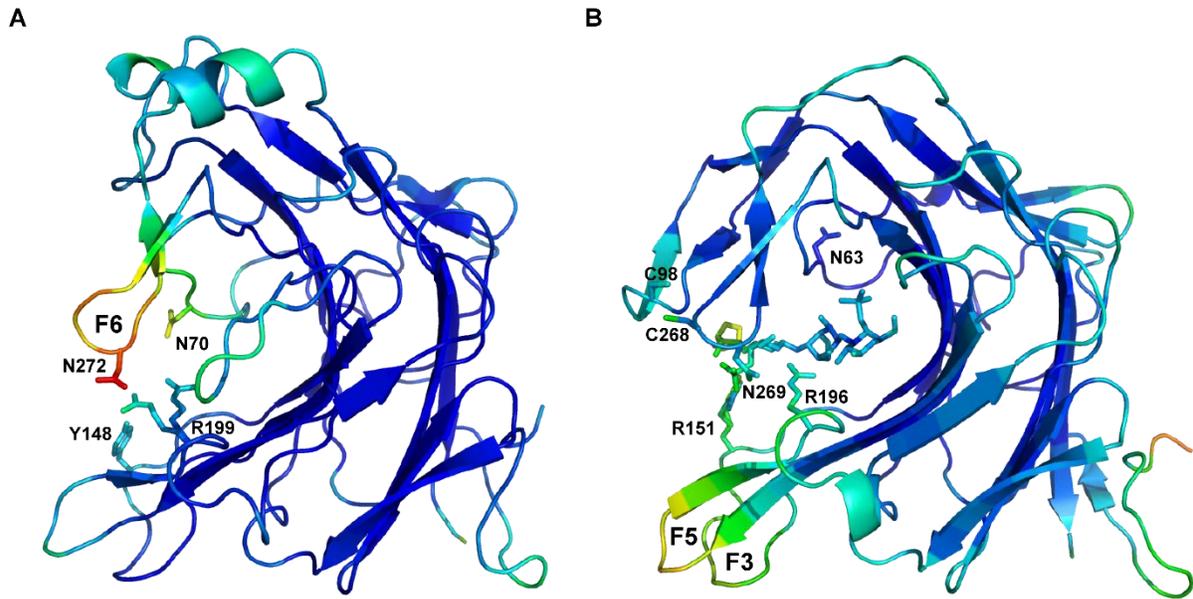


FIGURE S4. Structures of *ZgCgkAGH16* (A) and *PcCgkAGH16-E168D* in complex with a κ -neocarratetraose (B) tinted by B factor intensity. Key residues are annotated, as well as the loops with the highest B-factors.

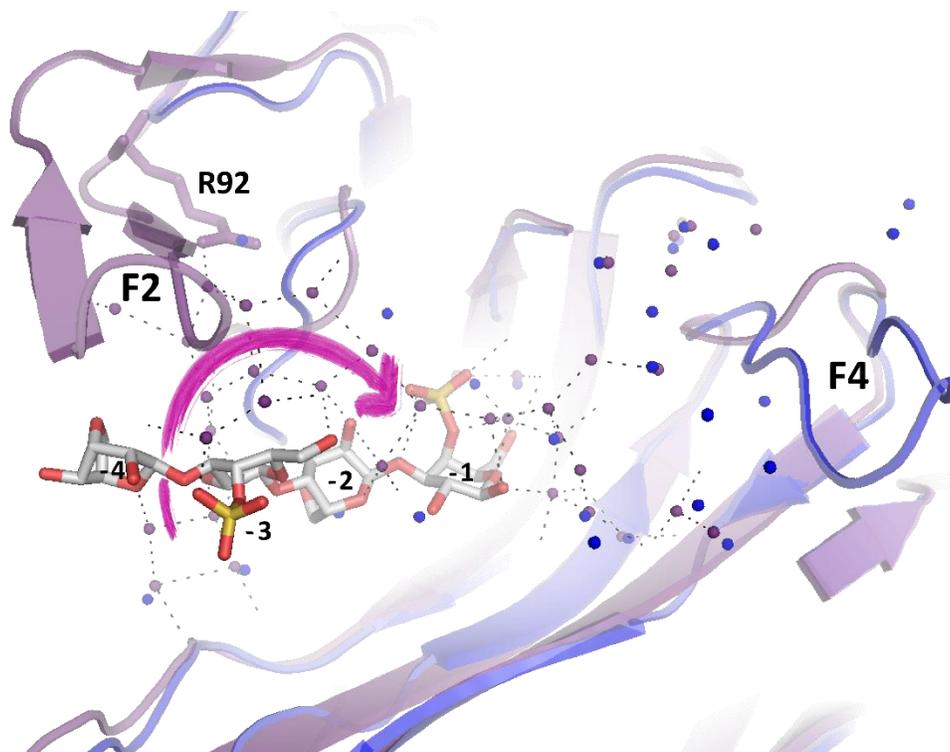
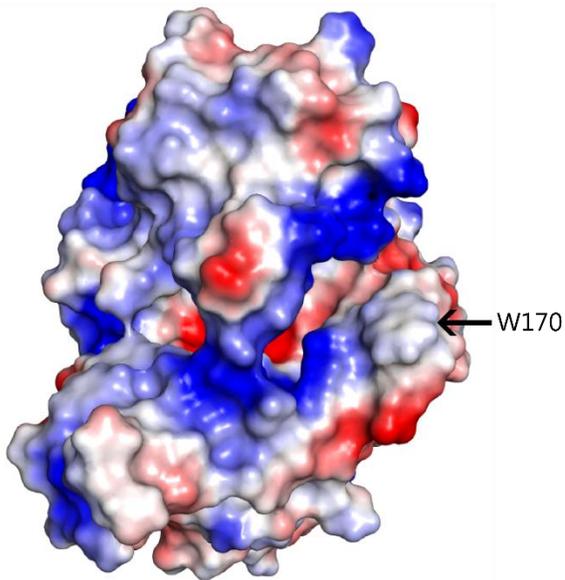


FIGURE S5. Water mesh within the catalytic tunnel of *PcCgkAGH16-E168D* (cartoon in slate and water in purple) superposed with *ZgCgkAGH16* and its associated water molecules (in blue and marine blue respectively). Possible hydrogen bonding are represented as black dashes. The R92 is shown for its role in the structuration of the water mesh at the entrance of the channel. Purple arrow underline the roll shape of the water mesh around the substrate. Neocarratetraose (DA-G4S-DA-G4S) is colored in light grey and the moieties are labelled with their corresponding positioning in the catalytic subsites from -4 to -1

A



B

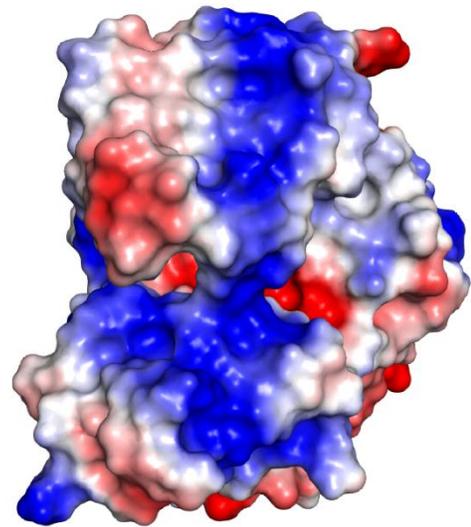


FIGURE S6. Electrostatic map of ZgCgkAGH16 (A) and PcCgkAGH16-E168D (B). Color code spreads from blue to red for basic to acidic residues. The position of W170 in ZgCgkAGH16 is indicated by an arrow.

Chapitre 3

Deux enzymes de bactéries marines spécifiques du β -carraghénane révèlent l'existence d'une nouvelle famille de glycoside hydrolase apparentée aux GH42

Chapitre 3 : Deux enzymes de bactéries marines spécifiques du β -carraghénane révèlent l'existence d'une nouvelle famille de glycoside hydrolase apparentée aux GH42

I. Introduction générale sur le β -carraghénane et sa dégradation

Le β -carraghénane est le seul motif de la famille des carraghénanes qui soit dépourvu de groupement sulfate. Il est présent en mélange avec des motifs κ au sein de la paroi de plusieurs familles d'algues rouges appartenant à l'ordre des Gigartinales, réparties dans les familles des Furcellariaceae : *Furcellaria lumbricalis* (Knutsen and Grasdalen 1987), des Endocladiaceae : *Endocladia muricata* (Renn *et al.* 1993), des Tichocarpaceae : *Tichocarpus crinitus* (Correc *et al.* 2012), des Acrotylaceae : *Hennedya crispa* (Chiovitti *et al.*, 2008) et des Solieriaceae : *Betaphycus gelatinus* = *Eucheuma gelatinae* (Greer and Yaphe 1984a) et *Eucheuma speciosa* (Renn *et al.* 1993). La faible teneur en sulfate des hybrides β/κ entraîne une capacité à la gélification à des concentrations plus faibles en KCl. De plus, la répartition des différents motifs β et κ au sein de ces carraghénanes hybrides génère des différences importantes dans leurs propriétés gélifiantes. Ainsi, une alternance régulière des motifs β et κ va permettre l'obtention de gel résistants, comme dans le cas du furcellarane extrait de *F. lumbricalis*. Au contraire, l'organisation des motifs β et κ en blocs distincts génère des gels peu résistants, comme dans le cas de *B. gelatinus* ou *T. crinitus* (Correc *et al.*, 2012; Yermak *et al.*, 1999). Cependant, malgré ces propriétés rhéologiques intéressantes, il reste le parent pauvre de la famille des carraghénanes, aussi bien en termes d'applications industrielles, que de nombre de publication portant sur son étude, comparativement au trois autres carraghénanes majoritaires, les κ -, ι - et λ -carraghénanes.

Outre la présence de motifs β au sein des parois de carraghénophytes, le travail précurseur de McLean et Williamson sur le catabolisme du carraghénane chez la bactérie *Pseudoalteromonas carrageenovora* a permis de mettre en évidence, dès 1979, la conversion du motif κ de l'extrémité non réductrice d'un néo-carratétraose en motif β , et ce par l'action d'une sulfatase non sécrétée dans le milieu de culture (McLean and Williamson 1979). Ils ont ensuite démontré l'existence d'une autre enzyme, également non sécrétée, et capable de reconnaître et de cliver ce motif β en amont d'un motif κ (McLean and Williamson 1981). Enfin, ils ont montré que la sulfatase précédemment caractérisée était également capable de convertir le disaccharide κ relargué en

disaccharide β . Bien que les gènes codant pour ces enzymes étaient inconnus à cet époque, ces travaux ont permis d'illustrer la stratégie générale des bactéries marines hétérotrophes (BMH) vivant au contact des macroalgues et qui, pour s'adapter à l'hétérogénéité de leurs substrats, expriment un panel d'enzymes agissant de façon conjointe et/ou séquentielle.

Cette stratégie observée de façon empirique a été confirmée par le séquençage récent de plusieurs génomes de BMH, dans lesquels de nombreuses enzymes impliquées dans des voies de dégradation des polysaccharides sont regroupés au sein d'îlots de gènes ('clusters') généralement co-régulés par un facteur de transcription. Récemment, une nouvelle sous-famille de GH16, active sur le furcellarane mais pas sur le κ -carraghénane pur, a été mise en évidence chez la bactérie *Paraglaciecola hydrolytica* par Mikkel Schultz-Johansen, au sein de l'équipe de Peter Stougaard à l'université de Copenhague (Schultz-Johansen *et al.*, in prep). Cette GH16 a été baptisée furcellaranase et semble spécifique du carraghénane désulfaté (β). Le cluster contenant cette GH16 code également une κ -carraghénase (GH16), une ι -carraghénase (GH82), une 'GH42-like', une 'Ig-like protein' et trois sulfatases. Or, chez la bactérie *P. carrageenovora*, il existe un regroupement similaire de gènes sur le plasmide dit 'carraghénolytique', mais présentant une synténie dégénérée. La présence au sein de ce cluster de gènes codant pour les enzymes du catabolisme des monosaccharides 3,6-anhydro-D-galactose et D-galactose, conjointement à la présence de trois sulfatases et des β , κ et ι -carraghénases, nous a amené à émettre l'hypothèse suivante : la dégradation des β , κ , et ι -carraghénanes aurait lieu de façon conjointe, aboutissant au même produit intermédiaire, de type β -carrabiose. De même les monosaccharides terminaux de 3,6-anhydro-D-galactose et D-galactose seraient dégradés par des voies parallèles mais se rejoignant pour les étapes finales du métabolisme. Cette hypothèse est également soutenue par Gobet *et al.* dans leur récente publication du génome de *P. carrageenovora*. Le modèle théorique proposé pour ces voies de dégradation conjointes est illustré en Figure 37.

II. Modèle de dégradation conjointe des ι -, κ - et β -carraghénanes par *P. carrageenovora*.

La première étape consiste en la dépolymérisation initiale du substrat en oligosaccharides de degrés de polymérisation compatibles avec leur importation dans le périplasme bactérien. C'est la κ -carraghénase PcCgkA qui effectue ce travail (étape 1 sur la figure 37). Sa caractérisation

détaillée dans plusieurs travaux réalisés au sein du laboratoire, a permis de déterminer son mode d'action précis - mécanisme endo-processif sur substrat gélifié avec rétention de la configuration anomérique, accompagné de la libération majoritaire d'oligosaccharides DP4 - (Potin *et al.* 1995; Barbeyron *et al.* 1998; Lemoine *et al.* 2009) ainsi que sa structure cristallographique (Michel *et al.* 2001a) (Matard-Mann *et al.*, 2017). Il a également été montré que cette enzyme est capable de générer des oligosaccharides de types ι/κ et β/κ en présence des polysaccharides hybrides correspondants (Bellion *et al.* 1982; Correc *et al.* 2012). Le schéma proposé en figure 37 commence donc par la représentation d'un carraghénane chimérique constitué de motifs ι , κ et β , symbolisant les différents substrats naturels que la κ -carraghénase a la capacité de dégrader.

Une fois les oligosaccharides générés dans le milieu, ils vont pouvoir être internalisés dans le périplasma de *P. carrageenovora* via des transporteurs TonB dépendants (TBDT) codés par les gènes mp0026, mp0031 et mp0046 (étapes 2.1, 2.2, 2.3). La protéine prédite comme possédant un repliement de type Ig_like (mp0027, non représentée sur le schéma) pourrait jouer un rôle similaire au SusD-like des *Bacteroidetes* en fixant les oligosaccharides pour faciliter leur internalisation par le TBDT.

Une fois dans le périplasma, différentes enzymes continuent la dégradation des oligosaccharides (étapes 3 à 7 du schéma).

Contrairement à la κ -carraghénase, aucune ι -carraghénase n'a jamais été observée sécrétée dans le milieu de culture de *P. carrageenovora*. Il a été rapporté chez d'autres souches bactériennes qu'une activité ι -carraghénolytique était observée soit dans le milieu de culture soit dans le culot bactérien, associé aux membranes (Bellion *et al.* 1982; Greer and Yaphe 1984b; Potin *et al.* 1991). Il est donc possible que, si elle est produite chez *P. carrageenovora*, ce soit dans le périplasma, du fait de la présence d'un peptide signal à l'extrémité N-terminale de sa séquence, clivable par la signal-peptidase SpI. Le gène codant pour cette enzyme a été cloné durant cette thèse. Malheureusement, la protéine recombinante est exprimée sous forme insoluble, empêchant son étude biochimique. Cependant la ι -carraghénase d'*Alteromonas fortis* a fait l'objet d'études biochimiques et structurales détaillées – montrant un mécanisme endoprocessif sur substrat gélifié avec inversion de la configuration anomérique et libération d'oligosaccharides DP4 et DP6 - (Barbeyron *et al.* 2000; Michel *et al.* 2001a, 2003; Rebuffet *et al.* 2010). Bien que celle de *P. carrageenovora* ne présente pas une forte identité de séquence avec cette dernière (24 % sur seulement 56 % de la séquence), elles appartiennent toutes deux à la famille des GH82.

Chapitre 3 : Etude comparée de deux β -carrabiose hydrolases

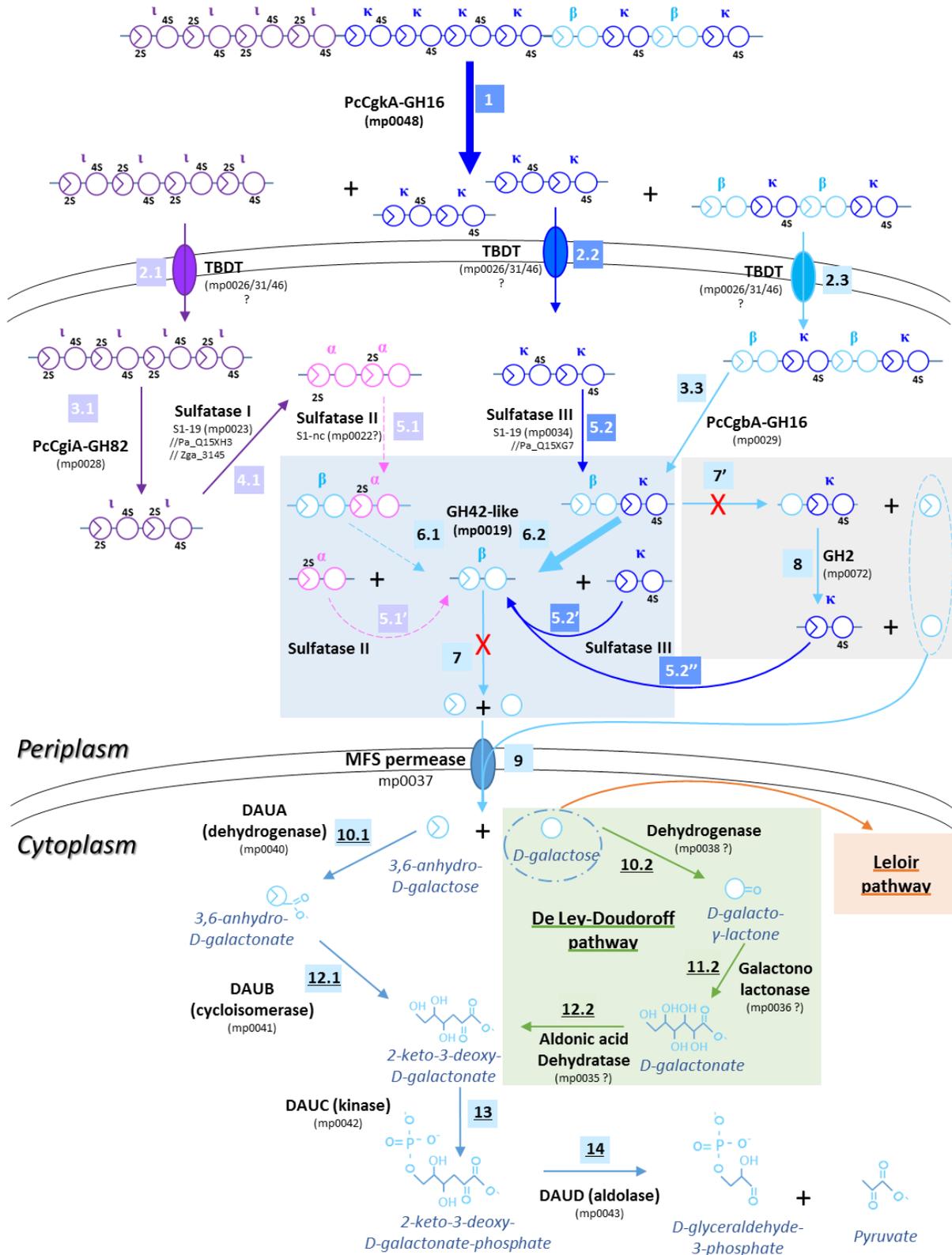


Figure 38 : Proposition de schéma théorique des voies de dégradation conjointes des 1-, κ - et β -carraghénanes par les enzymes du plasmide carraghénolytique de *P. carrageenovora*.

A ce jour, seule l'activité sur ι -carraghénane pur a été observée pour les enzymes de cette famille. Par conséquent, on peut penser que leurs produits de dégradation seront similaires (étape 3.1 du schéma). Concernant la dégradation des oligos β/κ , la première β -carraghénase de la famille des GH16 vient tout juste d'être identifiée chez *P. hydrolytica* (Schultz-Johansen, en préparation), mais les résultats préliminaires laissent penser à un mode d'action similaire aux deux carraghénases précédentes, tout du moins en ce qui concerne les produits de dégradation générés (étape 3.3). Afin de tenter de prédire le rôle des trois sulfatases du cluster, leurs séquences ont été comparées avec les séquences des sulfatases de *Z. galactanivorans* et de *Pseudoalteromonas atlantica* ayant déjà fait l'objet de caractérisations biochimiques. Ces trois sulfatases appartiennent à la famille S1 des sulfatases. Deux d'entre-elles appartiennent à la sous-famille 19 (S1-19), mais la troisième possède une séquence trop originale pour pouvoir être rattachée à l'une des sous-familles déjà identifiées au sein de la famille S1 (S1-nc).

La première sulfatase S1-19 du PUL (mp0023) possède le plus fort pourcentage d'identité de séquence avec les sulfatases également de type S1-19 Zga_3154 (54 %) de *Z. galactanivorans* et Q15XH3 (69 %) de *P. atlantica*. Or ces deux enzymes ont été caractérisées comme des 4S- ι -carraghénane-sulfatases, c'est-à-dire capables de retirer le groupement sulfate du carbone C4 de l'unité G d'un motif ι -carraghénane, générant ainsi un motif α -carraghénane (Préchoux *et al.* 2013; Ficko-blean *et al.*). On peut donc émettre l'hypothèse que, comme démontré pour le PUL carraghénolytique de *Z. galactanivorans*, les tétrasaccharides ι - ι sont convertis en tétrasaccharides α - α via la sulfatase mp0023 (Sulfatase I à l'étape 4.1).

La seconde S1-19 du PUL (mp0034) présente quant à elle un fort pourcentage d'identité de séquence avec la S1-19 Q15XG7 (64 %) de *P. atlantica*. Or cette enzyme a été caractérisée comme une 4S- κ -carraghénane-sulfatase, c'est-à-dire retirant le groupement sulfate du carbone C4 de l'unité G d'un motif κ -carraghénane, générant ainsi un motif β -carraghénane (Préchoux *et al.* 2016).

Légende de la figure 37 : Légende des flèches : en gras : activités enzymatiques démontrées expérimentalement ; en traits simples : activités théoriques prédites par homologies de séquences et démontrées chez d'autres organismes ; en pointillés : activités hypothétiques à vérifier. Entre parenthèses : gènes associés aux activités enzymatiques, en gras si vérifiés expérimentalement, et suivis d'un '?' si hypothétiques. Les étapes spécifiques aux différents carraghénanes sont colorées en violet/rose pour les ι/α et en bleu foncé/clair pour les κ/β . La voie de dégradation via les GH127/129 et la GH2 (Ficko-Blean *et al.*, 2017, sous presse) est représentée sur fond gris, la voie de dégradation via la GH42-like sur fond bleu et les voies de dégradation du galactose proposées par de Ley-Doudoroff sur fond vert et par Leloir sur fond orange.

On peut donc émettre l'hypothèse que mp0034 correspond à la sulfatase décrite par McLean et Williamson et permet la conversion du tétrasaccharide κ - κ en un hybride de type β - κ (Sulfatase III à l'étape 5.2).

La troisième sulfatase du PUL (mp0022) partage un pourcentage d'identité maximal avec la sulfatase non classée Q15XH4 (67 %) de *P. atlantica* dont l'étude a été initiée au cours de la thèse d'Aurélié Préchoux, mais dont le mode d'action précis reste encore inconnu. Elle semblerait cependant active sur un substrat de type ι -carraghénane (Préchoux 2012).

Au vu de ces éléments, nous proposons donc que la sulfatase non classée mp0022 soit responsable de la conversion d'un motif α en motif β en retirant le groupement sulfate du carbone C2 de l'unité D, aboutissant à la formation d'un tétrasaccharide de type β - α voir de type β - β (Sulfatase II à l'étape 5.1). C'est à cette étape qu'entre en jeu la GH42-like (mp0019). Comme détaillé dans le l'ébauche d'article ci-après, nous avons pu montrer que cette GH est capable de reconnaître et de cliver un motif néo- β -carrabiose du côté non réducteur d'un hexasaccharide hybride de type β - κ - κ . Nous proposons donc que la GH42-like corresponde à la β -carrabiose hydrolase décrite par McLean et Williamson, et soit responsable de la conversion du tétrasaccharide hybride de type β - κ (généré par la sulfatase III) en deux disaccharides, un β -carrabiose et un κ -carrabiose (étape 6.2). Ce dernier sera par la suite converti en β -carrabiose par cette même sulfatase (étape 5.2'). De façon plus hypothétique, nous proposons également que la GH42-like puisse également cliver le motif β d'un tétrasaccharide de type β - α ou de type β - β (étape 6.1), avec si besoin conversion du disaccharide α généré en β grâce à l'action de la sulfatase II (étape 5.1').

A ce stade, un verrou de taille se pose afin de compléter la voie métabolique. En effet, le disaccharide de départ de l'étape terminale de dégradation est un β -carrabiose. Or les enzymes capables de cliver l'unité 3,6-anhydro-D-galactose connues à ce jour appartiennent à la famille des GH127 et GH129-like, familles absentes du cluster de *P. carrageenovora* et de son génome global. Cette absence est d'autant plus surprenante que ce cluster renferme sept enzymes homologues à des enzymes connues pour catalyser les étapes finales de la conversion des monosaccharides de 3,6-anhydro-D-galactose et de D-galactose. En effet, il existerait une voie de dégradation du galactose alternative à la voie classique de Leloir, proposée par de Ley et Doudoroff (De Ley and Doudoroff 1957) (en orange sur le schéma), basée sur l'action successive d'une déshydrogénase (étape 10.2) convertissant le D-galactose en D-galacto- γ -lactone, puis d'une galactono-lactonase (étape 11.2) générant du D-galactonate et enfin d'une galactonic acid déshydratase (étape 12.2)

formant le 2-kéto-3-déoxy-D-galactonate. Ce composé est commun avec la voie de dégradation du 3,6-anhydro-D-galactose détaillée par Ficko-blean *et al.* dans le contexte du catabolisme du carraghénane chez *Z. galactanivorans* (Ficko-blean *et al.*, in press). Elle fait appel aux enzymes DAUA (étape 10.1) et DAUB (étape 12.1) pour obtenir le 2-kéto-3-déoxy-D-galactonate puis DAUC (étape 13) et DAUD (étape 14) pour générer le pyruvate et le D-glycéraldéhyde-3-phosphate qui seront ensuite intégrés au cycle classique de la glycolyse.

Toujours d'après ces mêmes travaux chez *Z. galactanivorans*, la voie de dégradation des carraghénanes alternative à celle utilisant la GH42-like (en gris sur le schéma) fait appel à l'action successive de GH127 et GH129 libérant un 3,6-anhydro-D-galactose (étape 7') puis d'une GH2 libérant un galactose (étape 8). Or il y a bien une GH2 codée sur le plasmide carraghénolytique de *P. carrageenovora* mais en aval du cluster et codée dans le sens inverse du cadre de lecture, aux côtés de gènes impliqués dans la dégradation du λ -carraghénane. Si cette voie alternative existait chez *P. carrageenovora*, elle rejoindrait la voie de la GH42-like à l'étape 5.2''. Cependant, dans ces deux voies, les GH127/129-like sont nécessaires, alors qu'elles sont absentes du génome de *P. carrageenovora*.

III. Conclusion

Ce cluster du PUL carraghénolytique de *P. carrageenovora* renferme donc toutes les enzymes nécessaires aux étapes initiales et terminales du catabolisme des ι -, κ - et β -carraghénanes, à l'exception des enzymes clés (GH127/129-like) permettant de cliver le β -carrabiose en sucres simples. Cela pourrait expliquer l'incapacité de *P. carrageenovora* à pousser en présence de ι -carraghénane et de furcellaran (β/κ -carraghénane) comme seule source de carbone. La croissance fortement différée observée en présence de κ -carraghénane (Gobet *et al.*, en préparation) irait également dans ce sens, tout en signifiant que la bactérie finit par trouver une solution alternative afin d'utiliser le κ -carraghénane. Cependant, on ne peut exclure la possibilité qu'il existe une autre enzyme capable d'effectuer l'étape manquante, n'appartenant pas à la famille des GH127/129. Dans ce cas, l'absence ou le retard de croissance de *P. carrageenovora* pourrait s'expliquer par un phénomène de co-régulation des gènes du métabolisme de ces substrats avec d'autres, comme ceux du λ -carraghénane par exemple, et nécessitant une induction par celui-ci pour être exprimés. Cette hypothèse est compatible avec les observations réalisées lors des travaux de Johnston et McCandless où même en présence de λ -carraghénane comme seule source de carbone, une activité

κ -carraghénolytique était observée dans le milieu de culture de *P. carrageenovora* en plus de celle de la λ -carraghénase (Johnston and McCandless 1973). De même, dans leurs travaux précurseurs, Weigl et Yaphe ont montré que le maximum d'activité κ -carraghénolytique extracellulaire n'avait lieu qu'en phase stationnaire de croissance de *P. carrageenovora*, alors que l'activité λ -carraghénolytique suivait la courbe de croissance bactérienne (Weigl and Yaphe 1966). De façon similaire, lors des premiers travaux réalisés sur les carraghénases de *Z. galactanivorans*, les κ - et λ -carraghénases natives étaient sécrétées dans le milieu de culture et co-purifiées suite à la croissance en présence de λ -carraghénane (Potin *et al.* 1991). Cette hypothèse d'induction de l'expression de la voie κ -carraghénolytique par le λ -carraghénane pourrait être confirmée par une analyse RNAseq des produits de gènes induits lors de la culture de *P. carrageenovora* en présence de λ -carraghénane.

La présence de ce cluster incomplet chez *P. carrageenovora*, qui plus est sur un plasmide, pourrait être la trace d'un transfert horizontal de gènes ayant transmis de façon partielle à *P. carrageenovora* un ensemble de gènes impliqués dans le catabolisme des carraghénanes.

D'un point de vue évolutif, cela pourrait également représenter un cas de coopération entre BMH où la synthèse de certaines enzymes sécrétées de façon extracellulaires, comme la κ -carraghénase, favoriserait la croissance d'autres bactéries capables de métaboliser les oligosaccharides générés. En retour ces dernières pourraient contribuer à la croissance de *P. carrageenovora* en favorisant la libération d'oligosaccharides métabolisables par celle-ci comme des λ -carraghénanes. En effet, *P. carrageenovora* est capable de pousser en présence de λ -carraghénane comme seule source de carbone, ce qui signifie que la voie métabolique de celui-ci est différente des autres carraghénanes (de par l'absence de pont 3,6-anhydro) et fonctionnelle.

Ce cluster incomplet pourrait également refléter un phénomène de succession temporelle dans les populations microbiennes pour s'adapter à la variabilité de leur substrat, comme ce qui a été observé dans le cadre des blooms phytoplanctoniques par Teeling (Teeling *et al.* 2012, 2016). Ainsi, chez l'algue rouge *Chondrus crispus*, la nature des carraghénanes constituant sa paroi varie en fonction de son cycle de reproduction. Le gamétophyte est constitué d'un mélange de κ/ι -carraghénanes alors que le tétrasporophyte présente une teneur élevée en λ -carraghénane. Au vu des observations précédentes sur le 'patrimoine enzymatique' de *P. carrageenovora*, il serait tentant d'imaginer une alternance dans la composition de la microflore associée à *C. crispus* en fonction de son cycle de vie, avec une prolifération de la population de *P. carrageenovora* au stade

sporophytique de l'algue, qui décroîtrait ensuite au profit d'autres populations capables de métaboliser les motifs ι - et κ -carraghénane au cours de son stade gamétophytique. Une étude métagénomique du microbiote associé à *C. crispus* au cours de son cycle de vie permettrait de répondre à cette question.

Néanmoins, afin de tester certaines de nos hypothèses, le gène mp0019 ainsi que deux des trois gènes codant pour les sulfatases du cluster de *P. carrageenovora* (mp0022 et mp0023, codant respectivement pour la S1-nc et une S1-19) ont été clonés avec succès au cours de cette thèse. C'est sur cette GH42-like que nous avons concentré notre étude, en partenariat avec l'équipe danoise de Peter Stougaard (Université de Copenhague, Department of Microbial Ecology and Biotechnology) qui a travaillé sur son homologue chez *P. hydrolytica*. Les résultats obtenus indépendamment ont été rassemblés dans l'ébauche d'article ci-après, qui devra être complétée par l'étude en parallèle des deux enzymes sur des substrats identiques avant de pouvoir faire l'objet d'une publication. L'étude de ces deux enzymes homologues a permis de proposer la constitution d'une nouvelle famille de GH, distincte de la famille des GH42, et dont l'organisation phylogénétique laisse penser qu'une spécificité pour le β -carraghénane existe pour au moins une des sous-familles (Clade A1 dans l'article). Ce travail ouvre également la voie à l'étude des trois sulfatases de ce cluster, en particulier la sulfatase non classée. Si les hypothèses évoquées plus haut sont confirmées, ceci permettrait de proposer une nouvelle voie de dégradation des ι -, κ - et β -carraghénanes aboutissant de façon conjointe à la formation d'un intermédiaire β -carrabiose.

Two marine bacterial enzymes specific of β -carrageenan allow the discovery of a new family of glycoside hydrolases, distantly related to GH42.

Maria Matard-Mann^{1,3}, Mikkel Schultz-Johansen², Pernille Kjersgaard Bech², Aurélie Préchoux¹, Tristan Barbeyron¹, Gervan Michel¹, Mirjam Czjzek*¹ and Peter Stougaard²

¹ Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, Roscoff, Bretagne, France.

² Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen,

Denmark

³ Amadéite SAS, « Pôle biotechnologique » du Haut du Bois, 56580 Bréhan, France

* Corresponding author

ABSTRACT

We report here the biochemical characterization of two marine bacterial enzymes, which are specific of desulfated carrageenan. These enzymes hydrolyze the β -1,4-D-galactose linkage of neo- β -carrabiose from furcellaran a hybrid carrageenan composed of β - and κ -carrageenan motives. Distantly related to the GH42-family, a phylogenetical analysis of these enzymes reveals that they constitute a new independent family, currently composed of at least 76 members, all of bacterial origin, belonging mainly to marine *Proteobacteria* and *Bacteroidetes*. These enzymes constitute the first exo-carrageenases ever cloned and uncover a new step in the carrageenolytic degradation pathway of marine bacteria.

KEYWORDS: carrageenan, beta-galactosidase, marine bacteria, GH42-like, sulfated polysaccharide, furcellaran

ABBREVIATIONS: DP, degree of polymerization; GH, glycoside hydrolase

For submission to FEBS letters <i>Research letters</i>

INTRODUCTION

Polysaccharides from marine algae play a key role in the global carbon cycle of the ocean. The carbon entrapped in cell walls of photosynthetic organisms is exported out of the cycle when organic matter sinks in the ocean or is reinjected in the cycle when metabolized by microorganisms and ‘higher’ consumers in the food chain (Azam and Malfatti 2007; Arnosti 2014). In order to extract the carbon stored in the polysaccharidic lattice or algal cell wall, marine bacteria have developed a wide set of enzymes adapted to the diversity and the complexity of algal polysaccharides (Michel and Czjzek 2013; Martin *et al.* 2014).

Carrageenans are amongst the most widely known algal polysaccharides synthesized by species of red macroalgae belonging to the order *Gigartinales* and their thickening and gellification properties have been harnessed essentially by the food industry since the 1930’s (Glicksman 1987; Bemiller 2009; Bixler and Porse 2010). Carrageenans are a family of linear D-galactans linked with alternating α -1,3 and β -1,4 glycosidic bonds. They are classified in families according to the presence of a 3,6-anhydrogalactose and to the number and position of sulfation motives per disaccharide unit. The main structures are called beta- (β -, no sulfate group), kappa- (κ -, one), iota- (ι -, two), and lambda- (λ -, three) carrageenans (Figure 1) (Craigie 1990). However, carrageenans naturally occurring into the algal cell walls are generally of hybrid nature, made of complex combinations of different types of carrageenans. Furcellaran, a polysaccharide originally extracted from *Furcellaria lumbricalis*, is an example of such a hybrid polymer, made of beta and kappa motives (Knutsen and Grasdalen 1987; Correc *et al.* 2012). In their seminal paper of 1981, McLean and Williamson have demonstrated the presence of an enzyme able of removing a beta-carrabiose motif from a hybrid tetrasaccharide of β/κ -carrageenan in a marine gammaproteobacterium, *Pseudoalteromonas carrageenovora* 9^T (ATCC 43555^T) (McLean and Williamson 1981). To our knowledge, until the study of Schultz-Johansen *et al.* in 2017 (in prep), no other study reporting this activity nor identifying associated genes have been published. Here we report the cloning of the historical neocarratetraose 4-O-monosulfate β -hydrolase (or neo- β -carrabiose hydrolase) of *P. carrageenovora* 9^T, along with a close homologous in the recently sequenced bacteria *Paraglaciecola hydrolytica* S66^T (Schultz-johansen *et al.* 2016), and their biochemical comparison. A phylogenetical analysis of these enzymes also reveals that they constitute a new family of glycoside hydrolases, distantly related to that of family GH42.

RESULTS AND DISCUSSION

Phylogenetical analysis reveals the existence of a new family of GH distantly related to GH42

- The genome of the marine bacterium *Paraglaiecola hydrolytica* S66^T (*Ph*) has recently revealed the existence of a new subfamily in the GH16 family, namely furcellaranases that are carrageenases acting on Furcellaran, a hybrid polymer of kappa/beta-carrageenan (Schultz-Johansen *et al.*, 2017, in prep). These enzymes were shown to be inactive on polysaccharide of pure κ -carrageenan but active on the hybrid κ/β -carrageenan, and thus constitute a new kind of endolytic enzyme specific of partially desulfated carrageenan. The gene cluster encoding these enzymes also contains many other enzymes involved in the carrageenan degradation pathway (Figure 2A). Interestingly, one gene product of this cluster (Ph1657), annotated as a GH42 β -galactosidase, presents 61 % of identity (76 % of similarity) with the protein encoded by the gene mp0019 of the carrageenolytic plasmid of *Pseudoalteromonas carrageenovora* 9^T (*Pc*), annotated as a distantly related GH42 (Gobet *et al.*, 2017, in prep) (Figure 2B). Blast searches with the sequence of the *P. carrageenovora* enzyme and the alignment of its closest homologous sequences allowed building a phylogenetic tree, which reveals a clear cut with the three sequences of ‘true’ GH42 used as an outside group (Figure 3). All the sequences analyzed come from bacterial origin. Based on the highest bootstrap values of the deepest nodes, two clades can be defined in this new family of GH (defined as “GH42-like”), clade A being composed of 45 sequences and clade B of 31 sequences which all belong to the phylum of *Bacteroidetes*. In clade A, 36 % of sequences belong to phyla of *Proteobacteria*, 24 % to *Bacteroidetes*, 9 % to *Firmicutes* and 4 % to *Planctomycetes*. This leads to think of a *Bacteroidetes* origin for this GH42-like family, but no clear subsequent evolutionary story can be inferred from clade A, due to the mixing of different bacterial phyla.

Interestingly, when we look at the genes assignment of these 76 homologous enzymes, only the one of *P. carrageenovora* is assigned as “distantly related to GH42”, whereas 56 % are annotated as hypothetical proteins, 36 % as GH42, 5 % as beta-galactosidases, and 1 as T9SS_C-terminal target domain containing protein. This is an illustrative example of the lack of knowledge about marine enzymes leading to mis-annotation of genes, but also to the fact that novel GH families probably specific of the marine ecosystems are overlooked (Barbeyron *et al.* 2000; Rebuffet *et al.* 2011). Indeed, all bacteria synthesizing this type of GH42-like enzymes almost exclusively live in seawater, with the exception of *Bacteroides ovatus*, *B. thetaiotamicron* and *Blautia sp.* Marseille-P3201 which are human gut bacteria, and *Paenibacillus spp.* species which are terrestrial bacteria widely spread in all environments. Concerning the genus *Paenibacillus*, Xu *et al.* (Xu *et al.* 2017) have recently reported the

presence of genes encoding for agar degradation in different terrestrial species of this genus. Xu *et al.* suggest that the presence of these marine genes could reveal an evolutionary strategy of genetic plasticity of the *Paenibacillus* genus allowing the colonization of a wide range of ecological niches. But no clear evolutionary history has been proposed to explain the presence of these genes in soil bacteria, isolated at a distance of more than 1300 km from the sea. On the other hand, the presence of genes encoding enzymes capable to degrade algal polysaccharides has already been observed in the microbiome of human populations that traditionally consuming high quantity of algae, such as Japanese. This case of horizontal gene transfer (HGT) between marine bacteria degrading algae to human gut bacteria has been demonstrated several times for enzymes degrading porphyran, agar and alginate (Hehemann *et al.* 2010, 2012b; Thomas *et al.* 2012). Interestingly, the first case of HGT of genes related to carrageenan degradation in human gut bacteria was reported by Matard-Mann *et al.* (*JBC* in press). In this study the phylogenetic analysis of κ -carrageenases also revealed the presence of κ -carrageenase encoding genes in *Bacteroides ovatus* and *B. thetaiotamicron*. In the here presented case of the new GH family, it is not yet clear whether this is a monospecific family. Consequently, the presence of homologous sequences in terrestrial bacteria doesn't mean they are active on marine substrate.

However, the comparative alignment of the 76 sequences of GH42-like family to true GH42 sequences reveals that only 9 residues are strictly conserved, including the two catalytic glutamic acid identified in *B. subtilis* YesZ E145/E295 (Shaikh *et al.* 2007), equivalent with E310/E495 and E309/E494 in *Pc* and *Ph* respectively. When comparing separately at the conserved residues of GH42-like and true GH42, we can also notice differences in the chemical nature of these conserved residues. Notably, family of GH42-like shows 35 % of charged residues, instead of 24 % for true GH42 family. This element is in good agreement with the fact that many marine polysaccharides are sulfated. Moreover, even if strictly conserved nonpolar residues of GH42-like family represent only 37 % instead of 48 % for true GH42, 71 % of them are aromatic residues compared with 59 % for true GH42. This element is also in good agreement with the presence of 3,6-anhydro-bonds in marine polysaccharide, that creates bicyclic saccharides with an increased hydrophobicity.

Altogether with the strong abundance of marine bacterial species in the family GH42-like and the already known case of HGT for the few terrestrial strains present in the tree, this allows the hypothesis that this new GH42-like family is specific of marine polysaccharidic substrates, of mixed chemical nature (charged and hydrophobic), and that the presence of related genes in some terrestrial bacteria might well represent a new case of HGT from marine environment, or

possibly a reminiscence of an ancestral adaptation to environmental niches connected with the marine ecosystems.

Biochemical activity reveals these enzymes are neo- β -carrabiose hydrolase – The analysis of the aminoacid sequence derived from genes encoding for the GH42-like enzymes of *Paraglaciecola hydrolytica* S66^T (Ph1657) and *P. carrageenovora* 9^T (Pcmp0019) revealed the presence of a signal peptide of 19 and 20 residues respectively, with the presence of a cysteine at +1 and a serine at +2 relative to the cleavage site of the signal peptidase SpII, characteristic of periplasmic lipoproteins of Gram-negative bacteria anchored in the outer-membrane (Seydel *et al.* 1999). Both enzymes have been successfully cloned and expressed as soluble recombinant proteins, without their signal peptide sequence, resulting in recombinant proteins of respectively 91.8 kDa and 91.5 kDa, with a pI of 5.41 and 5.52 respectively. Trials of crystallogenesis have been done but crystals were very difficult to obtain and diffracted with a poor resolution. The presence of a predicted N-terminal region of coiled-coil in both proteins could explain crystallogenesis troubles. Indeed, coiled-coil domains are known to mediate protein/protein interaction (Burkhard *et al.* 2001), which could complicate the crystalline assemblage of these proteins.

Preliminary assays showed that Pcmp0019 is active on the degradation products of furcellaran incubated with the κ -carrageenase from *P. carrageenovora* (Figure 4A). Subsequent activity tests on commercial DP6 having the sequence $\beta\kappa\kappa$, revealed that the products correspond to a DP2 β -carrabiose and DP4 κ -carratetraose. However, no activity was detected on DP6 $\kappa\kappa\kappa$ or DP4 $\kappa\kappa$ (Figure 4B). So the absence of a sulfate group on the G unit of the non-reducing end is compulsory for this enzyme to be able to cleave the subsequent β (1 \rightarrow 4) linkage. Moreover, to determine the biochemical conditions required for the activity of this enzyme, we used the recently identified 3,6-anhydro-D-galactosidases of *Zobellia galactanivorans* from GH127 and GH129 families (Ficko-Blean *et al.*, 2017, in press). These enzymes specifically remove the DA unit of the non-reducing end of β -carrageenan. Once they have cleaved the α (1 \rightarrow 3) linkage, Pcmp0019 is not able anymore to cleave the following β (1 \rightarrow 4) linkage, meaning that this enzyme requires the presence of a disaccharide in the non-reducing extremity to be able to cleave its substrate.

Similarly to Pcmp0019, in presence of a hybrid DP4 having the sequence $\beta\kappa$ -carratetraose, Ph1657 completely degrades the substrate to produce a DP2 β -carrabiose and a DP2 κ -carrabiose (Figure 5A). However, and in contrast to Pcmp0019, in presence of κ -oligosaccharides produced by the degradation of pure κ -carrageenan with the κ -carrageenase

of *P. carrageenovora*, this enzyme generates mainly DP4 and DP2 κ -oligosaccharides (Figure 5B). In addition, when added to DP4 κ -carratetraose, DP2 κ -carrabiose oligosaccharides are liberated, even if some DP4 κ remains not degraded. Thus, both enzymes belonging to Clade A of GH42-like seems to remove specifically a neo-betacarrabiose unit from hybrid beta/kappa oligosaccharides, with a less stringent specificity of Ph1657 that also seems able to cleave neo-kappacarrabiose.

Interestingly, thin layer chromatography following the reaction products with time showed that transglycosylation activity can be observed during the first four hours of incubation of Ph1657 with DP4 β κ -carratetraose oligosaccharides as sole substrate. Indeed, two new products of higher DP are synthesized, along with the cleavage of DP4 β κ in DP2 β -carrabiose and DP2 κ -carrabiose (Figure 6). They migrate at a height intermediate between DP4 and DP6 pure κ -carrabiose oligosaccharides, indicating that they might be of hybrid nature, most probably DP6 β κ -carrhexaose, along with some pure DP6 κ -carrhexaose. Mass spectrometry analysis will be necessary to identify more precisely the exact nature of these reaction products. However, after 24 h of incubation with the enzyme, these transglycosylation products are further hydrolyzed and completely disappear from the chromatogram by completion of the reaction course. This type of transglycosylation side-reaction is often observed for glycoside hydrolases proceeding by a catalytic mechanism with retention of the anomeric carbon at C1 (Bissaro et al. 2015). Contrary to inverting glycoside hydrolases, the transglycosylation product that is temporarily formed during the reaction is then hydrolysed (Cobucci-Ponzano et al. 2009). Consequently, we can postulate that these novel glycoside hydrolases are retaining enzymes. Seminal and more recent studies reported on the fine-tuning of the balance between both hydrolytic and transglycosylation activities, shifting the equilibrium towards transglycosylation by targeted mutagenesis of key residues involved in the catalytic mechanism (Mackenzie *et al.* 1998; Moracci *et al.* 1998), but also in the dynamic of water within the enzyme (David et al. 2017). These approaches might open the possibility to transform these GH42-like enzymes into transglycosidases, that could synthesize original oligosaccharides of pure β -carrageenan. This naturally desulphated carrageenan obtained without use of chemical (generally concentrated NaOH solution) could be of fundamental interest to study the impact of sulfatation on biological activities of carrageenan (Yuan et al. 2005 & 2011), providing a natural negative control.

CONCLUSION

Marine ecosystems are unique source of original polysaccharides absent from terrestrial environment. Notably, some heterogenous polysaccharides, like furcellaran, combine a charged

nature, due to the presence of sulfate groupments in κ motives, with a hydrophobic nature, due to the presence of 3,6-anhydrogalactoses in both κ and β motives. Marine microorganisms have thus developed very specific enzymes dedicated to these marine substrates which constitute new families, specific of marine environment, like the GH82 family specific for ι -carrageenan degradation. Here we report the discovery of a new family of GH distantly related to GH42 family. Cloning and biochemical study of two of these enzymes along with a phylogenetical study of 76 members of this new family have allowed to draw the first characteristics of this family. It seems to be characteristic of marine bacteria, with possibly some HGT to terrestrial strains, and subdivided in at least two clades, a subgroup of clade A being probably specific of β -carrageenan. Both enzymes studied show the ability to cleave the β -1,4 glycosidic linkage to remove a β -carrabiose moiety from hybrid oligosaccharides of β/κ carrageenan. However some differences of specificity appeared between them, with Ph1657 showing a partial tolerance to pure κ -oligosaccharides. The presence of a homologous of these enzymes in *Zobellia galactanivorans*, showing respectively 37 % and 38 % of identity with the enzymes of *Pc* and *Ph*, and belonging to clade B, gives promising prospects in order to determine if this new family of GH42-like can be polyspecific or restricted to partially desulfated carrageenan. Cloning of this enzyme is currently in progress.

MATERIEL AND METHODS

Phylogenetical analysis – The sequence of the GH42-like gene of *P. carrageenovora* 9^T (mp0019) was used to extracted homologous sequences with the blastp suite from NCBI server (NCBI Resource Coordinators 2017) (<https://blast.ncbi.nlm.nih.gov>) from the non-redundant protein sequences database. After manually curation, 76 sequences were aligned with MAFFT program available at: <http://www.ebi.ac.uk/Tools/msa/mafft/> and edited with BioEdit free software (Hall *et al.* 2011) available at: <http://www.mbio.ncsu.edu/BioEdit/bioedit.html> in order to eliminate all the insertion areas before calculation of the phylogenetic tree with MEGA software (<http://www.megasoftware.net/>). The sequences of three ‘true’ GH42, namely the exo- β -1,4-galactanase from *Dickeya dadantii dadantii* 3937 (Glasner *et al.* 2011), the β -1,6-galactosidase from *Bifidobacterium bifidum* S17 (Godoy *et al.* 2016) and the beta-galactosidase from *Bacillus subtilis* subsp. *subtilis* str. 168 (Shaikh *et al.* 2007), were used to form an outgroup. Bootstrap values inferior to 50 % do not appear on the tree.

Cloning, expression and purification of the two beta-neocarrabiohydrolases – The gene of *Paraglaciecola* sp. S66 encoding for the GH42-like Ph1657 (WP 068375692.1) and consisting of 2367 nucleotides after the cysteine in position (+1) after the peptide signal sequence, was

amplified from genomic DNA by PCR with forward primer 5'-GGCTTAAUTCTGAGCAAAGTAAACAAG-3' and reverse primer 5'-GGTTTAAUAAGCGCGTTGACCTC-3' and cloned into pET9a.His.USER vector fused to 8x His-tag in C-terminal extremity. The gene of *P. carrageenovora* encoding for the GH42-like Pcmp0019 and consisting of also 2367 nucleotides after Q22 (position +2 after the cleavage site of peptide signal) was amplified from genomic DNA by PCR with forward primer 5'-AAAAAAGGATCCCAAGAAAGTCCCATAGAGTTAAAAGA-3' and reverse primer 3'-TTTTTCTCGAGTTATTGAGCTGTAACCTTCAACAAGTAA-5' and cloned into pFO4 vector fused to 6x His-tag in N-terminal extremity. In both sequences, the signal peptide sequence encoded by 63 and 66 bp respectively has been excluded from cloning. *E. coli* strain BL21 (DE3) cells were transformed and grown in the auto-induced ZYP-5052 rich medium (Studier 2005) during 72 h at 20°C with the adapted antibiotic (Kanamycin or Ampicillin 50 μ g/ml respectively). After centrifugation, cell pellets were resuspended in A buffer (20 mM Phosphate buffer pH 7.4, 500 mM NaCl, 40 mM imidazole for Ph1657 and 50 mM Tris buffer pH 7.5, 300 mM NaCl, 50 mM imidazole for Pcmp0019) and cells were lysed by mean of a French-press in presence of an anti-protease mixture (Complete EDTA-free; Roche Applied Science), DNase (0.1 mg/ml) and lysozyme (0.2 mg/ml). After centrifugation, the supernatant was filtered on a 0.2 μ m membrane and subsequently purified by affinity chromatography on a HisTrap HP 1ml column (GE Healthcare®) with a gradient a B buffer differing from A buffer only in containing 500 mM imidazole. A final step of purification was done by a size exclusion chromatography on a Hiloal Superdex 200 (GE Healthcare®) equilibrated with 50 mM Tris buffer pH 8, 25 mM NaCl or 50 mM Tris buffer pH 7.5, 150 mM NaCl for Ph1657 and Pcmp0019 respectively. Proteins were finally concentrated to 8 mg/ml and 9.7 mg/ml respectively.

Enzymatic degradation of oligo-carrageenan with fluorophore assisted carbohydrate electrophoresis (FACE) analysis – Hybrid oligosaccharides of beta/kappa-carrageenotetraose (DP4) have been obtained from Dextra. Beta/kappa/kappa-carrageenohexaoses (DP6) were purified from digestion products of Furcellaran polymer by the kappa-carrageenase of *P. carrageenovora* on size exclusion chromatography, and structurally characterized by NMR analysis. 25 μ l of hybrid oligosaccharides at 2 % (w/v) were incubated with 2 μ l of enzyme concentrated at 8 and 1 mg/ml for Ph1657 and Pcmp0019 respectively, during 24 h at 20 °C. Reaction products were then labeled with a fluorochrome, according to the protocols described by Jackson *et al.*, (Jackson 1990, 1991). Briefly, aliquots of 25 μ l were dried in a speed-vacuum centrifuge, and the oligosaccharides were labeled with 8-aminonaphthalene-1,3,6-trisulfonate

(ANTS) or 2-aminoacridone (AMAC) depending on the cases. For derivatization with ANTS, the pellet was dissolved in 2 μ l of ANTS solution (0.15 M ANTS in acetic acid/water (3:17, v/v)). Once the pellet was dissolved, 5 μ l of 1 M sodium cyanoborohydride in dimethyl sulfoxide (DMSO) was added, and the mixture was incubated for 16 h at 37 °C. For the derivatization with AMAC, the pellet was dissolved in 2 μ l of AMAC solution (0.1 M AMAC in acetic acid/DMSO (3:17, v/v)) and 5 μ l of 1 M sodium cyanoborohydride in water also followed by 16 h of incubation at 37 °C. The samples were analyzed on a 30 % polyacrylamide running gel with a 4 % stacking gel and a Bio-Rad gel system at 4 °C. Gels are unveiled under a UV-lamp.

HPAEC analysis – Analyses of degradation products generated by both enzymes were conducted by HPAEC on an ICS-5000 system (*Dionex*). 20 μ l of samples were injected (AS50 Automatic Autosampler, *Dionex*) onto an AS11 analytical anion-exchange column (4 \times 250 mm, IonPac®, *Dionex*) equipped with an AG11 pre-column (4 \times 50 mm, IonPac®, *Dionex*). Elution were performed at a flow rate of 0.5 ml/min (Gradient Pump SP-5, *Dionex*) by NaOH multisteps gradients from 8 to 280 mM (40 min). Oligosaccharides were detected in conductivity mode (Analytical DC-5000 detector, *Dionex*) with a suppressor (AERS 500, 4 mm, *Dionex*) delivering a current from 50 to 300 mA. Acquisition of chromatograms was realized by the "Chromeleon" Software version 6.8 (*Dionex*).

FIGURES LEGENDS

Figure 1: Chemical structures of the four main carrageenans: beta (β), kappa (κ), iota (ι), lambda (λ). The neocarrabiose unit of β -, κ - and ι -carrageenan is made of a 3,6-anhydro-D-galactose(-2-sulfate) unit linked in β -1,4 with an α -D-galactose(-4-sulfate) unit, and λ neocarrabiose is made of D-galactose-2,6-disulfate unit linked in β -1,4 with an α -D-galactose-2-sulfate unit.

Figure 2: Genomic environment of GH42-like genes in *Paraglaciecola hydrolytica* S66^T from Ph1646 to Ph1677 (A) and *Pseudoalteromonas carrageenovora* 9^T from Pcmp0019 to Pcmp0053 (B).

Figure 3: Phylogenetic tree of 76 sequences of ‘GH42-like’ proteins. Clades A (circled and shaded in blue) and B (in orange) are clearly separated from the out group made of three sequences of true GH42 enzymes. Subgroup A1 in clade A contains the two enzymes studied in this work, and enlighten with a star. Their homologous in *Zobellia galactanivorans* is also marked with a star in clade B. Terrestrial bacteria strains are circled with dashes.

Figure 4: FACE gel using ANTS labelling of digestion products of *P. carrageenovora* GH42-like Pcmp0019. The reaction with GH42-like is annotated with +E legend; DP stands for degree of polymerisation. A: digestion of hybrid carrageenan made of β/κ motives: Lanes 1 and 2 show the oligosaccharidic profile of hydrolysed furcellaran after digestion by κ -

carrageenase alone (lane 1) or also with GH42-like (lane 2); Lane 4 shows a DP6 $\beta\kappa\kappa$, and its degradation products by GH42-like are on lane 5. Lane 3 is empty. B: trials on pure κ -carrageenan oligosaccharides: Lane 1 and 2: Profile of a mix of DP4 and DP6 without (lane 1) and with GH42-like (Lane 2) Lane 3 and 4: Profile of a mix of DP2 and DP4 without (lane 3) and with GH42-like (Lane 4).

Figure 5: FACE gel using AMAC labelling of digestion products of *P. hydrolytica* GH42-like Ph1657. A: digestion of different oligo-carrageenases; lanes are numbered from 1 to 9 and the reaction with enzyme is annotated with +E legend; B: trials on pure κ -carrageenan oligosaccharides. DP stands for degree of polymerisation.

Figure 6: TLC analysis of DP4 $\beta\kappa$ oligosaccharide digested by Ph1657 throught the time. Lanes 1, 10, 11, 12: standards oligosaccharids of DP4 $\beta\kappa$, DP2 κ , DP4 $\kappa\kappa$, DP6 $\kappa\kappa\kappa$; Lanes 2 to 9: 0 min, 5 min, 10 min, 30 min, 1 h, 4 h, 26 h, 48 h

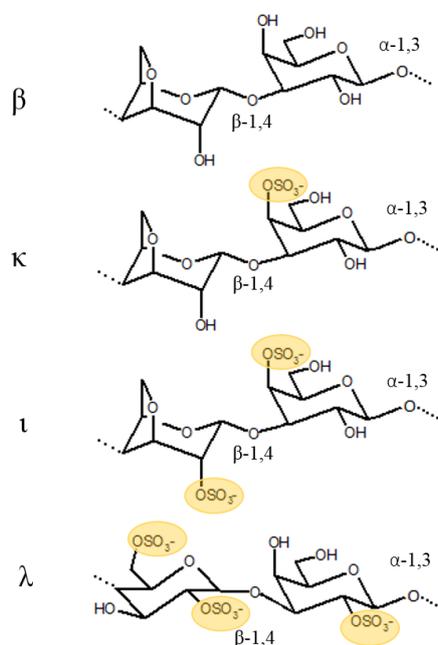
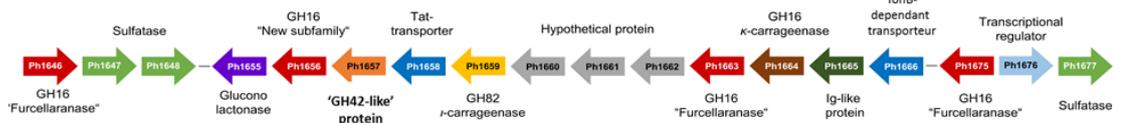


Figure 1

A



B

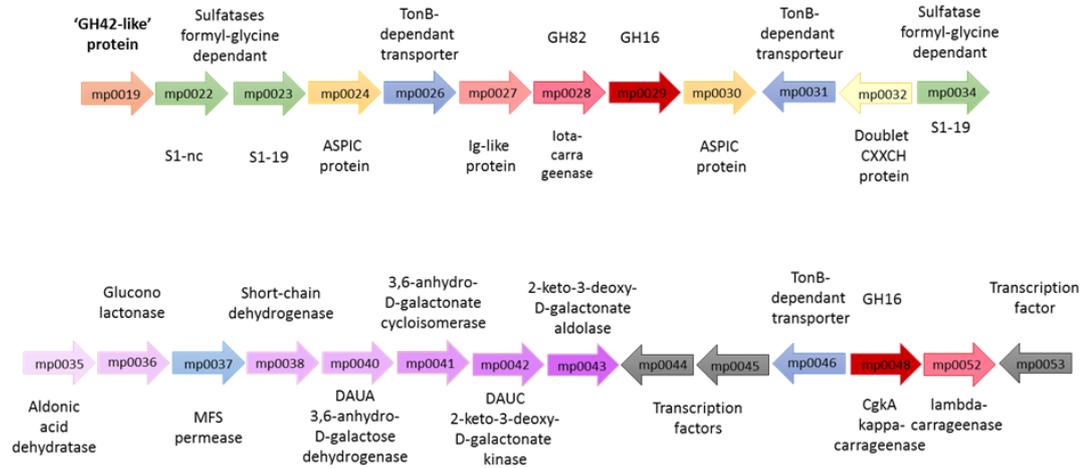


Figure 2

Chapitre 3 : Etude comparée de deux β -carrabiose hydrolases

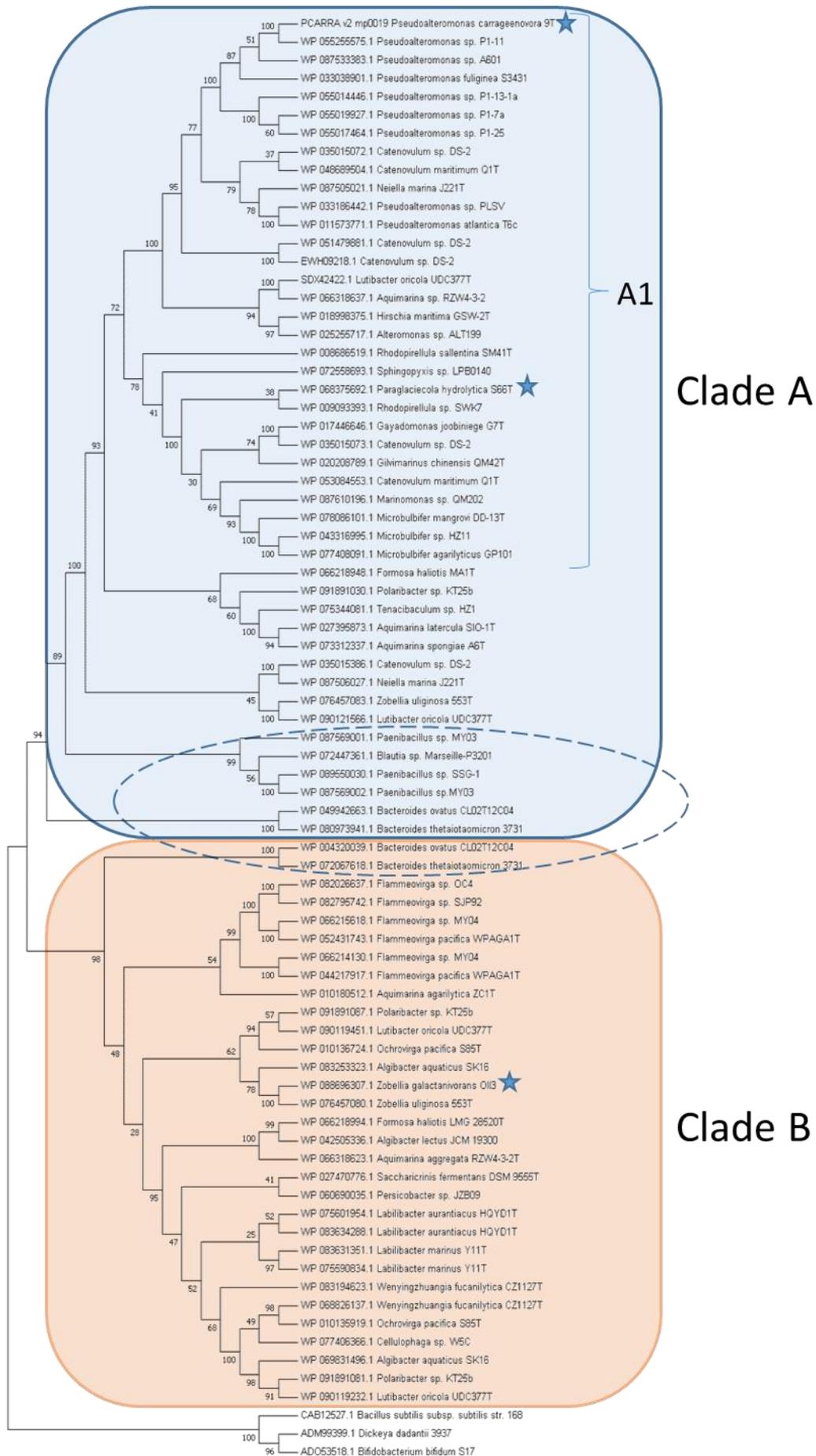


Figure 3
168

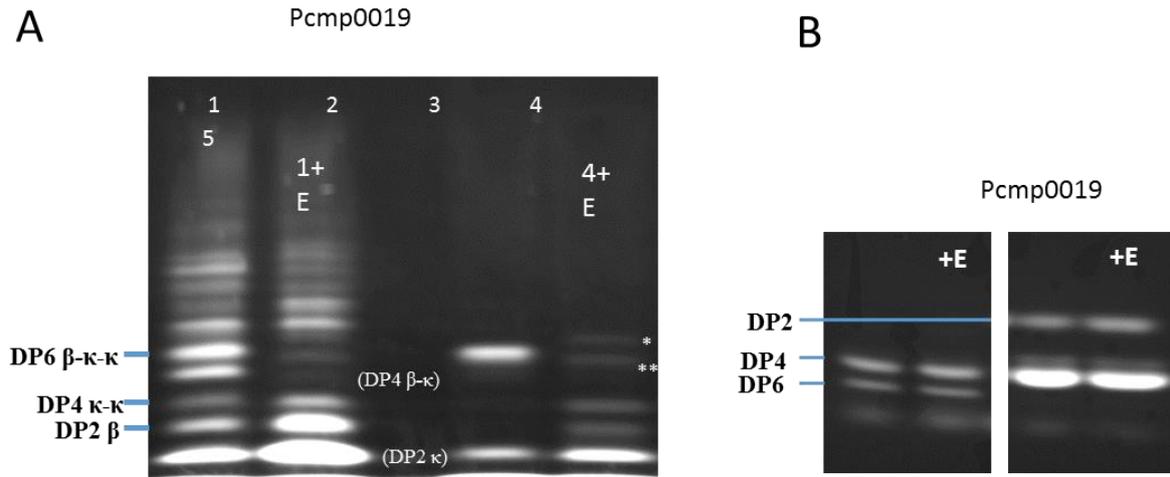


Figure 4

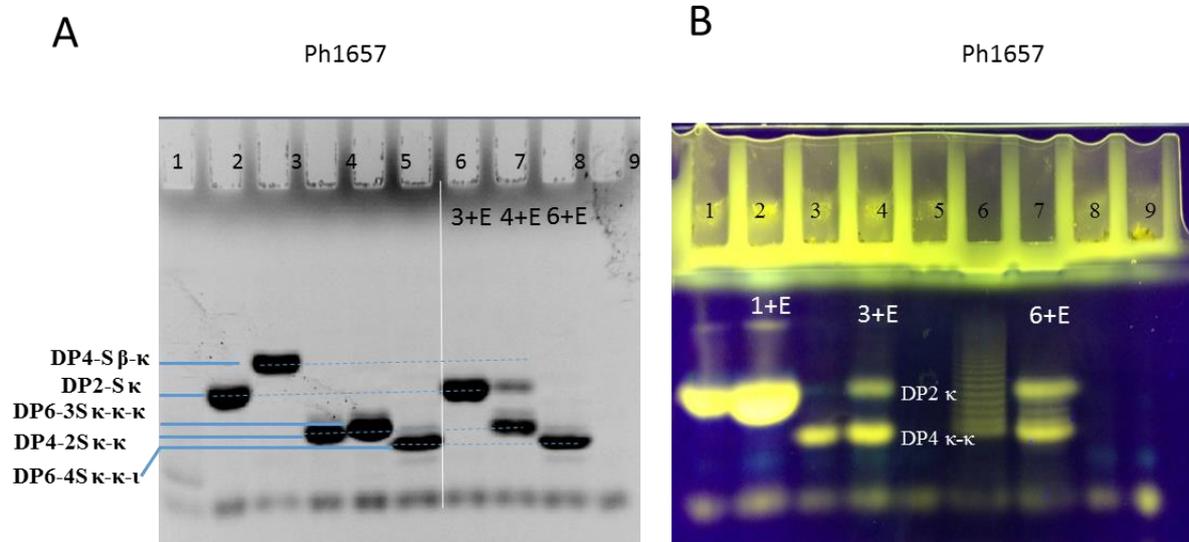


Figure 5

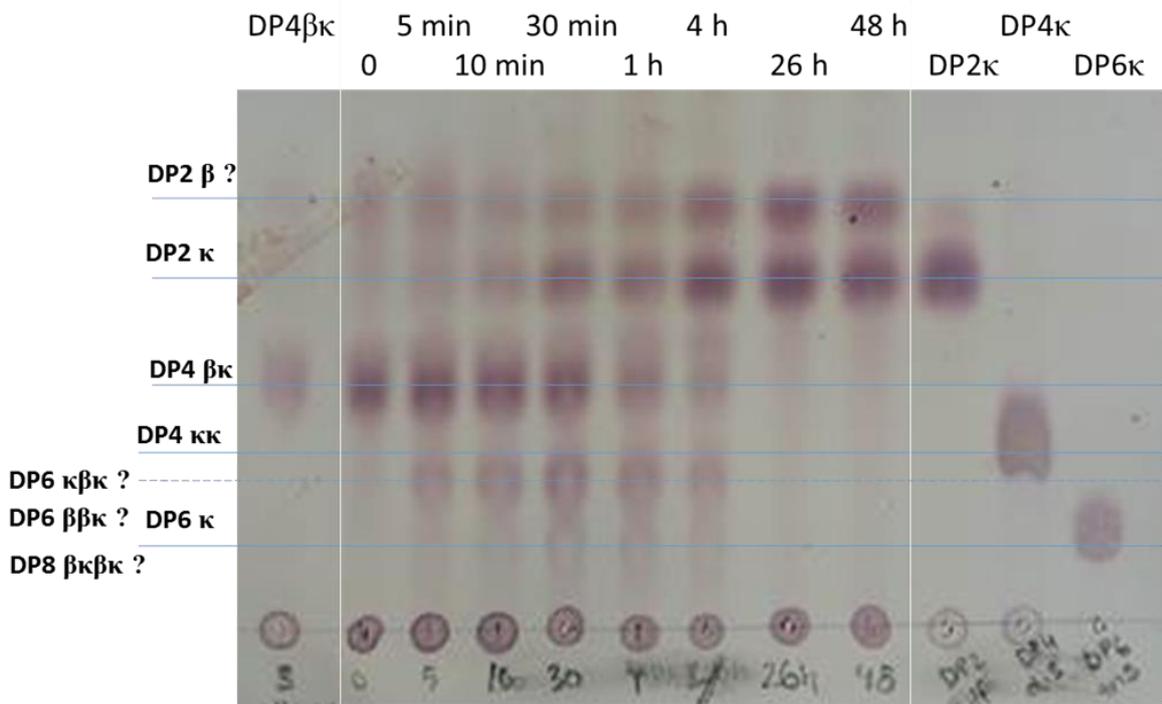


Figure 6

REFERENCES

1. Azam, F., and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* **5**, 782–91
2. Arnosti, C. (2014) Patterns of Microbially Driven Carbon Cycling in the Ocean: Links between Extracellular Enzymes and Microbial Communities. *Adv. Oceanogr.* **2014**, 1–12
3. Michel, G., and Czjzek, M. (2013) Polysaccharide-degrading enzymes from marine bacteria. in *Marine enzymes for biocatalysis: sources, biocatalytic characteristic and bioprocesses of marine enzymes*, pp. 429–464
4. Martin, M., Portetelle, D., Michel, G., and Vandenberg, M. (2014) Microorganisms living on macroalgae: diversity, interactions, and biotechnological applications. *Appl. Microbiol. Biotechnol.* **98**, 2917–35
5. Glicksman, M. (1987) Utilization of seaweed hydrocolloids in the food industry. *Hydrobiologia.* **151–152**, 31–47
6. Bemiller, J. N. (2009) One hundred years of commercial food carbohydrates in the United States. *J. Agric. Food Chem.* **57**, 8125–8129
7. Bixler, H. J., and Porse, H. (2010) A decade of change in the seaweed hydrocolloids industry. *J. Appl. Phycol.* **23**, 321–335
8. Craigie, J. S. (1990) Cell Walls. in *Biology of the red algae, New York*, pp. 221–257, Cambridge University Press
9. Grasdalen, H., and Knutser, S. H. (1987) Characterization of Water-extractable Polysaccharides from Norwegian *Furcellaria lumbricalis* (Huds.) Lamour. (Gigartinales, Rhodophyceae) by IR and NMR Spectroscopy. *Bot. Mar.* **30**, 497–506
10. Correc, G., Barabanova, A., Tuvikene, R., Truus, K., Yermak, I., and Helbert, W. (2012) Comparison of the structures of hybrid κ - β -carrageenans extracted from *Furcellaria lumbricalis* and *Tichocarpus crinitus*. *Carbohydr. Polym.* **88**, 31–36
11. McLean, M. W., and Williamson, F. B. (1981) Neocarratetraose 4-O-monosulfate beta-hydrolase from *Pseudomonas carrageenovora*. *Eur. J. Biochem.* **113**, 443–456
12. Schultz-johansen, M., Glaring, M. A., Bech, P. K., and Stougaard, P. (2016) Draft Genome Sequence of a Novel Marine Bacterium, *Paraglaciecola* sp. Strain S66, with Hydrolytic Activity against Seaweed. **4**, 4–5
13. Barbeyron, T., Michel, G., Potin, P., Henrissat, B., and Kloareg, B. (2000) ι -Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of κ -carrageenases. *J. Biol. Chem.* **275**, 35499–35505
14. 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
15. Xu, H., Qin, S., Lan, Y., Liu, M., Cao, X., Qiao, D., Cao, Y., and Cao, Y. (2017) Comparative genomic analysis of *Paenibacillus* sp. SSG-1 and its closely related strains reveals the effect of glycometabolism on environmental adaptation. *Sci. Rep.* **7**, 5720
16. 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**, 908–12

17. Hehemann, J.-H. J.-H., Kelly, A. G., Pudlo, N. A., Martens, E. C., and Boraston, A. B. (2012) Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19786–91
18. Thomas, F., Barbeyron, T., Tonon, T., Génicot, S., Czjzek, M., and Michel, G. (2012) Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine *Flavobacteriia* to their independent transfers to marine *Proteobacteria* and human gut *Bacteroides*. *Environ. Microbiol.* **14**, 2379–94
24. Shaikh, F. A., Müllegger, J., He, S., and Withers, S. G. (2007) Identification of the catalytic nucleophile in Family 42 β -galactosidases by intermediate trapping and peptide mapping: YesZ from *Bacillus subtilis*. *FEBS Lett.* **581**, 2441–2446
19. Seydel, A., Gounon, P., and Pugsley, A. P. (1999) Testing the “+2 rule” for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Mol. Microbiol.* **34**, 810–821
20. Burkhard P, Stetefeld J, Strelkov SV. 2001. Coiled coils: a highly versatile protein folding motif. *Trends in cell biology*, **11**: 82-88
21. Bissaro B, Monsan P, Fauré R, O'Donohue MJ. 2015. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochem J* **467**: 17–35.
22. Cobucci-Ponzano B, Conte F, Bedini E, Corsaro MM, Parrilli M, Sulzenbacher G, Lipski A, Dal Piaz F, Lepore L, Rossi M, et al. 2009. β -Glycosyl Azides as Substrates for α -Glycosynthases: Preparation of Efficient α -L-Fucosynthases. *Chem Biol* **16**: 1097–1108
23. Mackenzie LF, Wang Q, Warren RAJ, Withers SG. 1998. Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J Am Chem Soc* **120**: 5583–5584.
24. Moracci M, Trincone A, Perugino G, Ciaramella M, Rossi M. 1998. Restoration of the activity of active-site mutants of the hyperthermophilic β -glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles. *Biochemistry* **37**: 17262-17270.
25. David B, Irague R, Jouanneau D, Daligault F, Czjzek M, Sanejouand YH, Tellier C. 2017. Internal Water Dynamics Control the Transglycosylation/Hydrolysis Balance in the Agarase (AgaD) of *Zobellia galactanivorans*. *ACS Catal* **7**: 3357–3367.
26. Yuan H, Zhang W, Li X, Lü X, Li N, Gao X, Song J. 2005. Preparation and in vitro antioxidant activity of κ -carrageenan oligosaccharides and their oversulfated, acetylated, and phosphorylated derivatives. *Carbohydrate Research* **340**: 685-692.
27. Yuan H, Song J, Li X, Li N, Liu S. 2011. Enhanced immunostimulatory and antitumor activity of different derivatives of κ -carrageenan oligosaccharides from *Kappaphycus striatum*. *Journal of applied phycology* **23**: 59-65.
28. NCBI Resource Coordinators (2017) Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **45**, D12–D17
29. Hall, T., Biosciences, I., and Carlsbad, C. (2011) BioEdit: An important software for molecular biology. *GERF Bull. Biosci.* **2**, 60–61
30. Glasner, J. D., Yang, C. H., Reverchon, S., Hugouvieux-Cotte-Pattat, N., Condemine, G., Bohin, J. P., van Gijsegem, F., Yang, S., Franza, T., Expert, D., Plunkett, G., Francisco, M. J. S., Charkowski, A. O., Py, B., Bell, K., Rauscher, L., Rodriguez-

- Palenzuela, P., Toussaint, A., Holeva, M. C., He, S. Y., Douet, V., Boccara, M., Blanco, C., Toth, I., Anderson, B. D., Biehl, B. S., Mau, B., Flynn, S. M., Barras, F., Lindeberg, M., Birch, P. R. J., Tsuyumu, S., Shi, X., Hibbing, M., Yap, M. N., Carpentier, M., Dassa, E., Umehara, M., Kim, J. F., Rusch, M., Soni, P., Mayhew, G. F., Fouts, D. E., Gill, S. R., Blattner, F. R., Keen, N. T., and Perna, N. T. (2011) Genome sequence of the plant-pathogenic bacterium *Dickeya dadantii* 3937. *J. Bacteriol.* **193**, 2076–2077
31. Godoy, A. S., Camilo, C. M., Kadowaki, M. A., Muniz, H. dos S., Espirito Santo, M., Murakami, M. T., Nascimento, A. S., and Polikarpov, I. (2016) Crystal structure of β 1→6-galactosidase from *Bifidobacterium bifidum* S17: trimeric architecture, molecular determinants of the enzymatic activity and its inhibition by α -galactose. *FEBS J.* **283**, 4097–4112
 32. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–234
 33. Jackson, P. (1990) The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-trisulphonic acid. *Biochem. J.* **270**, 705–713
 34. Jackson, P. (1991) Polyacrylamide gel electrophoresis of reducing saccharides labeled with the fluorophore 2-aminocridone: subpicomolar detection using an imaging system based on a cooled charge-coupled device. *Anal. Biochem.* **196**, 238–244

Chapitre 4

Caractérisation d'une α -L-fucosidase de la famille GH29 au sein d'un PUL putatif de *Zobellia* *galactanivorans*

Chapitre 4 : Caractérisation d'une α -L-fucosidase de la famille GH29 au sein d'un PUL putatif de *Zobellia galactanivorans*

I. Introduction

Dans l'optique de découvrir de nouvelles enzymes de bactéries marines aux fonctions originales, une des stratégies possibles est d'étudier des GH encore non attribuées à une famille de CAZymes. Annotées comme GHnc (pour non classée ou *not classified*) dans les génomes, leurs séquences sont trop divergentes pour pouvoir être attribuées à l'une des 145 familles actuellement répertoriées dans la base de données CAZy.

Plusieurs éléments peuvent faciliter l'étude de telles enzymes : leur présence au sein de structures de type PUL, aux côtés d'autres enzymes de fonction prédite par homologie de séquence ; la disponibilité d'informations transcriptionnelles sur l'expression de certains gènes du PUL en présence de substrats définis ; l'obtention par clonage des différentes enzymes du PUL et leur expression sous forme soluble. Cependant, comme nous allons le voir au cours de ce chapitre, même si ces différentes conditions sont réunies, la recherche du substrat naturel d'une enzyme, encore plus quand on ignore tout de sa fonction, est une entreprise délicate.

En effet, face à la multitude de substrats naturels disponibles, ainsi qu'au manque de caractérisation chimique précise de ces substrats, les « banques de substrats » pouvant être constituées dans le temps imparti d'une thèse sont loin d'être exhaustives. De plus, comme évoqué lors de l'introduction, le nombre de CAZymes caractérisées et disponibles sous forme recombinante est très limité, en particulier dans le cas des substrats de type polymères de fucose : une seule fucanase, la GH107 MfFcnA de *Marineflexile fucanivorans* SWS, et une seule fucosidase, la GH29 TmaFuc de *Thermotoga maritima* MSB8 (cf. tableau 3 p. 60) ; cela offre peu d'outils de modification des substrats pour aider à identifier les fonctions de nouvelles enzymes. C'est donc en ayant conscience de ces différentes contraintes que j'ai commencé l'étude de deux GH d'un PUL putatif de la bactérie marine *Z. galactanivorans*, une GH29 et une GHnc.

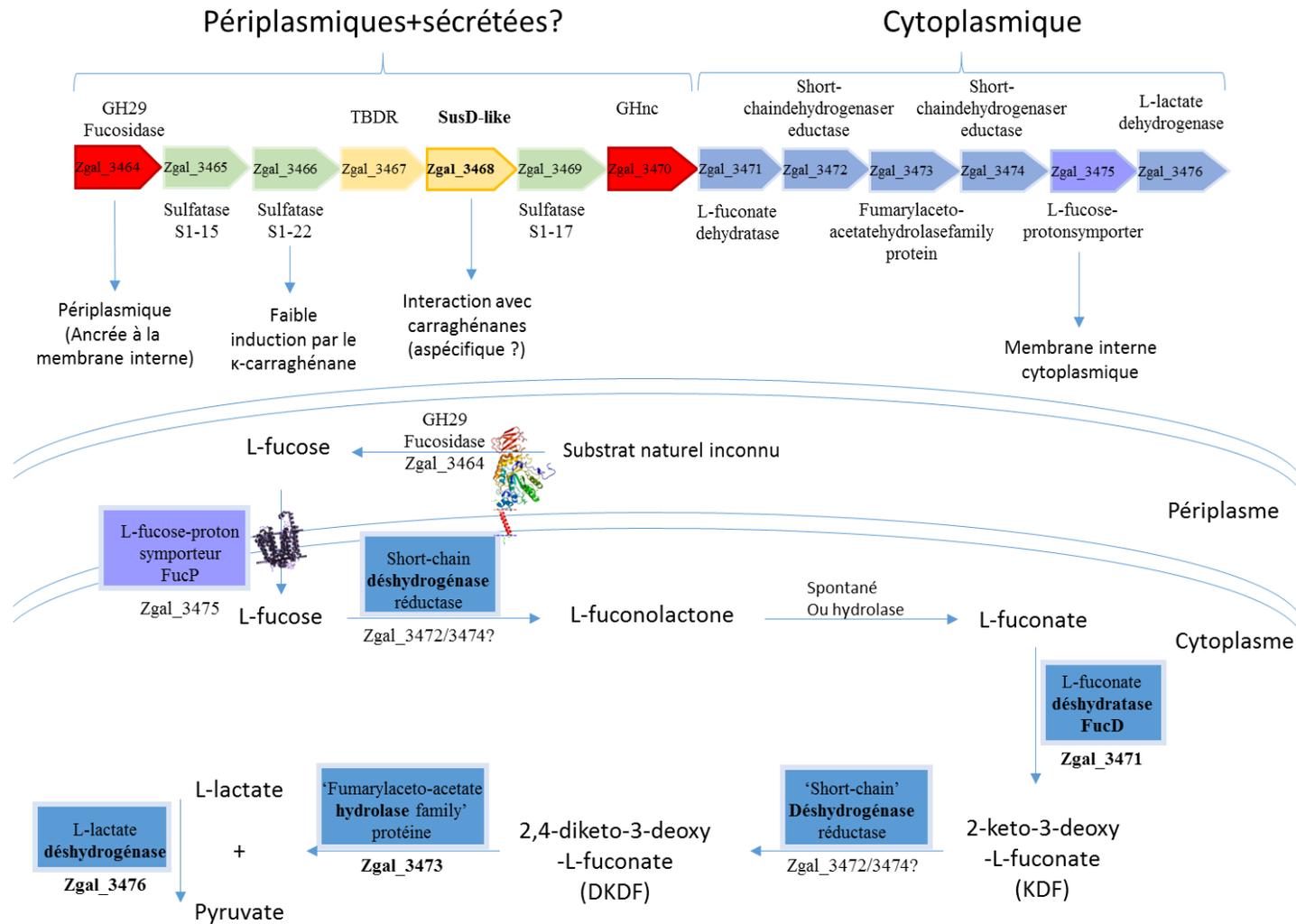


Figure 39 : Organisation du PUL putatif contenant la GHnc et la GH29, et localisation cellulaire des différentes protéines du PUL. La voie de dégradation cytoplasmique du fucose proposée pour les gènes Zgal_3471 à Zgal_3476 a été construite sur le modèle de la voie proposée par Yew *et al.* 2006.

II. Données et hypothèses de départ

1. Un PUL putatif fucanolytique

Parmi les cibles disponibles à l'issue de l'étape de clonage à moyen débit, l'une des plus prometteuses était une GHnc. Identifiée dans une structure de type PUL chez la Bacteroidetes marine *Z. galactanivorans*, son gène (Zgal_3470) se situe juste en aval d'un couple SusC/SusD-like (Figure 38). Le PUL putatif est borné en amont par un régulateur transcriptionnel de type AraC (Zgal_3463), et en aval par cinq gènes codants des protéines impliquées dans les étapes terminales du catabolisme du fucose : Zgal_3471 à Zgal_3476 (Figure 38). Ces cinq derniers gènes sont en effet prédits pour coder un proton-symporteur du fucose, des short-chain dehydrogenase reductases, une deshydratase, fumarylacetoacetate hydrolase et une deshydrogénase. Ces activités sont caractéristiques des dernières étapes du métabolisme du fucose (Yew *et al.* 2006).

Au sein de ce PUL putatif, le gène Zgal_3464 est prédit pour coder une α -L-fucosidase de la famille GH29 et trois autres gènes codent des sulfatases de la famille S1, appartenant aux sous-familles 15 (Zgal_3465, code une S1-15), 22 (Zgal_3466, code une S1-22) et 17 (Zgal_3469 code une S1-17).

Ces différents éléments amènent donc à penser que le constituant principal du substrat de ce PUL doit être le fucose, dont la sulfatation au sein des chaînes polysaccharidiques des macroalgues est fréquente (Deniaud-Bouët *et al.* 2017). Les substrats de la famille des fucoïdanes semblent les candidats les plus probables parmi les différentes familles de polysaccharides algaux.

Un autre élément en faveur de cette hypothèse vient de la comparaison de la structure de ce PUL avec celle d'autres bactéries marines. L'outil MAGE, mis en place par le Génoscope, permet d'effectuer une comparaison directe de la zone d'intérêt dans plusieurs génomes bactériens séquencés. Une représentation schématique des gènes de la bactérie cible chez les autres bactéries indique le degré de conservation et le caractère synténique de cette conservation. Grâce à cet outil, il est possible de constater que le cluster de gènes regroupant la GH29, deux des trois sulfatases (S1-15 et S1-17) ainsi que la GHnc, est conservé en synténie chez la bactérie marine *Mariniflexile fucanivorans*. Or, comme son nom l'indique, cette bactérie est capable de dégrader des fucanes, polymères de fucose sulfatés (Barbeyron *et al.* 2008).

D'autre part, les α -L-fucosidases (GH29) sont des enzymes de type exolytique. La dégradation d'un polymère de type fucoidane nécessite donc la présence d'une fucanase capable de cliver la chaîne de façon endolytique. Les fragments de taille intermédiaire ainsi générés pourront ensuite servir de substrat à d'autres enzymes de type exolytique.

Notre hypothèse de départ est donc que la GHnc pourrait réaliser la première étape de dégradation du substrat, avec une activité de type fucanase endolytique, accompagnée dans un second temps par la fucosidase GH29, avec une activité fucosidase exolytique.

2. Éléments contradictoires

Toutefois, quelques éléments semblent discordants avec cette théorie de « PUL fucanolytique ». Tout d'abord, la GHnc n'est absolument pas conservée chez les autres bactéries utilisées comme comparatifs, hormis chez *M. fucanivorans*, contrairement à la plupart des autres gènes du PUL (par exemple Zgal_3465 codant pour la S1-15 ou encore les gènes impliqués dans la dégradation terminale du fucose Zgal_3471 à 3476).

Une autre observation, issue de la comparaison avec les autres génomes bactériens, est que le couple SusC/SusD-like n'est pas strictement conservé en synténie avec les autres gènes du PUL (*Zobellia uliginosa* et *Z. russellii* par exemple). Néanmoins, cet argument n'est pas forcément contradictoire, puisque, comme évoqué au paragraphe III.2.2 de l'introduction, la plasticité génétique des bactéries est telle que de nombreux remaniements peuvent amener au déplacement d'un couple SusC/SusD à d'autres endroits du génome. Et, hormis la GHnc, la zone génétique englobant le PUL est particulièrement bien conservée en synténie chez *Z. amurskyensis* et *Z. laminariae*, deux bactéries capables de métaboliser le L-fucose (Nedashkovskaya *et al.* 2004).

Cependant, d'autres éléments, expérimentaux cette fois, posent aussi question.

Tout d'abord, la protéine de type SusD (codée par le gène Zgal_3468) impliquée dans la fixation des glucides, a été étudiée par Armando Salmeán dans le cadre d'une stratégie de criblage en double aveugle. Le but était d'identifier de nouvelles cibles protéiques spécifiques des polysaccharides, à l'aide de micropuces de type ComPP ('comprehensive microarray polymer profiling') (Moller *et al.* 2007). Ces travaux ont démontré que la protéine SusD-like de ce PUL putatif pouvait se fixer aux carraghénanes sulfatés (iota, kappa et lambda). Cependant, les résultats expérimentaux ne sont pas parvenus à déterminer une spécificité pour un type de carraghénane en particulier. De même, par comparaison avec les profils de fixation obtenus pour d'autres cibles, la fixation de la protéine SusD-like aux carraghénanes pourrait laisser

penser à une interaction relativement aspécifique, de type électrostatiques, liées à la présence des groupements sulfates chargés négativement (Salmeán *et al.* 2017).

Une autre étude, portant sur la culture de *Z. galactanivorans* en présence de différents substrats algaux (Thomas *et al.* 2017), mais n'incluant pas de fucoidanes, a également mis en évidence l'induction spécifique de l'expression d'une des sulfatases du PUL (S1-22 codée par Zgal_3466) par le κ -carraghénane (Tableau 6). Cependant, il faut noter que le niveau d'induction du gène de cette sulfatase, bien que significativement plus élevé qu'en présence de glucose, n'atteint pas des niveaux d'induction aussi élevés que d'autres gènes impliqués dans l'opéron alginolytique, par exemple, en présence d'alginate. De plus, cette sulfatase induite par le κ -carraghénane est aussi celle qui n'est pas conservée en synténie chez *M. fucanivorans*.

Ces différents éléments, bien qu'intrigants, ne semblent donc pas suffisamment probants face aux faisceaux d'éléments venant soutenir l'hypothèse d'un PUL fucanolytique.

Gene	Gene product	mean log ₂ -fold change compared to glucose					
		alginate	laminarin	agar	porphyran	kappa-carrag	iota-carrag.
ZGAL_3463	AraC-type transcriptional regulator	0,1	0,1	0,3	0,2	0,5	0,2
ZGAL_3464	Alpha-L-fucosidase, family GH29	1,1	0,6	0,5	1,3	1,1	0,8
ZGAL_3465	Sulfatase, family S1-15	0,7	0,9	0,2	1,4	1,2	0,6
ZGAL_3466	Sulfatase, family S1-22	0,8	0,4	0,2	1,2	2,0	0,8
ZGAL_3467	TonB-dependent Receptor	0,0	0,1	0,2	0,5	0,9	0,4
ZGAL_3468	SusD/RagB family lipoprotein	0,0	0,6	0,4	0,6	0,7	0,4
ZGAL_3469	Sulfatase, family S1-17	0,3	0,2	0,5	0,5	1,1	0,6
ZGAL_3470	Conserved hypothetical periplasmic protein	0,2	-0,2	0,1	1,0	1,0	0,7
ZGAL_3471	L-fuconate dehydratase	0,3	-0,3	-0,1	0,9	0,9	0,3
ZGAL_3472	Short-chain dehydrogenase/reductase	0,4	0,4	0,4	1,4	1,1	0,7
ZGAL_3473	Fumarylacetoacetate hydrolase family protein	0,1	0,1	0,2	-0,4	-1,1	-0,2
ZGAL_3474	Short-chain dehydrogenase/reductase	-1,1	-0,8	-0,4	-2,6	-3,5	-1,3
ZGAL_3475	L-fucose-proton symporter	0,3	0,2	0,3	0,3	0,4	0,6
ZGAL_3476	L-lactate dehydrogenase [cytochrome]	0,3	0,1	0,0	0,2	0,5	0,6

Tableau 6 : Résultats des analyses microarray lors de la croissance de *Zobellia galactanivorans* en milieu minimum pour les gènes du PUL sensu stricto. Les valeurs du tableau correspondent au log₂ de la différence avec les niveaux d'expressions observés lors de la croissance en glucose. Les valeurs significatives sont colorées en vert lorsqu'il s'agit d'une augmentation du niveau d'expression et en bleu lorsqu'il s'agit d'une diminution.

L'étude des deux GH de ce PUL, la GH29 et la GHnc, a donc eu pour but : 1) de vérifier la prédiction de l'activité fucosidase de la GH29, et 2) de tenter de déterminer l'activité de la GHnc, avec comme hypothèse principale qu'il devait s'agir d'une fucanase, active sur un polysaccharide de type fucoïdane d'algues brunes.

III. Résultats expérimentaux

1. Constructions clonées

Les deux GH du PUL sont des protéines modulaires, possédant toutes deux un peptide signal à leur extrémité N-ter qui a été exclu lors du clonage. Les différents modules ont été clonés séparément selon la méthodologie expliquée au chapitre 1, et les constructions correspondantes sont illustrées sur la figure 39.

La GH29 est constituée d'un module catalytique suivi d'un module supplémentaire en C-terminal, qui, bien que n'étant pas prédit comme un CBM lorsque sa séquence est entrée dans la base de données du NCBI, présente une prédiction de structure secondaire similaire à celles des CBM35 (prédictions Hyper et Phyre2).

La GHnc est prédite comme constituée d'un CBM35 du côté N-terminal puis d'un module sans homologie avec un domaine protéique connu.

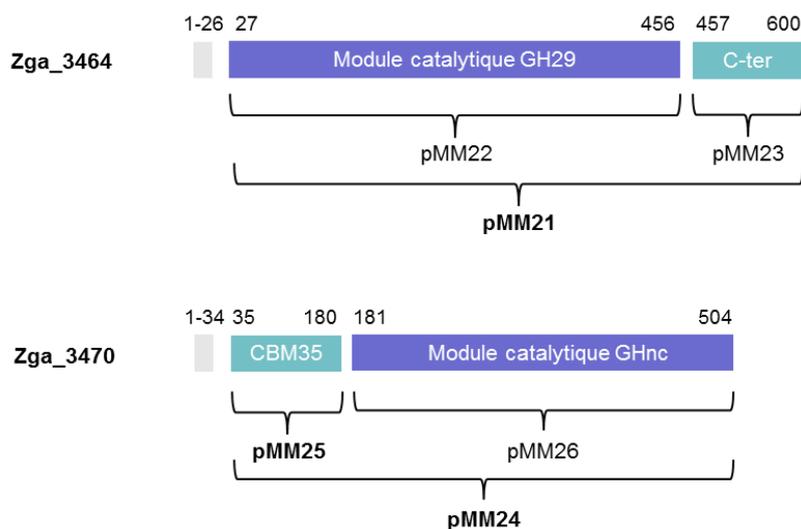


Figure 40 : Représentation modulaire de la GH29 (Zgal_3464) et de la GHnc (Zgal_3470) avec les constructions clonées durant la thèse (pMM21 à pMM26). En gris, le peptide signal, en mauve le module catalytique et en vert le module additionnel non catalytique.

Les tests d'expression à petite échelle ont montré que les constructions pMM21, pMM24 et pMM25 étaient particulièrement bien exprimées et sous forme soluble. Les deux constructions n'incluant que les modules catalytiques, bien que fortement exprimées, ne sont que partiellement solubles. Cette observation peut laisser penser à un mauvais repliement du module catalytique, imputable au choix du site de troncature de la séquence. À titre comparatif, lors du clonage de la κ -carraghénase de *Z. galactanivorans*, présentée au chapitre 2, la troncature après l'acide aminé N310 générait une protéine majoritairement insoluble, malgré un taux d'expression élevé. Le clonage indépendant d'une autre forme de l'enzyme avec une troncature après l'acide aminé S307 a suffi pour obtenir une protéine totalement soluble et toujours aussi fortement exprimée. Le choix des bornes lors du clonage de protéines modulaires est donc capital. Outre l'impact sur la solubilité de la protéine, ce choix peut également avoir un impact sur l'activité de l'enzyme.

2. Tests d'activité

2.1 Sur substrats synthétiques

Les deux constructions pMM21 et pMM22 de la GH29 ont été testées sur substrats synthétiques de type para-nitrophol : pNP-glucose, pNP-galactose et pNP-fucose. Ce test est couramment utilisé pour caractériser des enzymes exolytiques (Sakurama *et al.* 2012a; Guillotin *et al.* 2014; Cao *et al.* 2014). Les résultats, présentés en figure 40, montrent une activité spécifique de la forme entière de l'enzyme (pMM21) sur le pNP-fucose, confirmant ainsi la prédiction de l'activité fucosidase de cette GH29. La légère activité galactosidase observée sur le lysat bactérien brut est imputable à la β -galactosidase native d'*E.coli*, comme le confirme la disparition de cette activité suite à l'étape de purification sur chromatographie d'affinité.

La construction pMM22, outre sa propension à l'insolubilité, a par ailleurs perdu toute activité catalytique sur pNP-fucose. Ce phénomène n'est pas systématique, puisque pour reprendre l'exemple de la κ -carraghénase de *Z. galactanivorans*, la première construction partiellement insoluble ne présentait pas de perte significative d'activité. Cette différence d'effet pourrait s'interpréter à la lumière de la structure. En effet, si l'extrémité tronquée est impliquée dans un élément de structure impactant le site catalytique, cela pourrait expliquer la perte d'activité associée à une troncature déstabilisante d'un point de vue structurel.

Forts de l'observation réalisée sur le module catalytique de la GH29, les tests d'activité de la GHnc ont donc été réalisés uniquement sur sa forme entière (pMM24). La figure 41 illustre les résultats des tests d'activités obtenus sur des lysats bactériens non purifiés.

Contrairement à la GH29, la GHnc ne montre aucune activité sur pNP-fucose. Encore une fois, la légère activité galactosidase est vraisemblablement due à l'enzyme native d'*E.coli*.

Cette absence d'activité de la GHnc sur pNP-fucose pourrait indiquer que cette enzyme ne clive que des polysaccharides ou des oligosaccharides d'une longueur minimale, et non un monosaccharide terminal de type pNP-fucose. Ce qui serait compatible accord avec l'hypothèse de départ, à savoir qu'il s'agisse d'une enzyme de type fucanase.

Une autre possibilité est que la GHnc soit une fucosidase appartenant au groupe B. Ce sous-groupe des fucosidase de la famille de GH29 se caractérise par des enzymes très spécifiques pour leur substrat. Lors des tests d'activité, elles sont inactives sur substrat synthétique de type pNP (Sakurama *et al.* 2012b). D'après cette même classification, la GH29 appartiendrait quant à elle au groupe A, qui possède une plus grande tolérance, clivant des substrats de structures variées, synthétiques et naturels.

Étant donné la présence de trois sulfatases dans le PUL, j'ai également voulu les inclure dans les tests d'activité sur substrats naturels. En effet, certaines GH nécessitent l'action préalable de sulfatases avant de pouvoir cliver leur substrat. Sur les trois sulfatases, seules deux avaient été clonées avec succès lors de travaux antérieurs effectués par Robert Larocque, la S1-22 (Zgal_3466) et la S1-17 (Zgal_3469). Elles ont donc été produites en cultures de 250 ml et leur activité a été testée sur substrat synthétique de type MUF-sulfate (4-methyl-umbelliferyl-sulfate), qui fluoresce sous UV en présence d'une activité sulfatase. La figure 42 illustre les résultats de ces tests. Si la fluorescence est nette dans le cas de la S1-17 incubée à 37°C, celle de la S1-22 est plus discutable. De plus, les tests ayant été réalisés sur des lysats bactéries, il aurait fallu utiliser comme contrôle un lysat contenant le vecteur pFO4 vide pour tenir compte de l'activité des sulfatases endogènes d'*E. coli*. Néanmoins les deux sulfatases ont été intégrées aux tests sur substrats naturels. En effet, l'absence d'activité sur MUF n'est pas obligatoirement synonyme d'une inactivité de la sulfatase, la MUF restant un substrat synthétique avec les limites que cela implique.

DO ₄₀₀	GH29	
	lysat	purifiée
pNP-Glc	0,057	0,050
pNP-Gal	0,216	0,049
pNP-Fuc	3,634	3,498

Figure 41 : Test d'activité de la GH29 (pMM21) sur pNP avant purification (lysat) et après purification (purifiée) sur colonne d'affinité.

DO ₄₀₀	GH29	GHnc	pFO4
pNp-Glc	0,074	0,078	0,073
pNp-Gal	0,305	0,493	0,321
pNp-Fuc	3,282	0,108	0,093

Figure 42 : Tests d'activité de lysats bactériens clarifiés sur substrats pNP.

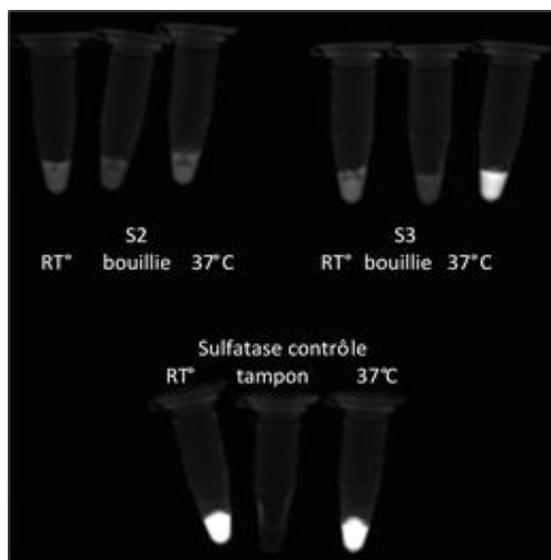


Figure 43 : Tests d'activité sur MUF des deux sulfatases du PUL disponibles, la S1-22 et la S1-17. La sulfatase contrôle positif est issue de *P. carrageenovora*.

2.2 Sur substrats naturels

2.2.1 Polysaccharides algaux

Comme décrit au chapitre 1, nous avons utilisé une mini-banque d'extraits polysaccharidiques issus de plusieurs espèces d'algues brunes, afin de tenter de déterminer le substrat naturel de nos deux GH. Les extraits provenaient des ordres des Fucales (N°11-12-13-14-17), des Laminariales (N°15-18) et des Ectocarpales (N°16).

Après un essai préliminaire non concluant avec la GHnc seule, nous avons décidé de procéder à des essais avec différentes combinaisons des quatre enzymes du PUL à notre disposition, à savoir les deux GH et deux des trois sulfatases. Les lysats bactériens clarifiés des enzymes ont été incubés 24 h à température ambiante en présence des différents substrats. Des gels FACE et C-PAGE ont ensuite été réalisés, à nouveau sans résultats. La figure 43 fournit un exemple de gel FACE obtenu suite à l'incubation simultanée des quatre enzymes avec neuf substrats différents. Bien qu'un profil de digestion semble apparaître en présence des enzymes (bandes blanches multiples dans les pistes avec enzymes), l'exacte conservation du profil entre les différents substrats testés laisse penser qu'il s'agit en fait d'un artefact du gel et non d'un réel profil de digestion.

Afin de prendre en compte la possibilité que la GHnc ne soit pas une enzyme de type endolytique, nous avons également ajouté au mélange réactionnel la fucanase A2 de *Mariniflexile fucanivorans*. Cependant, celle-ci est spécifique des fucoïdanes extraits de l'algue brune *Pelvetia canaliculata* (FS28), comme on peut le voir sur la figure 44, puisqu'aucune bande n'est visible dans la piste correspondant à la réaction de la fucanase sur fucoïdane extrait de *Fucus vesiculosus* (Ve). Mais le profil de digestion du substrat FS28 obtenu par l'action de la fucanase A2 ne présente aucune différence en présence des quatre autres enzymes du PUL.

De nombreux essais ont été faits, avec souvent des problèmes liés à la qualité des gels obtenus. Mais aucun d'eux n'a permis d'observer un profil différentiel reproductible entre les substrats avec et sans enzymes.

Plusieurs tentatives d'explication peuvent être avancées :

- Concernant les substrats tout d'abord : Au vue de la diversité des substrats algaux, il est très possible que notre mini-banque de 8 substrats extraits d'algues brunes ne contienne pas le substrat naturel des enzymes. D'autre part, bien que l'extraction des différents polysaccharides ait été menée en parallèle, afin d'assurer un maximum de similarité en termes d'abondance final en substrat, il est également possible que la concentration en motifs accessibles par nos enzymes soit trop faible pour permettre d'observer un profil de digestion par C-PAGE ou FACE.

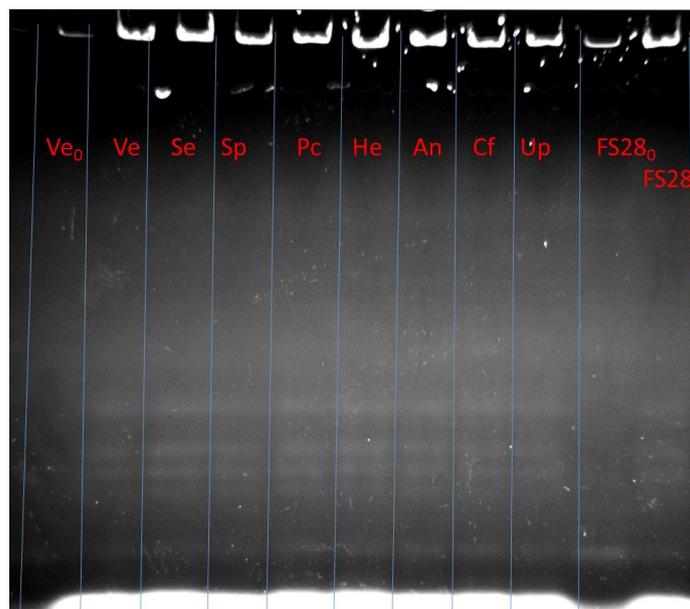


Figure 44 : Gel FACE avec marquage ANTS des substrats incubés en présence des quatre enzymes du PUL disponibles. Les indices '0' correspondent aux pistes avec substrats seuls sans enzymes. Ve : *Fucus vesiculosus*, Se : *F. serratus*, Sp : *F. spiralis*, P. carrageenovora : *Pelvetia canaliculata*, He : *Himanthalia elongata*, An : *Ascophyllum nodosum*, Cf : *Corda filum*, Up : *Undaria pinnatifida*, FS28 : fraction commerciale enrichie en fucanes sulfatés de *P. canaliculata*.

Oligo	+Fcn	Fcn+4enz	4 enz O/N
Fuc	Ve	FS28	Ve
		FS28	Ve
			FS28
			4enz

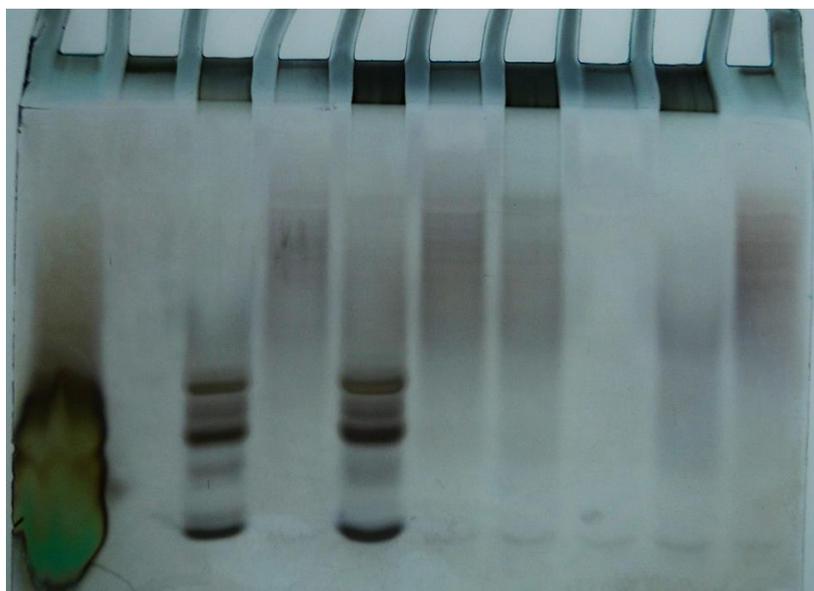
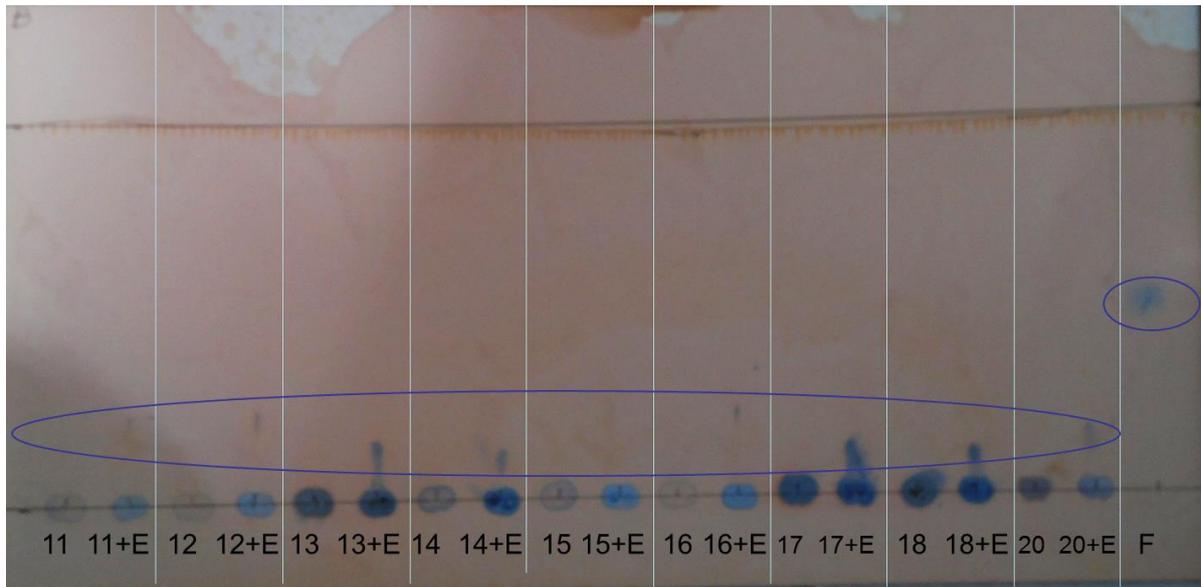


Figure 45 : Gel C-PAGE de différentes conditions d'incubations sur deux substrats, Ve (*F. vesiculosus*) et FS28 (*P. canaliculata*), avec la fucanase A2 (+Fcn), les quatre enzymes du PUL (4 enz O/N) ou le mélange des cinq (Fcn+4enz). Un mélange d'oligosaccharides issus de la dégradation de fucoïdanes par la fucanase A2 est inclus en piste 1, les substrats seuls sont déposés en pistes 8 et 9, et le mélange des 4 enzymes non purifiées est déposé en piste 10.

A



B

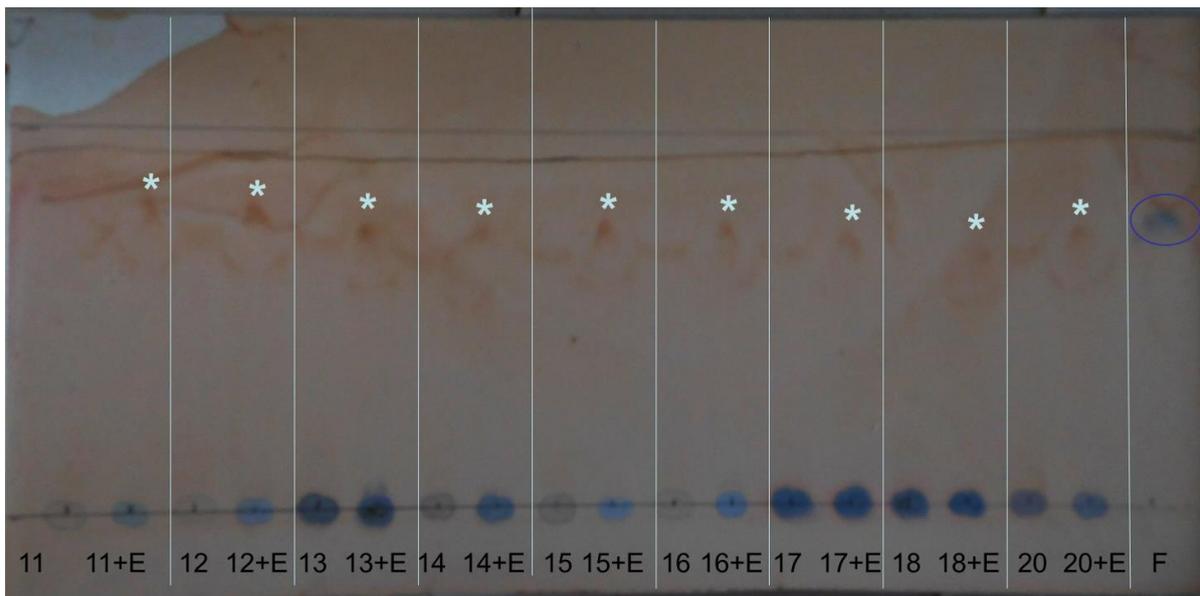


Figure 46 : Plaques TLC réalisées après incubation de la GH29 avec neuf polysaccharides algaux, et dans deux conditions différentes de tampon de migration, A : Butanol/EtOH/H₂O, B : isoprop/EtOH/H₂O. Les numéros des substrats font référence au tableau fourni au chapitre 1 (p.89), sauf le numéro 20 qui est un polysaccharide inconnu fourni par un laboratoire partenaire. La présence d'enzyme est indiquée par '+E'. Les étoiles (*) soulignent la présence de halos bruns migrant au niveau attendu pour le fucose, dans les pistes avec enzyme.

Pour pallier cette éventualité, des essais sur plaque TLC, plus adaptée pour la visualisation des monosaccharides, ont donc été réalisés pour la GH29 (Figure 45). Cependant ils n'ont pas non plus permis de trancher de façon définitive quant à la libération de fucose. En effet, bien qu'on observe des différences dans les profils de migration avec et sans enzyme, celles-ci sont difficilement interprétables. En A, des taches bleues migrent au-dessus de la ligne de dépôt en présence d'enzyme, mais bien en deçà du niveau de migration du fucose. En B, un halo avec une tache brune au centre apparaît à la hauteur du fucose en présence d'enzyme. Cependant, la tache correspondant au fucose en piste 19 apparaît bleue et non brune sur cette même plaque de TLC.

Des analyses par spectrométrie de masse (MS) ont donc été réalisées afin de déterminer si du fucose était libéré en présence de la GH29. Mais les trois réactions analysées (substrats N°13, 17 et 18, présentant les taches les plus intenses sur les plaques TLC en figure 45) se sont révélées négatives à la présence de fucose. Étant donnée la sensibilité de la technique de MS, cela signifie donc que la GH29 du PUL ne peut agir sur les polysaccharides algaux testés, tout du moins dans leur état initial après extraction.

- Concernant les enzymes : différentes raisons pourraient expliquer l'absence apparente de digestion des substrats naturels sur C-PAGE et FACE.

Tout d'abord, la GHnc pourrait ne pas être active dans les conditions testées (manque de co-facteur, problème de stabilité, pH et/ou température non optimale, ...). Une autre hypothèse serait que, tout comme la GH29, la GHnc soit en fait une exo-enzyme, retirant un motif terminal d'extrémité de chaîne ou de ramification latérale. Or, les gels C-PAGE et FACE sont d'avantage adaptés pour les enzymes de type endolytique, qui génèrent une grande quantité de fragments. Des tests sur TLC pourraient être envisagés, comme pour la GH29.

L'usage infructueux de la fucanase A2 de *M. fucanivorans* pourrait simplement signifier que les oligosaccharides issus de *P. canaliculata* ne contiennent pas les motifs reconnus par les deux GH du PUL. Il se peut donc que les fucoïdanes de *P. canaliculata* ne soient pas substrats des deux GH du PUL, tout du moins dans les conditions testées. L'absence d'autres fucanases empêche malheureusement le criblage de différents substrats qui pourraient être plus adaptés à nos enzymes.

Enfin, l'action des sulfatases, et en particulier de celle que nous ne possédions pas, pourrait être indispensable avant l'action des deux autres GH. Concernant les deux sulfatases utilisées, seule une des deux présentait une activité MUF significative. Cependant, d'autres travaux ont montré qu'une activité sulfatase sur substrat synthétique n'était pas forcément le gage d'une activité

sur substrat naturel, la réciproque étant également vraie (A. Préchoux, communication personnelle).

Pour pallier à l'absence de ces différentes enzymes (fucanases, sulfatases), une solution envisageable pour déterminer le substrat naturel des GH du PUL serait de réaliser une hydrolyse ménagée de différents polysaccharides algaux suivie de leur désulfatation chimique.

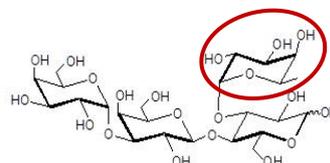
2.2.2 Oligosaccharides fucosylés issus du lait

En l'absence d'oligosaccharides algaux fucosylés de structure caractérisée, et afin de caractériser plus finement le type de liaison clivé par la GH29, nous avons testé son activité sur différents oligosaccharides fucosylés issus du lait humain. Leurs structures sont présentées en figure 46. Les deux types de liaisons α -1, 2 et α -1, 3 ont ainsi pu être testées, chacune dans deux environnements osidiques distincts : chaîne principale constituée de lactose ou de lactose/N-acétylglucosamine pour la liaison α -1, 2 et de galactose ou de lactose pour la liaison α -1, 3. Mais aucune de ces quatre structures n'a été reconnue par la GH29, comme en témoigne le profil TLC inchangé avec ou sans enzyme (Figure 47).

Etant donné que les oligosaccharides testés ne sont pas d'origine algale, il est possible que les oses présents dans la chaîne principale, en particulier le N-glucosamine, représentent une contrainte stérique empêchant l'activité de l'enzyme.

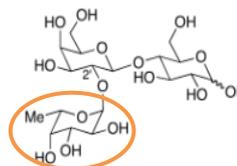
1: α -3Galactosyl-3Fucosyllactose

Liaison α 1-3



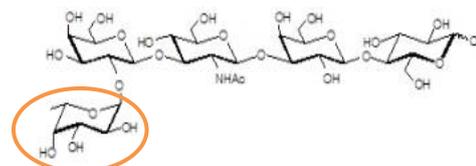
2: 2'-fucosyllactose = 2' FL

Liaison α 1-2



3: Lacto-N-fucopentaose I = LNFP I

Liaison α 1-2



4: 3Fucosyllactose = 3-FL

Liaison α 1-3

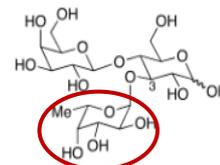


Figure 47 : Structure chimique des quatre oligosaccharides fucosylés issus du lait testés avec la GH29.



Figure 48 : TLC réalisée après incubation de la GH29 avec quatre oligosaccharides fucosylés. F : fucose, 1 : α 1-3Galactosyl-3Fucosyllactose, 2: 2'Fucosyllactose, 3 : Lacto-N-fucopentaose I, 4 : 3-Fucosyllactose, +E : substrat avec enzyme

3. Caractérisation biochimique de la GH29

Nous avons donc pu caractériser dans un second temps l'activité enzymatique de la GH29 sur pNP-fucose et calculer ses paramètres cinétiques (Figure 48A) :

$V_m = 35 \pm 4 \mu\text{M}\cdot\text{s}^{-1}$, $K_m = 628 \pm 182 \mu\text{M}$, $k_{cat} = 1530 \text{ s}^{-1}$, $k_{cat}/K_m = 2,4 \cdot 10^6 \text{ M}^{-1}\cdot\text{s}^{-1}$.

Ces valeurs sont dans le même ordre de grandeur que d'autres fucosidases déjà caractérisées et capables de cliver la liaison α -1,2 (Katayama *et al.* 2004 ; Cao *et al.* 2014). Sa température optimale est de 30°C, avec une décroissance de l'activité dès 35°C (Figure 49A). Le pH optimal est de 7 (Figure 49B), la concentration en NaCl de 200 mM (Figure 50A) et le tampon optimal est le tampon à pH 7 est le phosphate, suivi du PIPES et du MOPS (Figure 50B). La stabilité de la construction entière a également été évaluée au cours du temps, dans différentes conditions de température de stockage (Figure 48B). Au bout d'un mois, la congélation à -80°C semble la meilleure solution avec 81 % d'activité résiduelle. Mais la congélation à -20 °C ou à -80°C semble avoir un fort impact dès 2 mois de stockage (37 % et 35 % d'activité résiduelle respectivement) alors qu'à 4°C l'enzyme conserve 70 % de son activité initiale. Au bout de trois mois de stockage il n'y a plus de différence entre les trois températures (30 % d'activité résiduelle). Le stockage à 4°C semble donc suffisant pour garantir une bonne stabilité de l'enzyme jusqu'à deux mois après sa purification.

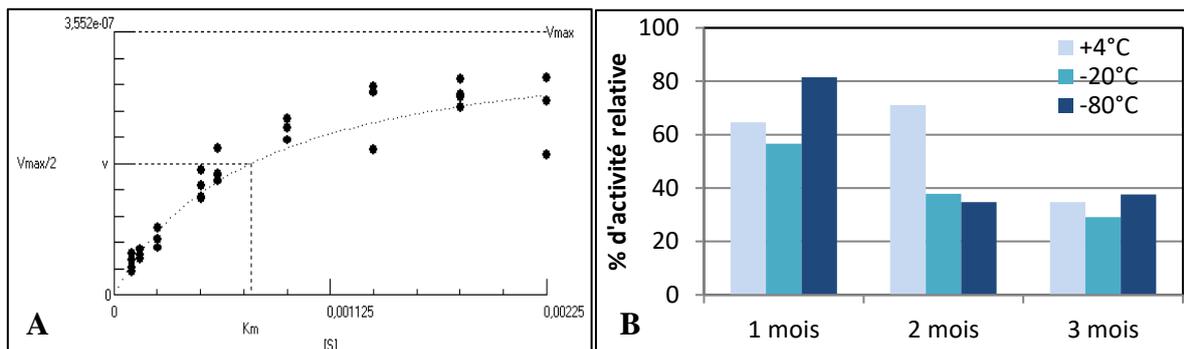


Figure 49 : Détermination des paramètres cinétiques de la GH29 par régression hyperbolique à l'aide du logiciel Hyper32 (A) et suivi de sa stabilité dans le temps en fonction de la température de stockage (B).

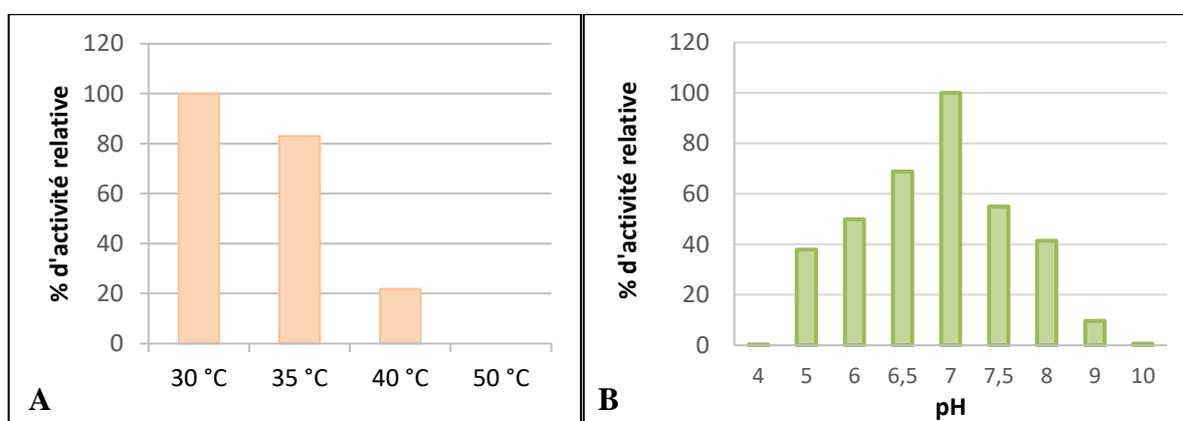


Figure 50 : Activité enzymatique de la GH29 sur pNP-fucose en fonction de la température d'incubation (A) et du pH (B).

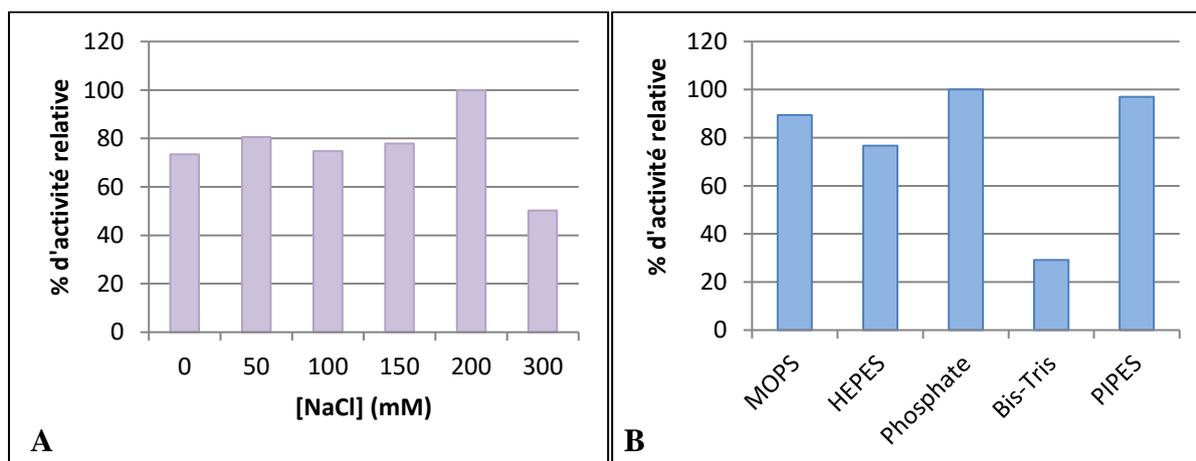


Figure 51 : Évaluation de la concentration optimale de NaCl (A) et du meilleur tampon (B) pour l'activité de la GH29 sur pNP-fucose.

4. Initiation de l'étude structurale

En parallèle de l'analyse biochimique des GH, leur cristallogénèse a été initiée en vue d'une étude structurale. Des cribles commerciaux de cristallogénèse ont été testés pour les deux constructions entières pMM21 et pMM24 (GH29 et GHnc respectivement).

Dans le cas de la GH29, des conditions de cristallisation ont été obtenues et reproduites manuellement permettant d'obtenir des cristaux exploitables (Figure 51). Ceux-ci ont diffracté à 2,1 Å, avec un groupe d'espace tétragonal (P41 ou P43) et un jeu de données complet a pu être obtenu sur la ligne Proxima2 du synchrotron SOLEIL. En utilisant la fucosidase de *Bacteroidetes thetaiotaomicron* (code PDB 4PEE) comme modèle (28 % d'identité), le groupe d'espace a pu être spécifié comme étant P43 ($a=b=117,62$, $c=91,22$), avec 2 molécules par unité asymétrique, donnant une solution par remplacement moléculaire (programme PHASER, TZF=9.9, Rfac=59,8 %). Cependant, la reconstruction de la structure complète nécessite encore un gros travail au vu de la faible similitude entre le modèle et la structure de la GH29 étudiée ici.

Concernant la GHnc, aucune condition de cristallisation n'a malheureusement pu être obtenue, malgré le recours à un nombre conséquent de cribles commerciaux.

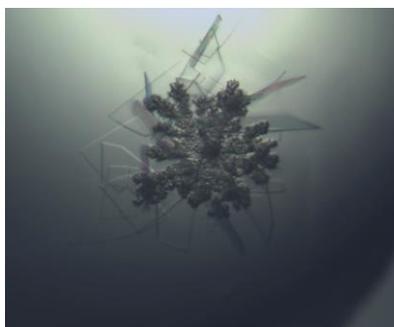


Figure 52 : Cristaux en plaquettes obtenus pour la construction pMM21 de la GH29.

IV. Discussion-Conclusion du chapitre

Cette étude exploratoire d'un PUL putatif chez *Z. galactanivorans* a permis de caractériser une nouvelle α -L-fucosidase, appartenant à la famille des GH29. La caractérisation biochimique sur substrat pNP-fucose a permis de déterminer les paramètres cinétiques de l'enzyme ainsi que ses conditions optimales d'activité et sa stabilité lors du stockage. La résolution de sa structure cristallographique est actuellement en cours. Il s'agit seulement de la deuxième fucosidase d'origine marine caractérisée à ce jour, ce qui offre de nouvelles perspectives pour la caractérisation des polysaccharides marins fucosylés.

La recherche des substrats naturels des deux GH du PUL putatif, bien qu'infructueuse, a néanmoins permis de tirer plusieurs conclusions importante pour la poursuite de l'étude de ce PUL :

1. La 'taille' d'un PUL peut être trompeuse et les substrats complexes

À mesure que le nombre de PULs caractérisés augmente, de nouveaux cas apparaissent où ce n'est plus un mais plusieurs PULs, parfois distants sur le génome qui sont co-régulés afin de dégrader un même substrat (Ndeh *et al.* 2017; Grondin *et al.* 2017). Bien que notre hypothèse de départ était que la GHnc soit une enzyme endolytique, il n'est pas à exclure que ça ne soit pas le cas, et qu'elle possède, tout comme la GH29, une activité exolytique. La possible absence de GH endolytique au sein de notre PUL putatif pourrait être un nouvel exemple illustrant la complexité des voies de dégradation des glucides au sein des génomes bactériens. Dans ce cas précis, étant donnée la présence des enzymes du catabolisme terminal du fucose au sein du PUL, il est possible qu'il s'agisse en fait un 'sous-PUL' terminal conçu pour le catabolisme d'oligosaccharides de fucose hautement sulfatés. Il n'est pas non plus à exclure que les motifs de fucose pris en charge par la GH29, soient en fait branchés sur une chaîne principale de nature très différente (galactose, acide uronique,...). En effet, en dehors des fucoïdanes au sens strict (à savoir moins de 10 % de résidus autres que le fucose), de nombreux autres polysaccharides algaux présentent des 'décorations' fucose. Ainsi, *Saccharina latissima* produit des fucogalactanes, fucoglucuromananes et fucoglucuronanes (Bilan *et al.* 2010), et *Dictyota dichotoma* renferme des polysaccharides riches en acides uroniques (Abdel-Fattah *et al.* 1978), pour ne citer que deux exemples d'algues ne faisant pas partie de notre « mini-banque ». L'enrichissement de cette mini-banque de substrats, initié au cours de cette thèse, devra donc

continuer afin de permettre d'augmenter les chances d'identifier le substrat naturel de ces enzymes.

2. L'importance des sulfatases

Pour maximiser les chances de succès dans la caractérisation d'un PUL, il faut avoir accès à l'ensemble des enzymes qu'il contient, et ce sous forme active. La présence de sulfatases au sein d'un PUL apporte un degré de difficulté supplémentaire. En effet, pour être actives, celles-ci nécessitent une étape de modification post-traductionnelle qui consiste en la formylation d'un résidu glycine du site catalytique (Appel and Bertozzi 2015). Or, dans certains cas, la machinerie post-traductionnelle d'*E. coli* ne suffit pas à assurer cette modification. Il faut alors employer une stratégie de co-expression de la sulfatase avec l'enzyme FGE (*formylglycine-generating enzyme*) responsable de la modification post-traductionnelle. La mise au point des conditions d'expression permettant d'obtenir des sulfatases recombinantes actives représente donc souvent un travail de longue haleine (Ficko-Blean *et al.*, sous presse).

3. La technique de détection est cruciale pour le criblage de substrats naturels

Le recours à une technique de détection des oligosaccharides telle que l'HPLC, pourrait être une approche intéressante pour la recherche d'activité enzymatique sur substrats naturels. En effet, la confection manuelle des gels C-PAGE principalement, mais aussi des FACE et en moindre mesure des TLC, peut entraîner de fortes hétérogénéités dans la qualité de visualisation des résultats : artefacts du gel, problèmes de migration, qualité de la révélation fortement dépendante du manipulateur pour les C-PAGEs. Ces techniques restent néanmoins très utiles en première approche ou pour des enzymes dont les substrats sont déjà connus. Mais dans des contextes plus exploratoires, où les conditions de concentration des substrats en particulier doivent être mises au point, l'HPLC pourrait être une alternative plus efficace.

**Conclusion générale
et perspectives**

Conclusion générale et perspectives

Malgré l'écart certain des connaissances entre les écosystèmes terrestres et marins, les nouveaux outils d'analyse de type « omique » (méta/génomique, méta/transcriptomique, métabolomique) offrent la possibilité à la communauté scientifique d'explorer la complexité des interactions à l'échelle d'écosystèmes entiers. Ainsi, à la suite des études sur les interactions entre la flore intestinale humaine et leur hôte (Gensollen *et al.* 2016; Schroeder and Bäckhed 2016; Sánchez *et al.* 2017), l'étude d'écosystèmes peu connus, tels que les communautés microbiennes marines associées aux algues, se développe de plus en plus (Teeling *et al.* 2012; Egan *et al.* 2013; Zhu *et al.* 2017). Ces outils d'analyse à grande échelle des données génomiques sont cruciaux pour avoir une vue d'ensemble des interactions au sein d'un même génome : co-régulation et co-expression de gènes impliqués dans des voies métaboliques associées ou encore phénomènes de duplication et d'évolution divergente de gènes aboutissant à l'émergence de nouvelles fonctions. À l'échelle des communautés bactériennes, cette connaissance permet également de retracer le cheminement évolutif ayant abouti au paysage microbien actuel : transferts horizontaux de gènes, coopération entre espèces pour la dégradation de certains substrats complexes.

Cependant, les génomes de bactéries marines séquencés contiennent souvent près de 40 % de gènes annotés comme codant des protéines putatives. Et un pourcentage non négligeable de ces gènes est présent au sein de clusters impliqués dans la dégradation de polysaccharides marins. C'est à ce niveau que la limite des analyses des données génétiques est atteinte. En effet, face à l'originalité des substrats marins, les protéines impliquées dans leur dégradation ne possèdent pas toujours d'homologues terrestres. Et même pour celles présentant des pourcentages d'identité suffisant pour être classées dans des familles de CAZymes connues, l'activité enzymatique n'est pas forcément prévisible. La caractérisation biochimique et structurale des enzymes identifiées dans les génomes est donc nécessaire. Et c'est là que se situe le goulet d'étranglement actuel. En effet, si la parallélisation de l'étape de clonage et d'expression de protéines recombinantes est réalisable grâce à des approches à moyen débit, comme celle suivie au cours de cette thèse, la caractérisation biochimique et structurale des protéines restent des étapes où le traitement individuel de chaque protéine est nécessaire. Au cours de cette thèse j'ai pu expérimenter les différents aspects de ces problématiques.

I. La stratégie moyen débit

Tirant profit des connaissances récemment acquises en termes de génomes de bactéries marines hétérotrophes (BMH) séquencés (Barbeyron *et al.* 2016; Schultz-johansen *et al.* 2016; Gobet *et al.*, in prep) et d'expériences de transcriptomique réalisées dans le cadre de l'étude du catabolisme des polysaccharides algaux (Thomas *et al.* 2017), j'ai eu recours à une stratégie de sélection et de clonage à moyen débit de plus de 30 cibles protéiques. Cette approche, présentée en détail au chapitre 1, a permis de mettre à disposition pour des études ultérieures 23 nouvelles enzymes, en plus des 4 étudiées durant cette thèse. Parmi elles, 19 sont parfaitement solubles, rendant possible une perspective de résolution structurale par cristallogénèse et diffraction des RX. Les nombreux avantages de cette stratégie ne doivent cependant pas faire oublier le contexte biologique sous-jacent. En effet, en particulier dans le cadre de l'étude de PULs (ou de clusters de type PULs) codant pour des enzymes de fonction inconnue, l'action coopérative de l'ensemble des enzymes doit être gardée à l'esprit si l'on veut réussir à identifier le substrat d'une enzyme donnée. Les cibles écartées dans une première approche de clonage en parallèle devront donc faire l'objet d'un clonage à part pour pouvoir se lancer dans la caractérisation complète d'un cluster.

II. La caractérisation d'enzymes de substrat inconnu

L'étude parallèle de deux GH d'un PUL potentiellement fucanolytique (chapitre 4), m'a permis de toucher du doigt la complexité de la recherche d'un substrat enzymatique adéquat. La multitude de substrats potentiels, la diversité et le caractère coopératif des enzymes impliquées dans la dégradation et enfin le faible nombre d'enzymes caractérisées à notre disposition (sulfatases, fucanase) sont autant d'éléments qui complexifient la tâche. À l'issue de cette étude préliminaire, de nombreuses pistes restent donc ouvertes pour mener à la caractérisation de la GHnc. Néanmoins, la caractérisation sur pNP-fucose de la GH29 de ce PUL offre un nouvel outil pour l'étude des polysaccharides algaux. En effet, il s'agit seulement de la 2^{ème} fucosidase de bactérie marine caractérisée, après celle de *Thermotoga maritima* (Sulzenbacher *et al.* 2004). La démonstration de son activité sous forme recombinante permet d'envisager son utilisation pour la détermination de son substrat naturel ainsi que celui de la GHnc. Une fois ce substrat identifié, elle pourrait servir d'outil d'analyse de ce substrat pour déterminer la structure chimique fine de celui-ci, mais également ses activités biologiques potentielles. En effet, les groupements fucoses décorant les polysaccharides sont des épitopes

importants dans les mécanismes immunitaires chez les plantes (Klarzynski *et al.* 2003), et les fucoidanes en général ont de nombreux effets immunomodulateurs et antioxydants sur les cellules animales (Deniaud-Bouët *et al.* 2017). Or, dans une perspective de valorisation des polysaccharides algaux, l'identification des éléments chimiques responsables d'une activité biologique représente un enjeu majeur pour la commercialisation d'un produit. Ainsi, la GH29 étudiée durant cette thèse pourra être un outil précieux afin d'expliquer les activités biologiques potentielles de son substrat.

III. Les voies de dégradation des carraghénanes

L'étude d'enzymes impliquées dans la dégradation des carraghénanes a permis de renforcer la compréhension des stratégies de dégradation de cette famille de substrats.

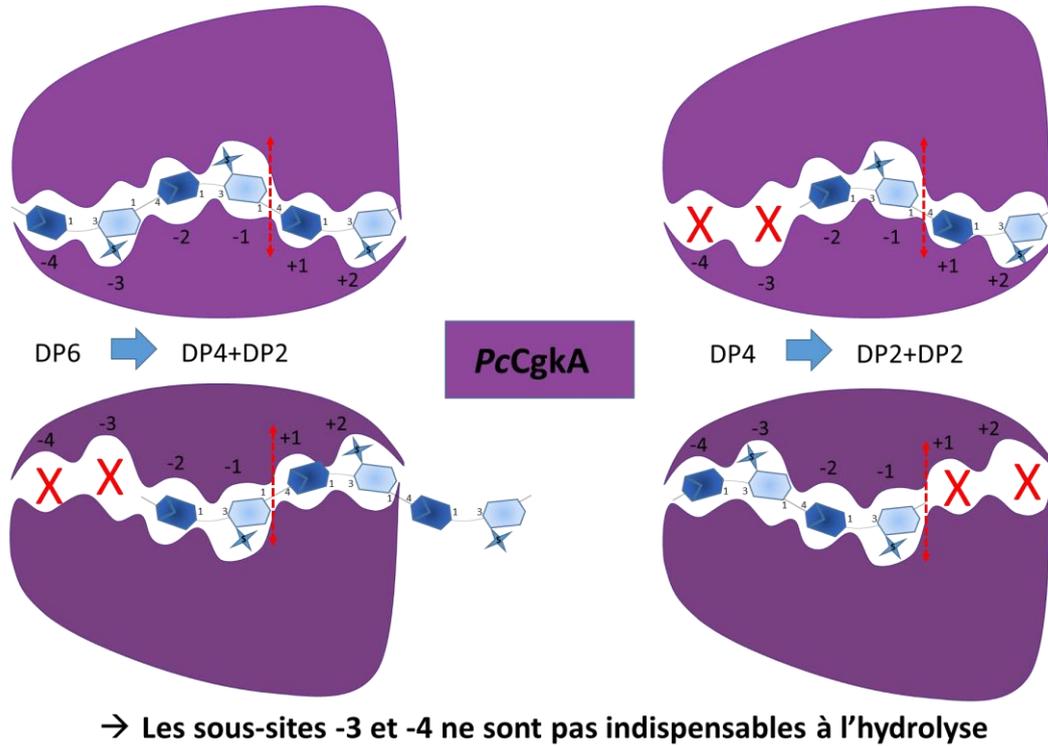
1. Les κ -carraghénases de *Z. galactanivorans* et de *P. carrageenovora*

En effet, grâce à l'étude biochimique et structurale comparée des κ -carraghénases de *Z. galactanivorans* (ZgCgkA) et de *P. carrageenovora* (PcCgkA), j'ai pu montrer l'existence d'au moins deux types de stratégies évolutives au sein de la sous-famille des κ -carraghénases (Chapitre 2, article accepté dans JBC).

La résolution de la structure cristallographique de ZgCgkA a permis de proposer une explication aux différences d'activités observées entre les deux enzymes. En effet, des modifications dans l'emplacement des structures de type 'feuillet-boucle-feuillet' bordant le tunnel catalytique génèrent des différences de flexibilité des enzymes, ainsi qu'une modification du nombre et de la position des sous-sites catalytiques (figure 52). Il en découle une plus grande flexibilité de ZgCgkA, qui pourrait expliquer sa meilleure efficacité catalytique en solution, mais également les différences de taille des oligosaccharides qu'elle peut cliver. La processivité de ZgCgkA a également pu être démontrée sur microgel, pour la première fois.

Ces résultats m'ont ainsi amenée à proposer un déplacement processif de ZgCgkA sur son substrat, dans le sens opposé de celui de PcCgkA, les sous-sites positifs servant de point d'ancrage pour la progression de l'enzyme. De façon intéressante, les différences structurales mises en évidence chez ces deux enzymes sont conservées à l'échelle phylogénétique. Les différences d'efficacité catalytique qui leurs sont associées pourraient donc être interprétées à la lumière du contexte écologique bactérien.

A



B

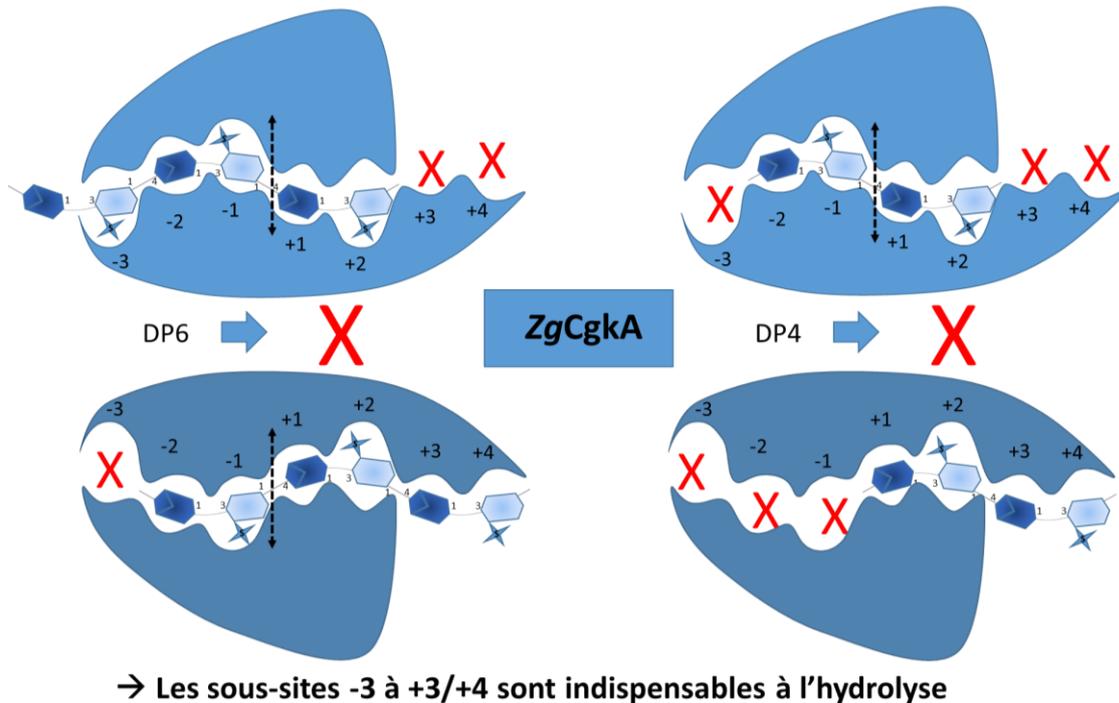


Figure 53 : Schémas résumant les interactions possibles entre les deux κ -carraghénases étudiées, PcCgkA (A) et ZgCgkA (B), et des oligosaccharides κ de DP6 et DP4, en prenant en compte les deux positionnements possibles des enzymes pour chaque DP. A : La κ -carraghénase de *P. carrageenovora* est capable de cliver les DP6 et DP4. B : La κ -carraghénase de *P. carrageenovora* par contre est incapable de cliver les DP6 et DP4.

2. Les GH42-like et la dégradation des carraghénanes par *P. carrageenovora*

En effet, le genre de *Bacteroidetes* est connu pour le nombre élevé de CAZymes codées dans les génomes de certains de ses représentants, dont *Z. galactanivorans*. Ces bactéries ont été associées à la dégradation de substrats complexes et récalcitrants à la dégradation d'autres genres bactériens (grands fonds marins par exemple). Cependant, au contact des macroalgues, les biofilms abritent aussi bien des *Bacteroidetes* que des *Gammaproteobacteria*, auxquelles appartient *P. carrageenovora*. Il est donc probable qu'un système de coopération ait lieu entre les différents genres bactériens partageant un même biofilm. Il est aussi possible, comme cela a été observé dans le cadre des blooms phytoplanctoniques (Teeling *et al.* 2016), qu'un système de successions temporelles se mette en place en termes d'abondance relative des différents phyla. Ce système pourrait avoir un lien avec la variabilité de la nature du substrat au cours du temps, tel que le changement du type de carraghénane en fonction du stade du cycle de reproduction de l'algue par exemple. La compréhension du contexte écologique, mais aussi de la complexité des substrats polysaccharidiques algaux se révèle donc primordiale pour élucider leurs voies de dégradation par les BMH.

En m'appuyant sur les travaux antérieurs réalisés chez *P. carrageenovora*, ainsi que les résultats récents du séquençage de son génome, j'ai ainsi pu proposer un modèle de dégradation conjointe des κ -, ι - et β -carraghénanes chez cette bactérie, potentiellement induit par la voie de dégradation du λ -carraghénane. Et grâce au travail réalisé en collaboration avec l'équipe danoise de Peter Stougaard, nous avons également initié la caractérisation de deux enzymes homologues chez *P. carrageenovora* et *P. hydrolytica*, responsables d'une des étapes clés du métabolisme des carraghénanes : l'hydrolyse de la liaison β -(1,4) entre un néo- β -carrabiose et un néo- κ -carrabiose ou tétraose. J'ai ainsi pu proposer la création d'une nouvelle famille de GH apparentée aux GH42, et dont au moins une sous-famille serait spécifique du β -carraghénane.

Les carraghénanes extraits des algues rouges carraghénophytes sont des molécules aux applications extrêmement variées (cf. introduction). Dans une perspective de recherche appliquée, le clonage de la κ -carraghénase de *Z. galactanivorans* représente donc un atout majeur. En effet, son efficacité catalytique est nettement supérieure à celle de *P. carrageenovora*, et les fragments libérés sont de taille différente. Les oligosaccharides générés pourraient ainsi avoir des effets biologiques nouveaux, en particulier dans le cas de la digestion de carraghénanes hybrides de type ι/κ . Cette enzyme représente également un nouvel outil afin de caractériser la structure fine de ces polysaccharides hybrides, encore mal connus. De par son

activité endolytique, elle pourrait enfin être utilisée pour enrichir des polysaccharides hybrides de type κ/ι en motifs ι . En effet, leur taux de sulfatation supérieur à celui des motifs κ leur confère généralement des activités biologiques plus importantes.

IV. Perspectives

Une meilleure connaissance de la diversité et de la spécificité des enzymes actives sur polysaccharides algaux représente un enjeu majeur pour le développement des biotechnologies bleues. Que ce soit pour des applications bien-être ou santé, à destination humaine, animale ou végétale, l'utilisation des enzymes marines représente un formidable secteur d'innovation. Et les polysaccharides marins, en particulier sulfatés, ouvrent des perspectives plus que prometteuses pour résoudre les grandes problématiques de santé publique, telles que la résistance aux antibiotiques, la lutte contre les maladies cardio-vasculaires, l'obésité, les cancers ou les MST.

Mais pour cela, le travail de recherche fondamental est essentiel car face à la diversité des substrats marins, les outils enzymatiques à notre disposition sont encore bien peu nombreux. Au travers de ce travail de thèse, je pense avoir apporté ma contribution à une meilleure compréhension de certains mécanismes biochimiques relatifs à l'efficacité catalytiques des κ -carraghénases. Les réflexions menées sur les voies de dégradation des carraghénanes offrent également des perspectives prometteuses pour de futures études. Le décryptage précis des différentes enzymes et des systèmes d'induction de ces voies métaboliques sont en effet nécessaires afin 1) de comprendre les mécanismes qui sous-tendent les interactions algues/bactéries, et 2) de disposer de nouveaux outils de modification des polysaccharides marins sulfatés.

En ce début de XXI^{ème} siècle, le regain d'intérêt pour les richesses marines peut amener une réflexion plus globale sur l'évolution de l'humanité et de la vie sur Terre en général. C'est le caractère marin de notre « planète bleue » qui la distingue de toutes les autres planètes du système solaire. En effet, c'est dans la mer que sont apparues les premières formes de vie il y a près de 4 milliards d'années. Dans cet océan primordial, à l'abri des radiations solaires destructrices, les premières molécules chimiques ont pu se combiner, donnant naissance aux briques élémentaires du vivant. Parmi elles, les glucides. Molécules aux propriétés auto-répliquatives, catalytiques, informatives, structurantes et protectrices, les glucides et leurs polymères, les polysaccharides, seraient à la base même du vivant selon certains. De ce monde « sucré » serait ensuite (ou en parallèle) né le monde à ARN puis à ADN que l'on connaît

aujourd'hui (Stern and Jedrzejak 2008). Il n'en reste pas moins vrai que la biomasse dominante sur terre en termes de carbone organique est représentée par les polysaccharides, la cellulose puis la chitine par ordre d'abondance (Horn *et al.* 2006b). Et si aujourd'hui les surfaces émergées ont pu être colonisées par une multitude d'espèces animales et végétales, c'est bien grâce à l'action des premiers organismes photosynthétiques marins qui ont permis la formation de la fameuse couche d'ozone. Au vu des près de 4 milliards d'années d'évolution de la vie sur Terre, la colonisation des continents par nos ancêtres tétrapodes représente bien la conquête toute récente (~400 millions d'années...) d'un nouveau monde (Shubin *et al.* 2014).

De notre origine aquatique, nous avons conservé le souvenir au travers de nos neuf premiers mois d'existence dans la poche placentaire. Et bien qu'adaptés à la vie terrestre à l'air libre, nos cellules sont constituées à 70 % d'eau. La composition minérale de notre plasma sanguin présente également de nombreuses similitudes avec celle de l'eau de mer. Enfin, à l'instar de nombreux polysaccharides algaux, notre matrice extracellulaire renferme des GAGs sulfatés. Si la mer a pu véhiculer dans de nombreuses cultures la peur face à l'inconnu et à l'insondable, mettant l'homme face à ses propres limites, elle a aussi été, et est encore, la 'mer-nourricière', le second 'poumon de la planète', la régulatrice des climats et la médiatrice des échanges commerciaux et culturels entre les hommes. Aujourd'hui comme hier, elle ouvre la route vers de nouveaux mondes à explorer. Car la connaissance des écosystèmes marins est cruciale pour pouvoir y trouver de nouvelles solutions aux besoins de l'humanité. Mais cette connaissance est aussi nécessaire pour assurer une exploitation durable des ressources marines, dans le respect des besoins et de la diversité des espèces qui y vivent.

“Earth provides enough to satisfy every man's needs, but not every man's greed.” *Mahatma Gandhi*

Annexes

Annexes

Annexe 1 : Solutions tampon

- **Sph (spheroplasting) : autoclaver puis stocker à 4°C**

Utilisé pour reprendre le culot bactérien afin d'obtenir des sphéroplastes.

Tris-HCl, pH 8	50 mM
Saccharose	25%

- **TES : autoclaver puis peut être gardé à température ambiante**

Tampon d'incubation du lysozyme. L'apport de NaCl et d'EDTA permet de commencer à inhiber les nucléases lors de l'extraction de l'ADN nucléoïdique.

Tris-HCl pH8	50 mM
EDTA	5 mM
NaCl	50 mM

- **Solution de lyse : peut être gardé à température ambiante**

Apporte beaucoup d'EDTA sur les sphéroplastes après le SDS et avant l'ajout de protéinase K

Tris-HCl pH 8	50 mM
EDTA	100 mM
NaCl	100 mM

- **TE : peut être gardé à température ambiante**

Utilisé à la place de l'eau pour stocker à 4°C ou à -20°C l'ADN

Tris-HCl pH 8	10 mM
EDTA	1 mM

Annexe 2 : Protocoles de biologie moléculaire

1. Extraction d'ADN génomique bactérien :

Après 48h de culture de la bactérie désirée en milieu Zobell, à température ambiante et sous agitation à 200 rpm, 100 ml de culture sont centrifugées à 3000 g pendant 15 min. Le culot obtenu est congelé à -20°C.

Le culot est repris dans 25 ml de tampon Sph et 5 ml de tampon TES additionné de 50 mg de lysozyme, et laissé à incuber 15 minutes à température ambiante. Cette première étape permet d'obtenir des sphéroplastes par digestion du peptidoglycane par le lysozyme. Le saccharose augmente la pression osmotique pour empêcher les sphéroplastes ainsi formés d'éclater.

6 ml d'EDTA sont ensuite ajoutés à 500 mM (soit 100 mM final) afin de protéger l'ADN de l'action des DNases et le mélange est incubé sur glace 10 min.

La lyse membranaire est achevée par l'ajout de 3,6 ml de SDS 20%, soit 2% final, et de 25 ml de solution de lyse. Le SDS dissout les lipides membranaires et la concentration élevée en EDTA favorise cette dissolution par chélation des ions magnésium, tout en protégeant l'ADN ainsi libéré. La dénaturation des protéines se fait par addition de 40 mg de protéinase K et une incubation d'1h à 50°C.

Après l'incubation, l'addition de 16 ml de perchlorate de sodium 5 M, soit 1 M final, permet de finaliser la rupture des liaisons ADN-protéines et d'éliminer le SDS. La déprotéinisation de la solution est obtenue par ajout de 40 ml de phénol saturé, suivi d'une agitation manuelle de 5 min. Il faut trouver le bon compromis entre force et douceur pour permettre l'émulsion des 2 phases et donc l'extraction des protéines, sans casser l'ADN. 40 ml d'un mélange chloroforme/alcool isoamylique (24/1) sont ajoutés et puis le mélange est agité encore 5 min, avant de centrifuger 15 minutes à 10 000 g, à température ambiante. En effet, les phénolates précipitent à froid.

En fin de centrifugation, les protéines sont concentrées à l'interface entre la phase organique et la phase aqueuse. Celle-ci est pipetée à l'aide d'un cône de 5 ml dont l'embout a été agrandi pour ne pas casser l'ADN. L'ajout de 80 ml de chloroforme/alcool isoamylique (24/1) à la phase aqueuse contenant les acides nucléiques va achever la déprotéinisation et enlever toute trace de phénol. S'en suit une agitation de 5 min puis 5 min de centrifugation à 10 000 g, toujours à température ambiante.

La phase aqueuse obtenue est à nouveau pipetée et additionnée à 0,6 fois son volume d'isopropanol. L'ADN est récupéré à l'aide d'une baguette en verre, en mélangeant toujours dans le même sens jusqu'à homogénéisation des deux phases. L'ADN enroulé autour de la

baguette est lavé dans 30 ml d'éthanol 70%, afin d'éliminer les sels, puis déshydraté dans 30 ml d'éthanol absolu, avant de le laisser sécher à l'air durant 16h. L'ADN est finalement dissout dans 15 ml de Tris 1M, pH 8, par agitation douce durant 12h, avec chauffage à 50 °C si nécessaire. La pureté de l'ADN est évaluée au NanoDrop par mesure des ratios 260/280 et 260/230.

2. Gradient de CsCl :

Pour un tube de volume final de 12,5 ml, peser 10,206 g de CsCl (816,5 mg/ml final) et les dissoudre dans 4 ml de tampon TE.

Ajouter l'ADN à purifier, au maximum 500 ng, puis une solution de BET à 10 mg/ml pour être à 750 µg/ml final et compléter à 12,5 ml avec du tampon TE.

Après avoir scellé le tube, bien mélanger puis ultracentrifuger à 45000 rpm pendant 48h à 20°C (pour que le CsCl ne précipite pas).

La bande d'ADN est récupérée sous lampe UV au moyen d'une aiguille piquée au travers du tube scellé à hauteur de la bande fluorescente (Figure A2.2).

L'ADN est ensuite lavé dans 2 volumes de butanol saturé en eau, par agitation puis élimination du surnageant. Faire autant de lavages que nécessaire pour que les deux phases soient de la même couleur, plus un lavage supplémentaire.

Précipiter l'ADN dans 1 volume d'éthanol absolu et mettre à -20°C pour la nuit.

Centrifuger 30 min à 10000 g pour éliminer l'éthanol puis reprendre le culot dans de l'éthanol 70 % et centrifuger 10 min à 13000 g. Eliminer l'éthanol et laisser évaporer à température ambiante.

Resuspendre dans 20 µl de tampon TE et doser au Nanodrop.

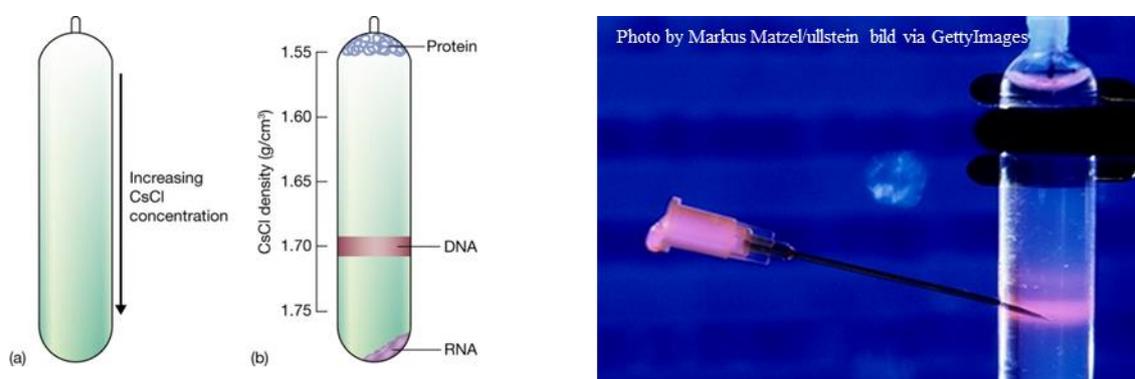


Figure A2.2 : Illustration du protocole de purification d'ADN génomique par gradient de CsCl

3. Amplification PCR des inserts à cloner dans pFO4 :

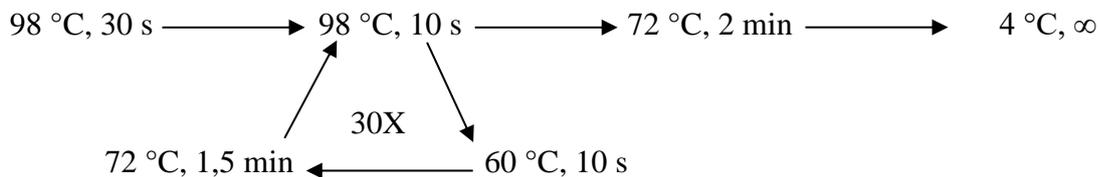
Composition du mélange pour PCR pour une réaction :

- 0,5 µl de Q5 polymérase (NEB)
 - 10 µl de buffer 5X pour Q5 polymérase
 - 1 µl de dNTPs à 10 mM
 - 36,5 µl de dH₂O
- } à multiplier par le nombre de réactions prévues + 5% de plus

Mélange réactionnel par réaction PCR (volume final = 50 µl):

- 48 µl du mélange pour PCR
- 0,5 µl de chaque amorce
- 1 µl d'ADN génomique dilué dans du TE à la dilution établie au préalable

Programme PCR :



Vérification de l'amplification sur gel d'agarose 1 %:

- 0,5 µl de tampon de charge 10X
 - 5 µl de la réaction PCR
- } Migration 30 min à 100 V

4. Digestion des inserts à cloner dans pFO4 :

Après purification des produits de PCR sur plaques de « PCR-cleaning » (QIAquick 96 plate, de QIAGEN) et dosage au Nanodrop, quatre solutions stock sont préparées, une pour chaque enzyme de restriction :

- 7 µl de tampon CutSmart 10X (NEB)
 - 0,5 µl de XhoI ou BamHI ou SalI
ou 1 µl de BglII
 - 5 µl de dH₂O ou 4,5 µl pour BglII
- } Multiplier par le nombre de réactions prévues + 5% de plus

Mélanger 60 µl d'insert purifié à 12,5 µl du mélange de restriction avec la 1^{ère} enzyme de restriction (BamHI ou SalI) et incubé 4h à 37°C, puis purifier avant d'incuber avec la 2^{ème} enzyme de chaque couple (XhoI ou BglII). La purification entre les deux incubations se fait sur les mêmes colonnes de « PCR-cleaning » régénérées par 2 lavages successifs en Tris 1 M.

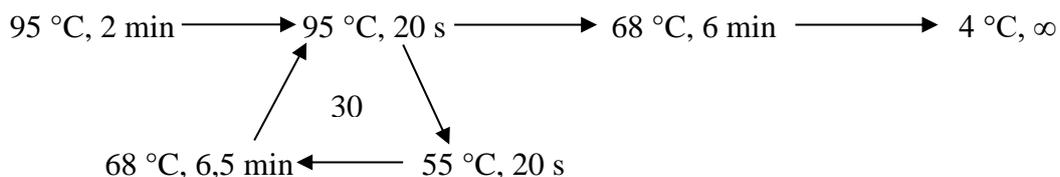
Après une dernière étape de purification sur colonne, un aliquot de chaque réaction est déposé sur gel d'agarose 1 % pour vérifier la présence et la taille des fragments, par comparaison avec des fragments de taille standard (Eurogentec). La concentration des réaction PCR est enfin évaluée au Nanodrop.

5. Amplification par PCR « inside-out » du vecteur pFO4 et digestion par le couple d'enzymes de restriction adapté à la stratégie de clonage

A partir d'un vecteur circulaire pFO4 diluée au 20^{ème} on réalise le mélange suivant :

- 2,5 µl de chaque amorce
- 20 µl de « 5X PCR buffer »
- 2 µl de Taq Herculase II fusion
- 1 µl de dNTP à 100 mM
- 2 µl d'ADN plasmidique
- 70 µl de dH₂O

On applique au mélange le programme PCR suivant :



Après la PCR, on procède à un lavage sur colonne « PCR cleaning » avant de doser au Nanodrop la quantité de plasmide linéaire obtenu. Un gel d'agarose à 0,7% est réalisé pour contrôler la qualité de l'ADN.

Le vecteur linéaire est ensuite digéré avec le couple d'enzyme de restriction adapté à la stratégie de clonage (dans notre cas BamHI/XhoI) selon le protocole suivant :

- 47 µl du vecteur pFO4 amplifié précédemment
- 1 µl de chaque enzyme de restriction
- 1 µl de DpnI (pour digérer le plasmide circulaire méthylé ayant servi de matrice durant la PCR)
- 6 µl de tampon CutSmart (NEB)
- 4 µl de dH₂O

Incuber 4h à 37°C, puis réaliser un gel d'agarose 0,7 %, en fusionnant 2 puits d'un peigne à dents larges afin de pouvoir déposer les 60 µl de réaction d'un seul tenant. Après migration,

découper sur plaque la bande correspondant au plasmide linéaire digéré puis réaliser une extraction de l'ADN sur une colonne QIAgen (Gel Extraction Kit) et doser au Nanodrop.

6. Ligation et transformation bactérienne

Pour chaque insert à cloner, ajouter à 5 ng (1 µl) 4 µl du mélange :

- 10 ng de vecteur linéaire (0,7 µl)
 - 0,5 µl de ligase diluée par 2,5 dans son tampon 1X
 - 0,5 µl de tampon de ligation 10X
 - 2,3 µl de dH₂O
- } Multiplier par le nombre de réactions prévues + 5% de plus

Incuber 3-4h à la température de la pièce puis transformer 45 µl de cellule compétente *E. coli* NEB5α avec les 5 µl de la réaction de ligation selon le protocole suivant :

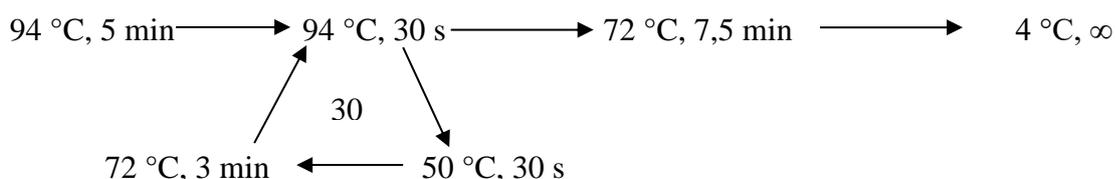
- Décongeler les cellules compétentes sur glace
- Mettre sur glace le tube de ligation contenant les 5 µl de réaction
- Ajouter 45 µl de cellule compétente et incuber sur glace 30 min
- Faire un choc thermique de 45 s à 42 °C et remettre sur glace au moins 2 min
- Etaler sur boîtes de Pétri LB + ampicilline préalablement mises à température ambiante
- Incuber à 37 °C sur la nuit

7. Criblage par PCR sur colonie

Mélange réactionnel pour 250 réactions (dans le cas de cette thèse) :

- 50 µl d'amorce T7Forward à 100 µM
- 50 µl d'amorce T7Reverse à 100 µM
- 2,5 ml de « 2X PCR mix » (contenant les dNTPs, la polymérase, le tampon) (Promega)
- 2,4 µl de dH₂O RNase-free

Aliquoter 20 µl du mélange dans chaque puits d'une plaque PCR 96 puits, et plonger dans chaque puits le cône ou l'ose plastique ayant servi à prélever 1 colonie bactérienne avant de lancer le programme PCR suivant :



Les produits de PCR sont ensuite visualisés sur gel d'agarose 1 %.

8. Réalisation des minipreps des clones et digestion par HincII

Les colonies positives suite au criblage PCR sont incubées à 37°C pendant une nuit dans 6 ml de milieu LB + ampicilline sous agitation à 200 rpm. On utilise 4 ml de culture pour réaliser l'extraction des plasmides (obtention de « minipreps ») selon le protocole du kit « QIAprep Spin Miniprep ». 1 miniprep sur 4 est contrôlée par dosage au Nanodrop.

Afin de contrôler la qualité des plasmides obtenus, 15 µl de chaque miniprep sont mélangés à 5 µl du mélange contenant l'enzyme HincII:

- | | | |
|---|---|---|
| <ul style="list-style-type: none"> - 2 µl de tampon CutSmart (NEB) - 3 U d'enzyme HincII - 2,7 µl de dH₂O | } | <p>Multiplier par le
nombre de réactions
prévues + 5% de plus</p> |
|---|---|---|

Après incubation pendant 2h à 37°C et migration sur gel d'agarose, le profil de digestion obtenu est contrôlé pour chaque miniprep. En effet, le nombre de bandes attendues sera différent avec ou sans insert.

Si le profil de digestion est conforme à ce qui est attendu, le 1 ml de culture restant (stocké à 4 °C entre temps) est utilisé pour réaliser un stock glycérol de chaque clone en cellule NEB5α (souche de transformation) en ajoutant 500 µl de glycérol 50 %.

A partir de 2 µl de la miniprep obtenue précédemment, 25 µl la souche d'expression BL21(DE3) d'*E.coli* sont transformés selon le protocole détaillé au point 3 de cette annexe.

Annexe 3 : Données de biologie moléculaire

1. Tableau des amorces utilisées lors du clonage moyen débit

Nom clone	Famille	pFO4	Stratégie de clonage	Protection/ <i>HE/STOP</i>	Codon matrice	Tm pb	Nom	Amorce
Zg-236_cat	GH16	pMM1	BamHI/XhoI	TTTTTCTCGAGTTA	GTACCTGCTGATTTTCCAGAC	70	OMM38	TTTTTCTCGAGTTAGTACTGCTGATTTTCCAGAC
Zg-3602_sp	GH97	pMM2	BamHI/XhoI	TTTTTCTCGAGTTA	TTATTTTGGTAAACTTATGGTGGC	70	OMM39	TTTTTCTCGAGTTATTTTTGGTAAACTTATGGTGGC
Zg-4674	AD3	pMM3	BamHI/XhoI	TTTTTCTCGAGTTA	TAATAAAAATACAACTCTTTGGCA	70	OMM40	TTTTTCTCGAGTTATCAAAAATACAACTCTTTGGCA
Pc0024_sp	dist GH42	pMM4	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM41	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
Pc0027_sp	S1-nc	pMM5	BamHI/XhoI	TTTTTCTCGAGTTA	TAATAAAAATCTTCACTGGTGTATAA	70	OMM42	TTTTTCTCGAGTTATAAAAATCTTCACTGGTGTATAA
Pc0028_sp	S1-19	pMM6	BamHI/XhoI	TTTTTCTCGAGTTA	TAATAAAAATCTTCACTGGTGTATAA	70	OMM43	TTTTTCTCGAGTTATAAAAATCTTCACTGGTGTATAA
Pc0033_sp	GH82, iotase	pMM7	BamHI/XhoI	TTTTTCTCGAGTTA	TTATAGCTAAATCACTAGCGCTT	70	OMM44	TTTTTCTCGAGTTATTAAGCTAAATCACTAGCGCTT
Pc0034_sp	GH16	pMM8	BamHI/XhoI	TTTTTCTCGAGTTA	TTATAGCTAAATCACTAGCGCTT	70	OMM45	TTTTTCTCGAGTTATTAAGCTAAATCACTAGCGCTT
Pc0058_sp	GHnc, lambdaase	pMM9	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM46	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
Pc0073_sp	GHnc, lambdaase	pMM10	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM47	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
Pc0075_sp	GH110, a-Gal	pMM11	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM48	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
Pc0076_sp	S1-8	pMM12	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	72	OMM49	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
Pc0077	GH2, b-Gal	pMM13	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM50	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF20580_sp	GH50	pMM14	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM51	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF110093_sp	GH16	pMM15	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM52	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF110095_sp	S1-nc	pMM16	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM53	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF110098_sp	GH16	pMM17	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM54	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF110100_sp	GH82, iotase	pMM18	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	72	OMM55	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF110102_sp	S1-19	pMM19	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM56	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
AF110113	GH127, b-arabinofura	pMM20	BamHI/XhoI	TTTTTCTCGAGTTA	TTAAGCTGTAATCAACAAGTAA	70	OMM57	TTTTTCTCGAGTTATTAAGCTGTAATCAACAAGTAA
Zg-3464_sp	GH29	pMM21	BglII/SalI	TTTTTGTGAC	TTACTCTACTCTCTCCGGTGC	72	OMM58	TTTTTGTGACCTACTCTCTCTCCGGTGC
Zg-3464_cat		pMM22	BglII/SalI	TTTTTGTGAC	CGCTTAGCAAAAGTCAACCGGGA	70	OMM59	TTTTTGTGACCTAGCTTAGCAAAAGTCAACCGGGA
Zg-3464_cbm		pMM23	BglII/SalI	TTTTTGTGAC	TTACTCTACTCTCTCCGGTGC	72	OMM58	TTTTTGTGACCTACTCTCTCTCCGGTGC
Zg-3470_sp	GHnc	pMM24	BglII/SalI	TTTTTGTGAC	TACCGAACTACTCTCCGAACTCT	70	OMM60	TTTTTGTGACCTTACCGAACTACTCTCCGAACTCT
Zg-3470_cbm		pMM25	BglII/SalI	TTTTTGTGAC	ACTTGGCTGTCACTCCAAAG	72	OMM61	TTTTTGTGACCTACTTGGCTGTCACTCCAAAG
Zg-3470_cat		pMM26	BglII/SalI	TTTTTGTGAC	TACCGAACTACTCTCCGAACTCT	70	OMM60	TTTTTGTGACCTTACCGAACTACTCTCCGAACTCT
Zg-4675	Ferredox red	pMM27	BglII/SalI	TTTTTGTGAC	TTAGCCAGGAGTTGGCCG	72	OMM62	TTTTTGTGACCTTTAGCCAGGAGTTGGCCG
Zg-4676	Ferredox	pMM28	BglII/SalI	TTTTTGTGAC	TTATTTGGCACTTTTAAACCACACC	72	OMM63	TTTTTGTGACCTTTATTTGGCACTTTTAAACCACACC
Zg-4677	Cyp450	pMM29	BglII/SalI	TTTTTGTGAC	TTATTTGGCACTTTTAAACCACACC	72	OMM64	TTTTTGTGACCTTTTATTTGGCACTTTTAAACCACACC
Pc0002_sp	GH110, S1-15	pMM30	BglII/SalI	TTTTTGTGAC	TTATTTGGCACTTTTAAACCACACC	72	OMM65	TTTTTGTGACCTTTTATTTGGCACTTTTAAACCACACC
Pc0074_sp	GH110, a-Gal	pMM31	BglII/SalI	TTTTTGTGAC	TTAGTTTTAGCTCTTTTAAAGTTGTTG	72	OMM66	TTTTTGTGACCTTTTATTTAGTTTTAGCTCTTTTAAAGTTGTTG
AF20540_sp	GH50	pMM32	BglII/SalI	TTTTTGTGAC	TTAGTTTTAGCTCTTTTAAAGTTGTTG	70	OMM67	TTTTTGTGACCTTTTATTTAGTTTTAGCTCTTTTAAAGTTGTTG
AF20553_sp	GH50	pMM33	BglII/SalI	TTTTTGTGAC	TTAGTTTTAGCTCTTTTAAAGTTGTTG	70	OMM68	TTTTTGTGACCTTTTATTTAGTTTTAGCTCTTTTAAAGTTGTTG
AF20561	Cyp450	pMM34	BglII/SalI	TTTTTGTGAC	TTAGTTTTAGCTCTTTTAAAGTTGTTG	70	OMM69	TTTTTGTGACCTTTTATTTAGTTTTAGCTCTTTTAAAGTTGTTG
AF20562	2Fe-2S ferredox	pMM35	BglII/SalI	TTTTTGTGAC	TTAGTTTTAGCTCTTTTAAAGTTGTTG	70	OMM70	TTTTTGTGACCTTTTATTTAGTTTTAGCTCTTTTAAAGTTGTTG
AF20563	Ferredox red	pMM36	BglII/SalI	TTTTTGTGAC	TTAGTTTTAGCTCTTTTAAAGTTGTTG	72	OMM71	TTTTTGTGACCTTTTATTTAGTTTTAGCTCTTTTAAAGTTGTTG
AF110107_cat	GH82, iotase	pMM37	BglII/SalI	TTTTTGTGAC	TTAATACCACCTTACATCACTACAATT	70	OMM72	TTTTTGTGACCTTAATACCACCTTACATCACTACAATT
AF110107_mod1-2	GH82, iotase	pMM38	BglII/SalI	TTTTTGTGAC	AGCCTCATCAAAGTACTCTCGTTT	70	OMM73	TTTTTGTGACCTTAAGCCTCATCAAAGTACTCTCGTTT
AF110107_mod3-4	GH82, iotase	pMM39	BglII/SalI	TTTTTGTGAC	GTTCACCACCTACAGAGATAGTC	72	OMM74	TTTTTGTGACCTTAGTTCACCACCTACAGAGATAGTC
AF110107_mod3-4-cat	GH82, iotase	pMM40	BglII/SalI	TTTTTGTGAC	TTAATACCACCTTACATCACTACAATT	70	OMM72	TTTTTGTGACCTTAATACCACCTTACATCACTACAATT

2. **Tableau de suivi des étapes du clonage** : La coloration verte indique que l'élément (colonie, PCR,...) a été obtenu, alors qu'une coloration rouge indique un non succès. En orange sont indiquées les anomalies dans les tailles (expliquées par une inversion des numéros des colonies). En jaune est indiqué le nombre de fragments attendu suite à la digestion HincII. Les tailles des inserts et des fragments sont colorées avec une gradation automatique en fonction de la valeur du nombre.

Gène	Plasmide	Taille insert	avec amorces T7	nb cfU	nb clones testés par PCR	nb positifs PCR	% positifs	N°colonie pour miniprep	Nb fragments et taille apres digestion par HincII						Nb miniprep digérées par HincII	Nb positifs confirmés par HincII
Zga_236	pMM1	849	1040	85	10	9	0,9	F1	3	3532	1969	1031			1	1
Zga_3602	pMM2	1914	2105	192	11	3	0,27273	9	4	3532	1838	1434	793	1	1	
Zga_4674	pMM3	1137	1328	185	4	3	0,75	1	3	3532	1900	1388		1	1	
Pca_p0024	pMM4	2394	2585	66	4	1	0,25	4	2	4545	3532			1	1	
Pca_p0027	pMM5	1476	1667	95	4	2	0,5	1	2	3627	3532			1	1	
Pca_p0028	pMM6	1389	1580	117	4	2	0,5	4	3	3532	2509	1031		1	1	
Pca_p0033	pMM7	1587	1778	115	4	1	0,25	3	3	3532	2413	1325		1	1	
Pca_p0034	pMM8	942	1133	63	10	6	0,6	F4	2	3532	3093			1	1	
Pca_p0058	pMM9	2799	2990	96	39	4	0,10256	31	2	4950	3532			1	1	
Pca_p0073	pMM10	2652	2843	104	39	1	0,02564	27	5	3532	1673	1439	1416	275	1	1
Pca_p0075	pMM11	1758	1949	111	4	2	0,5	2	4	3532	1989	997	923	1	1	
Pca_p0076	pMM12	1587	1778	300	4	2	0,5	4	3	3532	1945	1793		1	1	
Pca_p0077	pMM13	2469	2660	116	11	3	0,27273	6	3	3532	2521	2099		1	1	
Afo_20580	pMM14	2292	2483	90	11	1	0,09091	11	5	3532	1602	1596	845	400	1	1
Afo_110093	pMM15	873	1064	53	4	3	0,75	1	2	3532	3024			1	1	
Afo_110095	pMM16	1479	1670	69	4	3	0,75	3	3	3532	2191	1439		1	1	
Afo_110098	pMM17	948	1139	72	38	11?	0,36364	7&8	2	3532	3099			2	2	
Afo_110100	pMM18	1236	1427	74	4	3?	0,75	4	4	3532	2158	1119	110	1	1	
Afo_110102	pMM19	1449	1640	38	11	1?	0,09091	10	3	3532	2416	1184		1	1	
Afo_110113	pMM20	2001	2192	61	37	14	0,37838	4,5,8,12,21,24	3	3532	3169	983		6	5	
							43,0%									
Zga_3464	pMM21	1755	1946	66	4	1		3	3	3532	2713	1193		1	0	
				13	1	0,07692		15						1	1	
Zga_3464	pMM22	1389	1580	41	4	2	0,5	3	3	3532	2713	827		1	1	
Zga_3464	pMM23	393	584	32	4	1	0,25	4	2	3532	2544			1	1	
Zga_3470	pMM24	1437	1628	23	4	2	0,5	1	3	3532	2433	1155		1	1	
Zga_3470	pMM25	465	656	19	4	3	0,75	1	2	3532	2616			1	1	
Zga_3470	pMM26	999	1190	12	4	4	1	1	3	3532	1995	1155		1	1	
Zga_4675	pMM27	1275	1466	45	4	3	0,75	1	2	3532	3426			1	1	
Zga_4676	pMM28	330	521	73	4	1		1	2	3532	2481			1	0	
				13	5	0,38462		11						5	5	
Zga_4677	pMM29	1185	1376	155	4	3		2	4	3532	1671	1326	339	1	0	
								1&3						2	1	
Pca_p0002	pMM30	1518	1709	39	11	2	0,18182	5	4	3532	1573	1471	625	1	1	
Pca_p0074	pMM31	1815	2006	42	38	1	0,02632	14 & 34	2	3966	3532			2	2	
Afo_20540	pMM32	2280	2471	60	11	1		8	3	3532	2960	1471		1	0	
				13	0											
Afo_20553	pMM33	2292	2483	65	38	1bs		2	4	3532	2227	1323	893	1	0	
		1263	1454	18	4	2		4	3	3532	2287	1127		1	0	
Afo_20561	pMM34						0,5	1						1	1	
				14	1	0,07143		11						1	1	
Afo_20562	pMM35	333	524	5	4	4	1	1	2	3532	2484			1	1	
Afo_20563	pMM36	1347	1538	17	4	2	0,5	1	4	3532	1438	1271	789	1	1	
Afo_110107	pMM37	1416	1607	14	4	1	0,25	3	3	3532	2278	1289		1	1	
Afo_110107	pMM38	1038	1229	80	10	7	0,7	F3	2	3532	3189			1	1	
Afo_110107	pMM39	1296	1487	38	4	3	0,75	1	4	3532	1614	1567	266	1	1	
Afo_110107	pMM40	2784	2975	32	32	1bs		3	5	3532	1813	1567	1289	266	1	0
	moy	1496,025	1687,025	74,45			48,2%									37
							45,6%									40
																92,5%

3. Tableau des amorces internes utilisées pour le séquençage

		Séquence des amorces internes	Nb tot pb	Tm
pMM2	F	CGAAGAAGGCGATGCCGGG	19	64
	R	GGTATAGTCGCCATGACCGG	20	64
pMM4	F	CCAGAGGTTAGAAAAGGGCGG	21	66
	R	ACCAGCCCATGTTGGCATCC	20	64
pMM9	F	CTGCAATTTCTGTTATGAATTTGG	24	64
	R	AGCGCCTTTACCAATTA ACTCC	22	64
pMM10	F	CAATAGCAAATTTAAACGCATTGC	24	64
	R	TTCAACAATTGGTAAAACACCCG	23	64
pMM11	F	ACACCACGCTTTAGGAATGGG	21	64
	R	CAAATTCATCACCTTGACCGG	22	64
pMM12	F	GGTCAGATACAGCTACAAGGC	21	64
	R	GTTGCTGGTATATCTAAACCGC	22	64
pMM13	F	TAGATATACGGTAGTTACTACCG	23	64
	R	CGTAGGCTGACTGGTAAGCG	20	64
pMM14	F	CTGGTATGCGTTCGAACCCC	20	64
	R	GTTTAGAGTCTGGATGCTTTTCG	22	64
pMM20	F	GGTATTAGCACTTTACTACTACG	23	64
	R	TGGTACCACATACACTAATCCC	22	64

4. Tableau bilan des séquençages

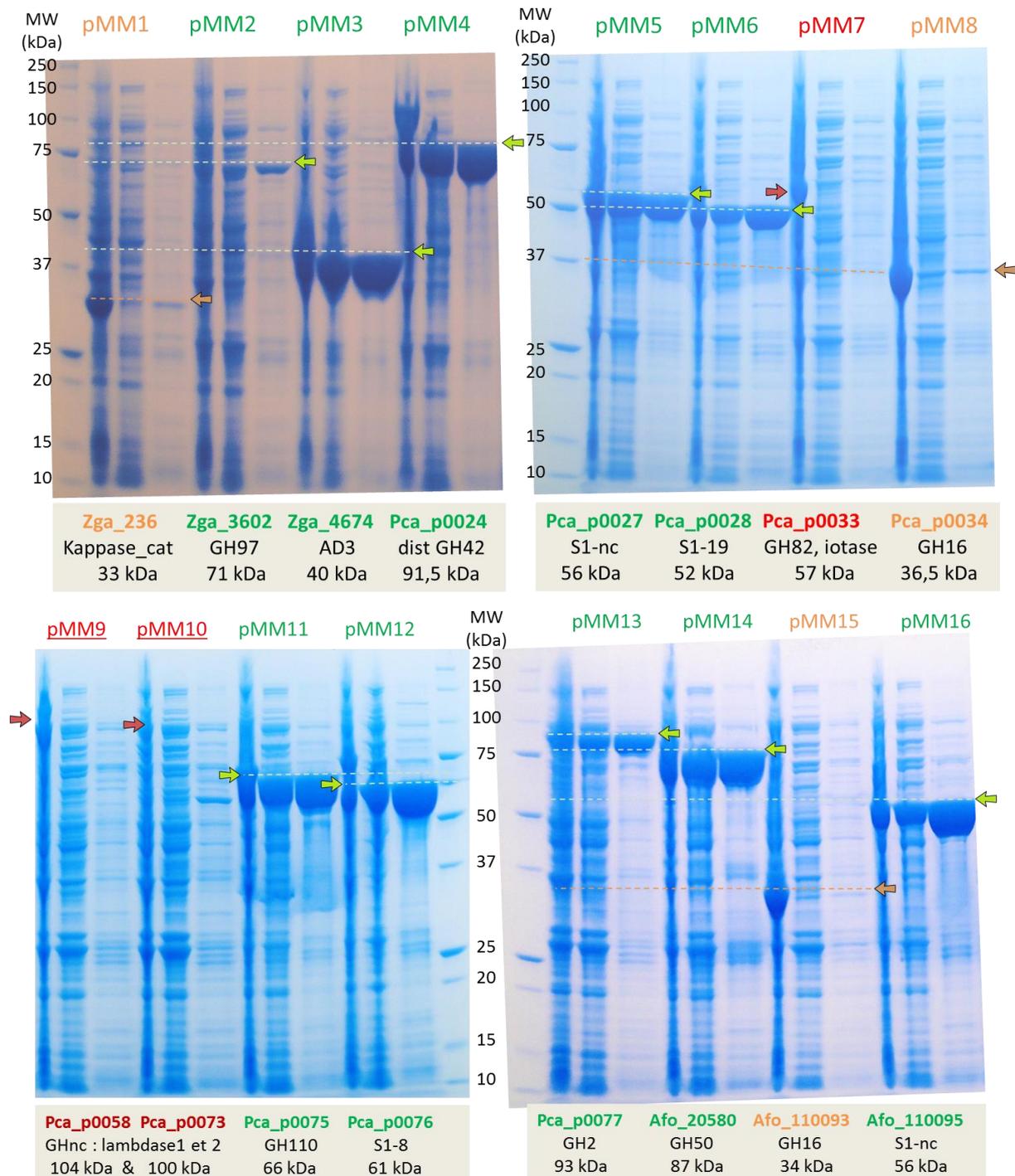
Code couleur similaire au tableau de l'annexe 3.2

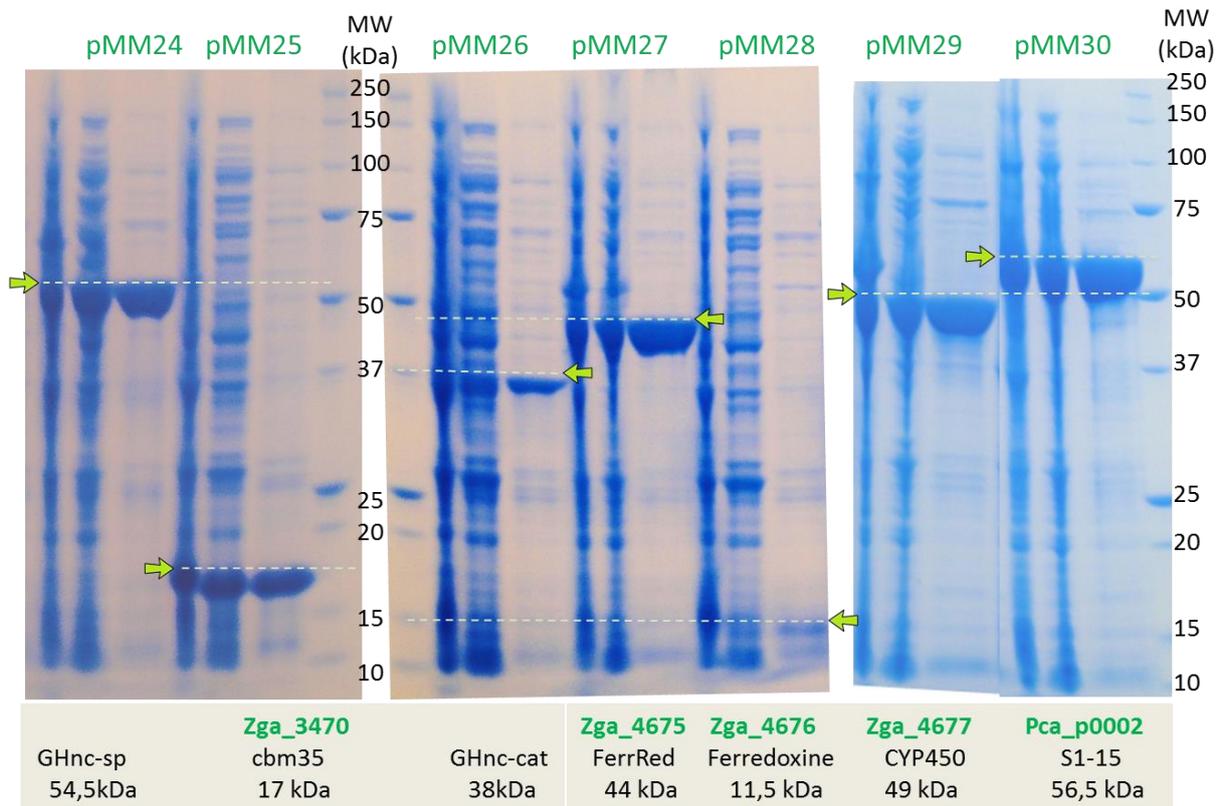
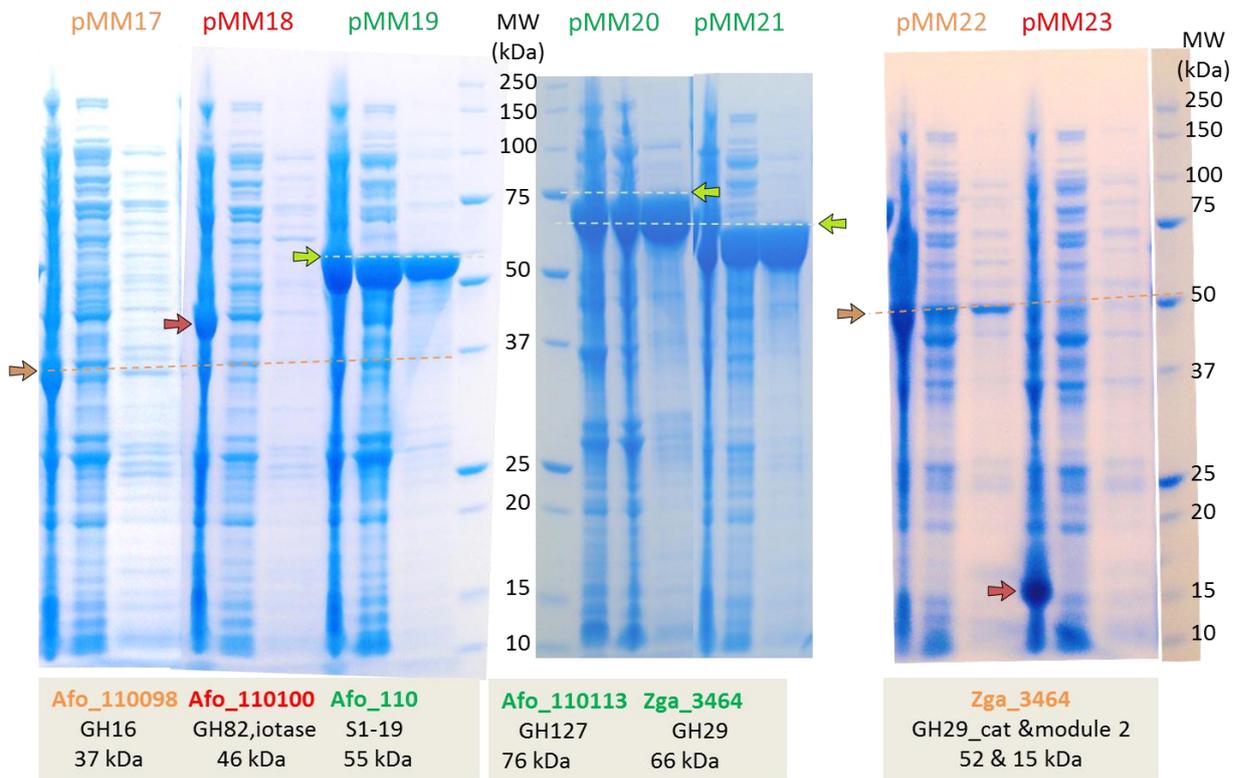
Cl:M48lasm	Taille séquence avec polyhis	Nb bases confirmées par séquençage		Nb pb séquencées
	Nb pb	Forward	Reverse	
1	849	841	910	1751
2	1914	783	623	1406
3	1137	888	898	1786
4	2394	855	960	1815
5	1476	856	840	1696
6	1389	607	789	1396
7	1563	890	783	1673
8	942	858	900	1758
9	2799	888	892	1780
10	2652	885	776	1661
11	1758	878	828	1706
12	1587	647	785	1432
13	2469	909	596	1505
14	2292	576	748	1324
15	873	871	873	1744
16	1490	752	759	1511
17	965	448	736	1184
18	1239	799	897	1696
19	1499	919	884	1803
20	2007	938	961	1899
21	1755	919	851	1770
22	1389	895	875	1770
23	393	393	393	786
24	1437	931	837	1768
25	465	465	465	930
26	999	609	755	1364
27	1275	626	925	1551
28	330	330	330	660
29	1185	985	782	1767
30	1518	845	865	1710
31	1815	952	996	1948
34	1263	950	898	1848
35	333	333	333	666
36	1347	581	971	1552
37	1416	756	864	1620
38	1038	674	656	1330
39	1296	830	876	1706

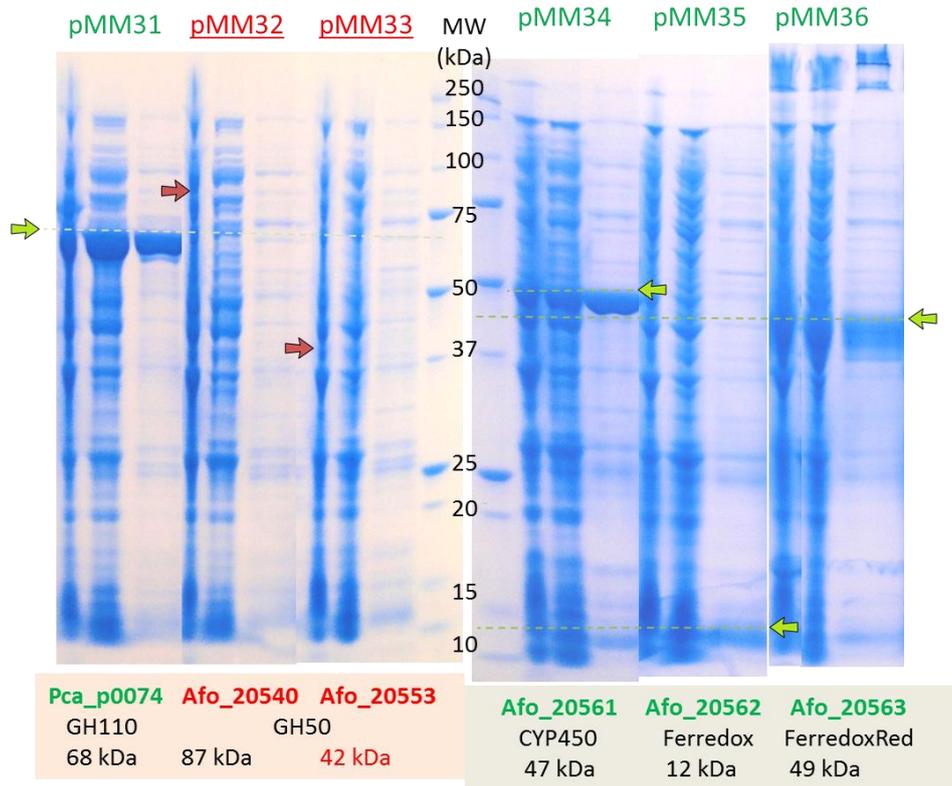
Nb bases confirmées par séquençage avec amorces internes		Nb pb séquencées
Forward	Reverse	
1459	pas séquencé	2082
pas séquencé	1899	2754
1786	1610	3396
1016	1648	2664
1747	1108	2855
1313	1563	2876
1730	1504	3234
1387	1488	2875
1802	1698	3500

Annexe 4 : Résultat des tests d'expression à petite échelle

MW : marqueurs de taille (kDa) ; En vert : cibles exprimées de façon majoritairement soluble ; En orange : cibles exprimées partiellement solubles ; En rouge : cibles exprimées totalement insolubles ou non exprimées (soulignées)







Annexe 5 : Protocoles d'extraction des polysaccharides algaux

1. Extraction de carraghénane de *Chondracanthus acicularis*

- Pour 32 g algues sèches, broyées au blender , 2,6 l H₂O déminéralisée
- Chauffer 4h à 60 °C (sous agitation ou agiter manuellement toute les heures)
- Centrifuger à 7 500 rpm (14 000 g), à 35 °C, pendant 1h
- Centrifuger une nouvelle fois le surnageant 1h pour se débarrasser d'un maximum de débris cellulaires
- Précipitation du surnageant dans l'isopropanol (2,8 l par litre de surnageant) + KCl (solution à 140 g/l, ajouter 11,1 ml par litre d'isopropanol), dans un erlen de 5l à col large, avec un agitateur à ailettes fixé sur une potence.
Ajouter progressivement le surnageant pendant que l'agitateur tourne dans l'isopropanol + KCl.

2. Préparation des AIR d'algues

Les AIR (*Alcool Insoluble Residues*) de *Palmaria palmata* et de différentes algues brunes ont été réalisés suivant le protocole fourni par Cécile Hervé :

- 50 g d'algues séchées et broyées le plus finement possible (blender, mortier et azote liquide, broyeur à billes,...)
- Reprendre deux fois de suite dans 800 ml d'éthanol 70 % et filtrer sur Büchner (frité N°3 ou 4 additionné d'un filet en nylon pour protéger le fritté)
- Reprendre deux fois de suite dans 800 ml d'éthanol 96 % et filtrer sur Büchner
- Sécher avec 100-200 ml d'acétone

3. Extraction des fucoïdanes d'algues brunes

Protocole mis au point par Mikkel Schultz-Johansen et suivi durant cette thèse pour réaliser les tests d'activité enzymatiques à partir des lysats bactériens clarifiés contenant les protéines recombinantes à tester.

- Répartir 20 mg d'AIR dans des eppendorfs de 2 ml
- Ajouter 1 ml de CaCl₂ à 2 % dans de l'eau mQ
- Agiter 4 h à 37 °C avec les tubes à l'horizontal pour maximiser l'agitation
- Centrifuger et récupérer le surnageant (environ 800 µl) contenant la fraction de fucoïdanes
- Filtrer sur membrane 10 kDa (spin filtrer) pour enlever les oligosaccharides. Pour cela :
 - déposer 500 µl de surnageant sur le filtre,
 - centrifuger à 13 000g pendant 10 min
 - rincer avec 450 µl de tampon (20 mM Tris-HCl pH8 par exemple)
 - centrifuger à nouveau et récupérer le rétentat concentré 10 fois (environ 50 µl)

Les tests d'activité enzymatiques sont réalisés par dilution au 10^{ème} des extraits précédents (2 µl +18 µl de tampon de type Tris) et ajout de 10 µl du surnageant d'un lysat bactérien clarifié par centrifugation. Les 30 µl finaux sont incubés 16 h à la température de la pièce. 10 µl sont ensuite mélangés avec 5 µl de colorant pour C-PAGE (sucrose +rouge phénol) puis migration sur gel C-PAGE.

Annexe 6 : Propriétés associées aux composés extraits des algues

En jaune : catégories d'applications principales ; En verts : extraits possédant les propriétés associées à des applications générales (vert foncé) ou plus spécifiques (vert clair)

		Protecteur Cardio vasculaire	Anti hypertenseur	Hypo cholestérol émiant	Anti arythmie	Anti athéro sclérose	Anti hyperlipidémiant	Hypo glycé miant	Activité anti-lipase	Hémo statique	Anti coagulant	Anti aggrégation plaquettaire	ACE inhibiteur	Fibrino lytique	Anti obésité	Anti diabète	Immuno stimulant	Prévention Troubles Autoimmuns
Minéraux	Iode																	
Protéines	Phycobilliprotéines																	
	Lectines																	
	Glycoprotéines																	
Acides aminés	Taurine																	
	Laminine																	
	Kainoïds																	
	MMAs																	
Peptides bioactifs endogènes	Carnosine																	
	Glutathione																	
	SECMA 1																	
	Analogues GABA																	
	Kahalalide F																	
	Almazole D																	
	Galaxamide																	
Dipeptides																		
Hydrolysat protéique																		
Acides gras	LN, SA, EPA																	
	Stérols saturés et insaturés																	
Composés phénolés	Phlorotannins																	
	Diterpènes																	
	Sesquiterpènes																	
	Meroditerpènes																	
	Composés hallogénés																	
Pigments	Caroténoïdes: Fucoxanthin, Xantophylles, β-carotène, lutéine																	
	Chlorophylles																	
	Phycobilliprotéines																	
Polysaccharides	Agars																	
	Carraghénanes																	
	Alginate																	
	Fucoïdanes																	
	Laminarine																	
Ulvanes																		
Polyols	Mannitol																	

		Prévention Arthrite	Anti oxydant	Pro Oxydant	Décontractant musculaire	Anti-inflammatoire	Anti allergique	Héliciteur végétal	Anti "Plant pathogène"	Insecticide	Anthelmintique	Anti fouling	Anti brouteur	Anti invertébré	Anti larvaire	Anti algal	Contre troubles digestifs	Stimule réparation tissulaire	Ophthalmoprotecteur
Minéraux	Iode																		
Protéines	Phycobilliprotéines																		
	Lectines																		
	Glycoprotéines																		
Acides aminés	Taurine																		
	Laminine																		
	Kainoïds																		
	MMAs																		
Peptides bioactifs endogènes	Carnosine																		
	Gluthatione																		
	SECMA 1																		
	Analogues GABA																		
	Kahalalide F																		
	Almazole D																		
	Galaxamide																		
Dipeptides																			
Hydrolysat protéique																			
Acides gras	LN, SA, EPA																		
	Stérols saturés et insaturés																		
Composés phénolés	Phlorotannins																		
	Diterpènes																		
	Sesquiterpènes																		
	Meroditerpènes																		
	Composés hallogénés																		
Pigments	Caroténoïdes: Fucoxanthin, Xanthophylles, β-carotène, lutéine																		
	Chlorophylles																		
	Phycobilliprotéines																		
Polysaccharides	Agars																		
	Carraghénanes																		
	Alginates																		
	Fucoïdanes																		
	Laminarine																		
	Ulvanes																		
Polyols	Mannitol																		

		Anti dépresseur	Traitement troubles nerveux	Neuro protecteur	Traitement déficience thyroïdienne	Anti photoa ging	Anti UV	Détox ifiant	Hépatoprotecteur	Anti cancer	Antitumoral	Anti angiogénique	Mitogénique	Cytotoxique	Anti prolifératif	Anti viral	Anti VIH	Anti fongique	Anti bactérien	Anti adhésion	Anti métalloprotéases	Anti douleur	Osmolite	Colorant naturel	Bio technologies
Minéraux	Iode																								
Protéines	Phycobilliprotéines																								
	Lectines																								
	Glycoprotéines																								
Acides aminés	Taurine																								
	Laminine																								
	Kainoids																								
	MMAs																								
Peptides bioactifs endogènes	Carnosine																								
	Gluthatione																								
	SECMA 1																								
	Analogues GABA																								
	Kahalalide F																								
	Almazole D																								
	Galaxamide																								
Dipeptides																									
Hydrolysat protéique																									
Acides gras	LN, SA, EPA																								
	Stérols saturés et insaturés																								
	Phlorotannins																								
Composés phénolés	Diterpènes																								
	Sesquiterpènes																								
	Meroditerpènes																								
	Composés hallogénés																								
	Caroténoïdes: Fucoxanthin, Xanthophylles, β-carotène, lutéine																								
Pigments	Chlorophylles																								
	Phycobilliprotéines																								
	Agars																								
Polysaccharides	Carraghénanes																								
	Alginates																								
	Fucoïdanes																								
	Laminarine																								
	Ulvanes																								
	Ulvanes																								
Polyols	Mannitol																								

Bibliographie

Bibliographie I

- Abdel-Fattah AF, Hussein MMD, Fouad ST. 1978. Carbohydrates of the brown seaweed *Dictyota dichotoma*. *Phytochemistry* **17**: 741–743.
- Adrien A, Bonnet A, Dufour D, Baudouin S, Maugard T, Bridiau N. 2017. Pilot production of ulvans from *Ulva* sp. and their effects on hyaluronan and collagen production in cultured dermal fibroblasts. *Carbohydr Polym* **157**: 1306–1314.
- Agirre J, Iglesias-Fernández J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. 2015. Privateer: software for the conformational validation of carbohydrate structures. *Nat Struct Mol Biol* **22**: 833–834.
- Ahmad M, Hirz M, Pichler H. 2014. Protein expression in *Pichia pastoris* : recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* **98**: 5301–5317.
- Alva V, Nam S-Z, Söding J, Lupas AN. 2016. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* **44**: gkw348.
- Alves A, Sousa RA, Reis RL. 2013. A practical perspective on ulvan extracted from green algae. *J Appl Phycol* **25**: 407–424.
- Anderson KL, Salyers AA. 1989a. Biochemical-Evidence that Starch Breakdown by *Bacteroides-Thetaiotaomicron* Involves Outer-Membrane Starch-Binding Sites and Periplasmic Starch-Degrading Enzymes. *J Bacteriol* **171**: 3192–3198.
- Anderson KL, Salyers AA. 1989b. Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. *J Bacteriol* **171**: 3199–3204.
- Appel MJ, Bertozzi CR. 2015. Formylglycine, a Post-Translationally Generated Residue with Unique Catalytic Capabilities and Biotechnology Applications Mason. *ACS Chem Biol* **10**: 72–84.
- Arnosti C. 2014. Patterns of Microbially Driven Carbon Cycling in the Ocean: Links between Extracellular Enzymes and Microbial Communities. *Adv Oceanogr* **2014**: 1–12.
- Arzel P. 1994. L'exploitation des algues en Bretagne. *Etud Rurales* **133**: 113–126.
- Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B. 2012. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol Biol* **12**: 186.
- Azam F, Malfatti F. 2007. Microbial structuring of marine ecosystems. *Nat Rev Microbiol* **5**: 782–91.
- Aziz A, Poinssot B, Daire X, Adrian M, Bézier A, Lambert B, Pugin A. 2003. Laminarin Elicits Defense Responses in Grapevine and Induces Protection Against *Botrytis cinerea* and *Plasmopara viticola*. *Mol Plant-Microbe Interact* **16**: 1118–1128.
- Barbeyron T. 1992. Caractérisation moléculaire d'une sulfatase, d'une kappa-carraghénase et d'une iota-carraghénase chez deux bactéries marines.
- Barbeyron T, Gerard A, Potin P, Henrissat B, Kloareg B. 1998. The Kappa-Carrageenase of the

- Marine Bacterium *Cytophaga drobachiensis*. Structural and Phylogenetic Relationships Within Family-16 Glycoside Hydrolases. *Mol Biol Evol* **15**: 528–537.
- Barbeyron T, Henrissat B, Kloareg B. 1994. The gene encoding the kappa-carrageenase of *Alteromonas carrageenovora* is related to β -1,3-1,4-glucanases. *Gene* **139**: 105–109.
- 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.
- Barbeyron T, Michel G, Potin P, Henrissat B, Kloareg B. 2000. ι -Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of κ -carrageenases. *J Biol Chem* **275**: 35499–35505.
- Barbeyron T, Thomas F, Barbe V, Teeling H, Schenowitz C, Dossat C, Goesmann A, Leblanc C, Oliver Glöckner F, Czjzek M, *et al.* 2016. Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* DsijT. *Environ Microbiol* **18**: 4610–4627.
- Bellion C, Hamer GK, Yape W. 1982. The degradation of *Eucheuma spinosum* and *Eucheuma cottonii* carrageenans by ι -carrageenases and κ -carrageenases from marine bacteria. *Can J Microbiol* **28**: 874–880.
- Bemiller JN. 2009. One hundred years of commercial food carbohydrates in the United States. *J Agric Food Chem* **57**: 8125–8129.
- Berteau O, Mulloy B. 2003. Sulfated fucans, fresh perspectives: Structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. *Glycobiology* **13**: 29–40.
- Bertness MD, Leonard GH, Levine JM, Schmidt PR, Ingraham AO. 1999. Testing the Relative Contribution of Positive and Negative Interactions in Rocky Intertidal Communities. *Ecology* **80**: 2711–2726.
- Bhatnagar M, Bhatnagar A. 2015. Wound Dressings from Algal Polymers. *Mar Algae Extr Process Prod Appl* **2–2**: 523–556.
- Bhattacharyya S, Liu H, Zhang Z, Jam M, Dudeja PK, Michel G, Linhardt RJ, Tobacman JK. 2010. Carrageenan-induced innate immune response is modified by enzymes that hydrolyze distinct galactosidic bonds. *J Nutr Biochem* **21**: 906–13.
- Bilan MI, Grachev AA, Shashkov AS, Kelly M, Sanderson CJ, Nifantiev NE, Usov AI. 2010. Further studies on the composition and structure of a fucoidan preparation from the brown alga *Saccharina latissima*. *Carbohydr Res* **345**: 2038–2047.
- Bissaro B, Monsan P, Fauré R, O'Donohue MJ. 2015. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochem J* **467**: 17–35.
- Bixler HJ, Porse H. 2010. A decade of change in the seaweed hydrocolloids industry. *J Appl Phycol* **23**: 321–335.
- Bjursell MK, Martens EC, Gordon JI. 2006. Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J Biol Chem* **281**: 36269–36279.
- Black GW, Rixon JE, Clarke JH, Hazlewood GP, Theodorou MK, Morris P, Gilbert HJ. 1996. Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. *Biochem J* **319**: 515–520.

- Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M. 2007. Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. *PLoS One* **2**.
- Bodian MY, Lafontaine N, Matard M, Mussio I, Rusig AM. 2013. Evaluation of the in vitro methods for micropropagation of *Chondracanthus acicularis* (Roth) Fredericq (Gigartinales, Rhodophyta): Tissue culture and production of protoplasts. *J Appl Phycol* **25**: 1835–1845.
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**: 769–81.
- Boyen C, Kloareg B, Polnefuller M, Gibor A. 1990. Preparation of alginate lyases from marine mollusks for protoplast isolation in brown-algae. *Phycologia* **29**: 173–181.
- Buck CB, Thompson CD, Roberts JN, Müller M, Lowy DR, Schiller JT. 2006. Carrageenan is a potent inhibitor of papillomavirus infection. *PLoS Pathog* **2**: 0671–0680.
- Burki F, Alegado RA, King N, Schlacht A, Herman EK, Mary J, Cavalier-smith T. 2014. The Eukaryotic Tree of Life from a Global Phylogenomic Perspective. *Cold Spring Harb Perspect Biol* **6**: a016147.
- Cao H, Walton JD, Brumm P, Phillips GN. 2014. Structure and substrate specificity of a eukaryotic fucosidase from *Fusarium graminearum*. *J Biol Chem* **289**: 25624–38.
- Cardenas D. 2013. Let not thy food be confused with thy medicine: The Hippocratic misquotation. *ESPEN J* **8**: e260–e262.
- Chauhan PS, Saxena A. 2016. Bacterial carrageenases: an overview of production and biotechnological applications. *3 Biotech* **6**.
- Cheatham MR, Rouse MN, Esker PD, Ignacio S, Pradel W, Raymundo R, Sparks AH, Forbes GA, Gordon TR, Garrett KA. 2009. Mini-Review Beyond Yield: Plant Disease in the Context of Ecosystem Services. *Phytopathology* **99**: 1228–1236.
- Chiovitti A, Kraft GT, Bacic A, Craik DJ, Liao M-L. 2008. Kappa-/beta-carrageenans from Australian red algae of the Acrotylaceae (Gigartinales, Rhodophyta). *Phycologia* **47**: 35–40.
- Cobucci-Ponzano, B., Conte, F., Bedini, E., Corsaro, M. M., Parrilli, M., Sulzenbacher, G., ... & Moracci, M. (2009). β -Glycosyl azides as substrates for α -glycosynthases: Preparation of efficient α -L-fucosynthases. *Chemistry & biology* **16**(10): 1097-1108.
- Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury J-M, Badger JH, *et al.* 2010. The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature* **465**: 617–621.
- Collén J, Porcel B, Carré W, Ball SG, Chaparro C, Tonon T, Barbeyron T, Michel G, Noel B, Valentin K, *et al.* 2013. Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc Natl Acad Sci U S A* **110**: 5247–52.
- Colombini I, Chelazzi L. 2003. Influence of marine allochthonous input on sandy beach communities. *Oceanogr Mar Biol an Annu Rev* **41**: 115–159.
- Correc G, Barabanova A, Tuvikene R, Truus K, Yermak I, Helbert W. 2012. Comparison of the structures of hybrid κ -/ β -carrageenans extracted from *Furcellaria lumbricalis* and *Tichocarpus crinitus*. *Carbohydr Polym* **88**: 31–36.

- 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.
- Courtois J. 2009. Oligosaccharides from land plants and algae: production and applications in therapeutics and biotechnology. *Curr Opin Microbiol* **12**: 261–273.
- Craigie JS. 1990. Cell Walls. In *Biology of the red algae*, New York, pp. 221–257, Cambridge University Press.
- Craigie JS. 2011. Seaweed extract stimuli in plant science and agriculture. *J Appl Phycol* **23**: 371–393.
- Crawford MA. 1992. The role of dietary fatty acids in biology: Their place in the evolution of the human brain. *Nutr Rev* **50**: 2–11.
- Cunha L, Grenha A. 2016. Sulfated seaweed polysaccharides as multifunctional materials in drug delivery applications. *Mar Drugs* **14**.
- Cunnane SC, Crawford MA. 2014. Energetic and nutritional constraints on infant brain development: Implications for brain expansion during human evolution. *J Hum Evol* **77**: 88–98.
- Czjzek M, Cicek M, Zamboni V, Burmeister WP, Bevan DR, Henrissat B, Esen A. 2001. Crystal structure of a monocotyledon (maize ZMGLu1) β -glucosidase and a model of its complex with p-nitrophenyl β -D-thioglucoside. *Biochem* **354**: 37–46.
- D'Elia JN, Salyers AA. 1996. Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. *J Bacteriol* **178**: 7180–7186.
- Dabin J. 2008. Etude structurale et fonctionnelle des polysaccharidases de *Rhodospirellula baltica*.
- David B, Irague R, Jouanneau D, Daligault F, Czjzek M, Sanejouand YH, Tellier C. 2017. Internal Water Dynamics Control the Transglycosylation/Hydrolysis Balance in the Agarase (AgaD) of *Zobellia galactanivorans*. *ACS Catal* **7**: 3357–3367.
- David S. 1995. *Chimie moléculaire et supramoléculaire des sucres : Introduction chimique aux glycosciences*. InterEditions.
- Davies G, Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**: 853–9.
- Davies GJ, Sinnot ML. 2008. Sorting the diverse. *Biochem J* 26–32.
- Davies GJ, Wilson KS, Henrissat B. 1997. Glycosyl Hydrolases. *Biochem J Lett* **321**: 557–559.
- De Ley J, Doudoroff M. 1957. The metabolism of D-galactose in *Pseudomonas saccharophila*. *J Biol Chem* **227**: 745–757.
- Deniaud-Bouët E, Hardouin K, Potin P, Kloareg B, Hervé C. 2017. A review about brown algal cell walls and fucose-containing sulfated polysaccharides: cell wall context, biomedical properties and key research challenges. *Carbohydr Polym* **175**: 395–408.
- Deniaud-Bouët E, Kervarec N, Michel G, Tonon T, Kloareg B, Hervé C. 2014. Chemical and enzymatic fractionation of cell walls from Fucales: Insights into the structure of the extracellular matrix of brown algae. *Ann Bot* **114**: 1203–1216.
- Dillehay TD, Ramirez C, Pino M, Collins MB, Rossen J, Pino-Navarro JD. 2008. Monte Verde: Seaweed, Food, Medicine, and the Peopling of South America. *Science (80-)* **320**: 784–

786.

- Din N, Gilkes NR, Tekant B, Miller RC, Warren RAJ, Kilburn DG. 1991. Non hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. *Nat Biotechnol* **9**: 1096–1099.
- Divne C, Ståhlberg J, Teeri TT, Jones T a. 1998. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *J Mol Biol* **275**: 309–325.
- Duggins DO, Simenstad CA, Estes JA. 1989. Magnification of Secondary Production by Kelp Detritus in Coastal Marine Ecosystems. *Science (80-)* **245**: 170–173.
- Egan S, Fernandes ND, Kumar V, Gardiner M, Thomas T. 2014. Bacterial pathogens, virulence mechanism and host defence in marine macroalgae. *Environ Microbiol* **16**: 925–938.
- Egan S, Harder T, Burke C, Steinberg P, Kjelleberg S, Thomas T. 2013. The seaweed holobiont: Understanding seaweed-bacteria interactions. *FEMS Microbiol Rev* **37**: 462–476.
- Elyakova LA, Zvyagintseva TN. 1974. A study of the laminarins of some far-eastern, brown seaweeds. *Carbohydr Res* **34**: 241–248.
- Emsley P, Cowtan K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr* **60**: 2126–2132.
- Evans FD, Critchley AT. 2014. Seaweeds for animal production use. *J Appl Phycol* **26**: 891–899.
- Falshaw R, Furneaux RH, Wong H, Liao M, Bacic A, Chandkrachang S. 1996. Structural analysis of carrageenans from Burmese and Thai samples of *Catenella nipa* Zanardini. *Carbohydr Res* **285**: 81–98.
- Fernandez E, Lopez D, Mijangos C, Duskova-Smrckova M, Ilavsky M, Dusek K. 2007. Rheological and Thermal Properties of Agarose Aqueous Solutions and Hydrogels. *J Polym Sci Part B Polym Phys* **46**: 322–328.
- Ficko-Blean E, Hervé C, Michel G. 2015. Sweet and sour sugars from the sea: the biosynthesis and remodeling of sulfated cell wall polysaccharides from marine macroalgae. *Perspect Phycol* **2**: 51–64.
- Ficko-blean E, Préchoux A, Thomas F, Rochat T, Larocque R. Carrageenan catabolism is conferred by a complex regulon in marine heterotrophic bacteria. 1–38.
- Field CB. 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science (80-)* **281**: 237–240.
- Fleurence J, Massiani L, Guyader O, Mabeau S. 1995. Use of enzymatic cell wall degradation for improvement of protein extraction from *Chondrus crispus*, *Gracilaria verrucosa* and *Palmaria palmata*. *J Appl Phycol* **7**: 393–397.
- Fleurence J, Levine I. 2016. *Seaweed in Health and Disease Prevention*. Academic Press.
- Frangoudes K. 2011. Seaweeds fisheries management in France Japan, Chile and Norway. *Cah Biol Mar* **52**: 517–525.
- Frangoudes K, Garineaud C. 2015. Governability of kelp forest small-scale harvesting in Iroise sea, France. In *Interactive Governance for Small-Scale Fisheries*, Vol. 5 of, pp. 435–456.
- Frantz C, Stewart KM, Weaver VM. 2010. The extracellular matrix at a glance. *J Cell Sci* **123**: 4195–4200.

- Freitas AC, Rodrigues D, Carvalho AP, Pereira L, Panteleitchouk T, Gomes AM, Duarte AC. 2015. Marine functional foods. In *Springer Handbook of Marine Biotechnology*, pp. 969–994, Springer Berlin Heidelberg, Berlin, Heidelberg.
- Freshour G, Bonin CP, Reiter W-D, Albersheim P, Darvill AG, Hahn MG. 2003. Distribution of fucose-containing xyloglucans in cell walls of the mur1 mutant of *Arabidopsis*. *Plant Physiol* **131**: 1602–12.
- Fu XT, Kim SM. 2010. Agarase: Review of major sources, categories, purification method, enzyme characteristics and applications. *Mar Drugs* **8**: 200–218.
- Fusetti F, Von Moeller H, Houston D, Rozeboom HJ, Dijkstra BW, Boot RG, Aerts JMFG, Van Aalten DMF. 2002. Structure of human chitotriosidase: Implications for specific inhibitor design and function of mammalian chitinase-like lectins. *J Biol Chem* **277**: 25537–25544.
- Gachon CMM, Sime-Ngando T, Strittmatter M, Chambouvet A, Kim GH. 2010. Algal diseases: Spotlight on a black box. *Trends Plant Sci* **15**: 633–640.
- Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9**: 747–755.
- Gattuso J-P, Frankignoulle M, Wollast R. 1998. Carbon and carbonate metabolism in coastal aquatic ecosystems. *Annu Rev Ecol Syst* **29**: 405–434.
- Gauthier A, Trouvelot S, Kelloniemi J, Frettinger P, Wendehenne D, Daire X, Joubert JM, Ferrarini A, Delledonne M, Flors V, *et al.* 2014. The sulfated laminarin triggers a stress transcriptome before priming the SA- And ROS-dependent defenses during Grapevine's induced resistance against *Plasmopara viticola*. *PLoS One* **9**.
- Genicot-Joncour S, Poinas A, Richard O, Potin P, Rudolph B, Kloareg B, Helbert W. 2009. The Cyclization of the 3,6-Anhydro-Galactose Ring of α -Carrageenan Is Catalyzed by Two D-Galactose-2,6-Sulfurylases in the Red Alga *Chondrus crispus*. *Plant Physiol* **151**: 1609–1616.
- Gensollen T, Iyer SS, Kasper DL, Blumberg RS. 2016. How colonization by microbiota in early life shapes the immune system. *Science (80-)* **352**: 539–544.
- Gilad R, Rabinovich L, Yaron S, Bayer EA, Lamed R, Gilbert HJ, Shoham Y. 2003. Cell, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *J Bacteriol* **185**: 391–398.
- Gilbert HJ, Knox JP, Boraston AB. 2013. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol* **23**: 669–77.
- Glasner JD, Yang CH, Reverchon S, Hugouvieux-Cotte-Pattat N, Condemine G, Bohin JP, van Gijsegem F, Yang S, Franza T, Expert D, *et al.* 2011. Genome sequence of the plant-pathogenic bacterium *Dickeya dadantii* 3937. *J Bacteriol* **193**: 2076–2077.
- Glasson CRK, Sims IM, Carnachan SM, de Nys R, Magnusson M. 2017. A cascading biorefinery process targeting sulfated polysaccharides (ulvan) from *Ulva ohnoi*. *Algal Res* 1–9.
- Glicksman M. 1987. Utilization of seaweed hydrocolloids in the food industry. *Hydrobiologia* **151–152**: 31–47.
- Godoy AS, Camilo CM, Kadowaki MA, Muniz H dos S, Espirito Santo M, Murakami MT, Nascimento AS, Polikarpov I. 2016. Crystal structure of β 1→6-galactosidase from

- Bifidobacterium bifidum S17: trimeric architecture, molecular determinants of the enzymatic activity and its inhibition by α -galactose. *FEBS J* **283**: 4097–4112.
- Gold P. 1992. Use of a novel agarose gel-digesting enzyme for easy and rapid purification of PCR-amplified DNA for sequencing. *Biotechniques* **13**: 132–4.
- Grant GT, Morris ER, Rees DA, Smith PJ., Thom D. 1973. Biological interactions between polysaccharides and divalent cations: the egg-box model. *FEBS Lett* **32**: 195–198.
- Greer CW, Yaphe W. 1984a. Characterization of Hybrid (Beta-Kappa-Gamma) Carrageenan from *Eucheuma gelatinae* J. Agardh (Rhodophyta, Solieriaceae) Using Carrageenases, Infrared and ¹³C-Nuclear Magnetic Resonance Spectroscopy. *Bot Mar* **27**: 473–478.
- Greer CW, Yaphe W. 1984b. Purification and properties of ι -carrageenase from a marine bacterium. *Can J Microbiol* **30**: 1500–1506.
- Groisillier A, Hervé C, Jeudy A, Rebuffet E, Pluchon PF, Chevotot Y, Flament D, Geslin C, Morgado IM, Power D, *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.
- Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H. 2017. Polysaccharide Utilization Loci: Fuelling microbial communities. *J Bacteriol* JB-00860.
- Guibet M, Colin S, Barbeyron T, Genicot S, Kloareg B, Michel G, Helbert W. 2007. Degradation of λ -carrageenan by *Pseudoalteromonas carrageenovora* λ -carrageenase: a new family of glycoside hydrolases unrelated to κ - and ι -carrageenases. *Biochem J* **404**: 105–114.
- Guillén D, Sánchez S, Rodríguez-Sanoja R. 2010. Carbohydrate-binding domains: Multiplicity of biological roles. *Appl Microbiol Biotechnol* **85**: 1241–1249.
- Guillotin L, Lafite P, Daniellou R. 2014. Unraveling the substrate recognition mechanism and specificity of the unusual glycosyl hydrolase family 29 BT2192 from *Bacteroides thetaiotaomicron*. *Biochemistry* **53**: 1447–1455.
- Hahn T, Lang S, Ulber R, Muffler K. 2012. Novel procedures for the extraction of fucoidan from brown algae. *Process Biochem* **47**: 1691–1698.
- Hall T, Biosciences I, Carlsbad C. 2011. BioEdit: An important software for molecular biology. *GERF Bull Biosci* **2**: 60–61.
- Hansen JE, Packard JE, Doyle WT. 1981. *Mariculture of red Seaweeds*. Report No. ed. California Sea Grant. University of California System (USA).
- Hardouin K, Burlot AS, Umami A, Tanniou A, Stiger-Pouvreau V, Widowati I, Bedoux G, Bourgougnon N. 2014. Biochemical and antiviral activities of enzymatic hydrolysates from different invasive French seaweeds. *J Appl Phycol* **26**: 1029–1042.
- Harnedy PA, Fitzgerald RJ. 2011. Bioactive proteins, peptides, and amino acids from macroalgae. *J Phycol* **47**: 218–232.
- Haroun-Bouhedja F, Ellouali M, Sinquin C, Boisson-Vidal C. 2000. Relationship between sulfate groups and biological activities of fucans. *Thromb Res* **100**: 453–459.
- Harris HMB, Duncan SH, P. Scott K, Martin JC, O. Sheridan P, J. Flint H, Browne HP, Bernalier-Donadille A, Lawley TD, O'Toole PW. 2016. Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. *Microb Genomics* **2**.

- Hartney SL, Mazurier S, Kidarsa TA, Quecine MC, Lemanceau P, Loper JE. 2011. TonB-dependent outer-membrane proteins and siderophore utilization in *Pseudomonas fluorescens* Pf-5. *BioMetals* **24**: 193–213.
- Hatada Y, Ohta Y, Horikoshi K. 2006. Hyperproduction and application of α -agarase to enzymatic enhancement of antioxidant activity of porphyrin. *J Agric Food Chem* **54**: 9895–9900.
- 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**: 908–12.
- Hehemann J-H, Correc G, Thomas F, Bernard T, Barbeyron T, Jam M, Helbert W, Michel G, Czjzek M. 2012a. Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium *Zobellia galactanivorans*. *J Biol Chem* **287**: 30571–84.
- Hehemann J-H, Kelly AG, Pudlo NA, Martens EC, Boraston AB. 2012b. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc Natl Acad Sci U S A* **109**: 19786–91.
- Hendrickson W, Horton JR, LeMaster DM. 1990. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**: 1665–1672.
- Henrissat B, Callebaut I. 1995. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci U S A* **92**: 7090–7094.
- Henrissat B, Davies G. 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol* **7**: 637–644.
- Henrissat B, Vegetales M, Grenoble F. 1991. A classification of glycosyl hydrolases based sequence similarities amino acid. *Biochem J* **280**: 309–316.
- Hildebrand A, Remmert M, Biegert A, Soding J. 2009. Fast and accurate automatic structure prediction with HHpred. *Proteins Struct Funct Bioinforma* **77**: 128–132.
- Holdt SL, Kraan S. 2011. Bioactive compounds in seaweed: Functional food applications and legislation. *J Appl Phycol* **23**: 543–597.
- Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sorlie M, Synstad B, Vriend G, Varum KM, Eijsink VGH, Sørli M, et al. 2006a. Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci* **103**: 18089–18094.
- Horn SJ, Sørbotten A, Synstad B, Sikorski P, Sørli M, Vårum KM, Eijsink VGH. 2006b. Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*. *FEBS J* **273**: 491–503.
- Hornell J. 1946. *Water Transport: origins and early evolution*. ed. Cambridge University Press.
- Hrmova M, Farkas V, Harvey AJ, Lahnstein J, Wischmann B, Kaewthai N, Ezcurra I, Teeri TT, Fincher GB. 2009. Substrate specificity and catalytic mechanism of a xyloglucan xyloglucosyl transferase HvXET6 from barley (*Hordeum vulgare* L.). *FEBS J* **276**: 437–456.
- Huang W, Boju L, Tkalec L, Su H, Yang HO, Gunay NS, Linhardt RJ, Yeong SK, Matte A, Cygler M. 2001. Active site of chondroitin AC lyase revealed by the structure of enzyme - Oligosaccharide complexes and mutagenesis. *Biochemistry* **40**: 2359–2372.

- Iwamoto M, Kurachi M, Nakashima T, Kim D, Yamaguchi K, Oda T, Iwamoto Y, Muramatsu T. 2005. Structure-activity relationship of alginate oligosaccharides in the induction of cytokine production from RAW264.7 cells. *FEBS Lett* **579**: 4423–4429.
- Jackson P. 1991. Polyacrylamide gel electrophoresis of reducing saccharides labeled with the fluorophore 2-aminocridone: subpicomolar detection using an imaging system based on a cooled charge-coupled device. *Anal Biochem* **196**: 238–244.
- Jackson P. 1990. The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphtalene-1,3,6-trisulphonic acid. *Biochem J* **270**: 705–713.
- Jeon YJ, Wijesinghe WAJP, Kim SK. 2011. Enzyme-assisted Extraction and Recovery of Bioactive Components from Seaweeds. In *Handbook of Marine Macroalgae: Biotechnology and Applied Phycology*, pp. 221–228.
- Jiao G, Yu G, Zhang J, Ewart HS. 2011. Chemical structures and bioactivities of sulfated polysaccharides from marine algae. *Mar Drugs* **9**: 196–223.
- Johnson KA, Goody RS. 2011. The original Michaelis constant: Translation of the 1913 Michaelis-Menten Paper. *Biochemistry* **50**: 8264–8269.
- Johnston KH, McCandless EL. 1973. Enzymic hydrolysis of the potassium chloride soluble fraction of carrageenan: properties of “lambda carrageenases” from *Pseudomonas carrageenovora*. *Can J Microbiol* **19**: 779–88.
- Kabisch A, Otto A, König S, Becher D, Albrecht D, Schüler M, Teeling H, Amann RI, Schweder T. 2014. Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes “*Gramella forsetii*” KT0803. *ISME J* **8**: 1492–502.
- Kalra EK. 2003. Nutraceutical - Definition and Introduction. *AAPS PharmSci* **5**: 1–2.
- Kang DH, Hyeon JE, You SK, Kim SW, Han SO. 2014. Efficient enzymatic degradation process for hydrolysis activity of the Carrageenan from red algae in marine biomass. *J Biotechnol* **192**: 108–113.
- Katayama, T., Sakuma, A., Kimura, T., Makimura, Y., Hiratake, J., Sakata, K., ... & Yamamoto, K. (2004). Molecular cloning and characterization of *Bifidobacterium bifidum* 1, 2- α -L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95). *Journal of bacteriology* **186**: 4885-4893.
- Kelly FC. 1961. Iodine in Medicine and Pharmacy since its discovery -1811-1961. *Proc R Soc Med* **54**: 831–836.
- Kelly LA, Mezulis S, Yates C, Wass M, Sternberg M. 2015. The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc* **10**: 845–858.
- Kidby D, Davidson D. 1973. A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal Biochem* **55**: 321–325.
- Kim HT, Lee S, Kim KH, Choi IG. 2012. The complete enzymatic saccharification of agarose and its application to simultaneous saccharification and fermentation of agarose for ethanol production. *Bioresour Technol* **107**: 301–306.
- Klarzynski O, Descamps V, Plesse B, Yvin J-C, Kloareg B, Fritig B. 2003. Sulfated fucan oligosaccharides elicit defense responses in tobacco and local and systemic resistance against tobacco mosaic virus. *Mol Plant Microbe Interact* **16**: 115–122.
- Kloareg B, Quatrano RS. 1988. Structure of the cell walls of marine algae and ecophysiological

- functions of the matrix polysaccharides. *Ocean Mar Biol Annu Rev* **26**: 259–315.
- Knutsen SH, Grasdalen H. 1987. Characterization of Water-extractable Polysaccharides from Norwegian *Furcellaria lumbricalis* (Huds.) Lamour. (Gigartinales, Rhodophyceae) by IR and NMR Spectroscopy. *Bot Mar* **30**: 497–506.
- Knutsen SH, Myslabodski DE, Larsen B, Usov AI. 1994. A Modified System of Nomenclature for Red Algal Galactans. *Bot Mar* **37**: 163–170.
- Knutser SH, Grasdalen H. 1987. Characterization of Water-extractable Polysaccharides from Norwegian *Furcellaria lumbricalis* (Huds.) Lamour. (Gigartinales, Rhodophyceae) by IR and NMR Spectroscopy. *Bot Mar* **30**: 497–506.
- Kobayashi T, Uchimura K, Koide O, Deguchi S, Horikoshi K. 2012. Genetic and Biochemical Characterization of the *Pseudoalteromonas tetraodonis* Alkaline κ -Carrageenase. *Biosci Biotechnol Biochem* **76**: 506–511.
- Korzen L, Peled Y, Shamir SZ, Shechter M, Gedanken A, Abelson A, Israel A. 2015. An economic analysis of bioethanol production from the marine macroalga *Ulva* (Chlorophyta). *Technology* **3**: 114–118.
- Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev Camb Phylos Soc* **28**: 416–436.
- Kraan S. 2012. Algal polysaccharides, novel applications and outlook. *Carbohydrates - Compr Stud Glycobiol Glycotechnol* 489–532.
- Kummu M, de Moel H, Salvucci G, Daniel Viviroli and Philip J Ward and Olli. 2016. Over the hills and further away from coast: global geospatial patterns of human and environment over the 20th–21st centuries. *Environ Res Lett* **11**: 34010.
- Kupper FC, Kloareg B, Guern J, Potin P. 2001. Oligoguluronates Elicit an Oxidative Burst in the Brown Algal Kelp *Laminaria digitata*. *Plant Physiol* **125**: 278–291.
- Labourel A, Jam M, Jeudy A, Hehemann JH, Czjzek M, Michel G. 2014. The β -glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site adapted for efficient degradation of algal laminarin. *J Biol Chem* **289**: 2027–2042.
- Labourel A, Jam M, Legentil L, Sylla B, Ficko-Blean E, Hehemann J-H, Ferrières V, Czjzek M, Michel G. 2015. Structural and biochemical characterization of the modular laminarinase ZgLamC from *Zobellia galactanivorans* suggests a preferred recognition of branched laminarin. *J Biol Chem* **71**: 173–184.
- Lahaye M, Robic A. 2007. Structure and function properties of Ulvan, a polysaccharide from green seaweeds. *Biomacromolecules* **8**: 1765–1774.
- Laine RA. 1994. Invited commentary: A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: The Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **4**: 759–767.
- Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O, Klintner S, Pudlo NA, Urs K, Koropatkin NM, *et al.* 2014. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* **506**: 498–502.
- Lawson CJ, Rees DA. 1970. An enzyme for the metabolic control of polysaccharide conformation and function. *Nature* **227**: 392–393.
- Lechat H. 1998. Nature et Organisation des Polysaccharides Pariétaux de l'Algue Rouge

Eucheuma cottonii.

- Lemoine M, Nyvall Collén P, Helbert W. 2009. Physical state of κ -carrageenan modulates the mode of action of κ -carrageenase from *Pseudoalteromonas carrageenovora*. *Biochem J* **419**: 545–553.
- Lieutaud P, Canard B, Longhi S. 2008. MeDor: a metaserver for predicting protein disorder. *BMC Genomics* **9**: S25.
- Liu GL, Li Y, Chi Z, Chi ZM. 2011. Purification and characterization of kappa-carrageenase from the marine bacterium *Pseudoalteromonas porphyrae* for hydrolysis of kappa-carrageenan. *Process Biochem* **46**: 265–271.
- Liu Z, Li G, Mo Z, Mou H. 2013. Molecular cloning, characterization, and heterologous expression of a new κ -carrageenase gene from marine bacterium *Zobellia* sp. ZM-2. *Appl Microbiol Biotechnol* **97**: 10057–10067.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: 490–495.
- Mackenzie LF, Wang Q, Warren RAJ, Withers SG. 1998. Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J Am Chem Soc* **120**: 5583–5584.
- Magnusson M, Carl C, Mata L, de Nys R, Paul NA. 2016. Seaweed salt from *Ulva*: A novel first step in a cascading biorefinery model. *Algal Res* **16**: 308–316.
- Major AP. 1977. *The book of seaweed*. Gordon & Cremonesi.
- Martens EC, Chiang HC, Gordon JI. 2008. Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. *Cell Host Microbe* **4**: 447–57.
- Martens EC, Kelly AG, Tauzin AS, Brumer H. 2014. The devil lies in the details: How variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. *J Mol Biol* **426**: 3851–3865.
- Martin M, Barbeyron T, Martin R, Portetelle D, Michel G, Vandenbol M. 2015. The cultivable surface microbiota of the brown alga *Ascophyllum nodosum* is enriched in macroalgal-polysaccharide-degrading bacteria. *Front Microbiol* **6**: 1–14.
- Martin M, Portetelle D, Michel G, Vandenbol M. 2014. Microorganisms living on macroalgae: Diversity, interactions, and biotechnological applications. *Appl Microbiol Biotechnol* **98**: 2917–2935.
- McLean MW, Williamson FB. 1979. Glycosulphatase from *Pseudomonas carrageenovora*. Purification and some properties. *Eur J Biochem* **101**: 497–505.
- McLean MW, Williamson FB. 1981. Neocarratetraose 4-O-monosulfate beta-hydrolase from *Pseudomonas carrageenovora*. *Eur J Biochem* **113**: 443–456.
- Mercier L, Lafitte C, Borderies G, Briand X, Esquerré-Tugayé MT, Fournier J. 2001. The algal polysaccharide carrageenans can act as an elicitor of plant defence. *New Phytol* **149**: 43–51.
- Merzendorfer H, Zimoch L. 2003. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J Exp Biol* **206**: 4393–4412.
- Michel G, Chantalat L, Duee E, Barbeyron T, Henrissat B, Kloareg B, Dideberg O. 2001a. The κ -carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* **9**: 513–25.

- Michel G, Chantalat L, Fanchon E, Henrissat B, Kloareg B, Dideberg O. 2001b. The ι -carrageenase of *Alteromonas fortis*. A beta-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J Biol Chem* **276**: 40202–9.
- Michel G, Czjzek M. 2013. Polysaccharide-degrading enzymes from marine bacteria. In *Marine enzymes for biocatalysis: sources, biocatalytic characteristic and bioprocesses of marine enzymes*, pp. 429–464.
- Michel G, Flament D, Vernet T, Dideberg O. 1999. Expression, purification, crystallization and preliminary X-ray analysis of the κ -carrageenase from *Pseudoalteromonas carrageenovora*. *Acta Crystallogr Sect D Biol Crystallogr* **D55**: 918–920.
- Michel G, Helbert W, Kahn R, Dideberg O, Kloareg B. 2003. The Structural Bases of the Processive Degradation of ι -Carrageenan, a Main Cell Wall Polysaccharide of Red Algae. *J Mol Biol* **334**: 421–433.
- Michel G, Tonon T, Scornet D, Cock JM, 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 JM, 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.
- Mišurcová L. 2011. Chemical Composition of Seaweeds. In *Handbook of Marine Macroalgae: Biotechnology and Applied Phycology*, pp. 171–192, John Wiley & Sons.
- Miyashita K, Narayan B, Tsukui T, Kamogawa H, Abe M, Hosokawa M. 2011. Brown Seaweed Lipids as Potential Source of Omega-3 PUFA in Biological Systems. In *Handbook of Marine Macroalgae: Biotechnology and Applied Phycology*, pp. 329–339, John Wiley & Sons.
- Moller I, Sørensen I, Bernal AJ, Blaukopf C, Lee K, Øbro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP, *et al.* 2007. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J* **50**: 1118–28.
- Morrissey J, Kraan S, Guiry MD. 2001. *A Guide to Commercially Important Sea weeds on the Irish Coast*. Bord Iascaigh Mhara/Irish Sea Fisheries Board
- Mouritsen OG, Dawczynski C, Duelund L, Jahreis G, Vetter W, Schröder M. 2013. On the human consumption of the red seaweed dulse (*Palmaria palmata* (L.) Weber & Mohr). *J Appl Phycol* **25**: 1777–1791.
- Mouritsen OG, Richard D. 2015. *Algues marines : propriétés, usages, recettes*. Delachaux et Niestlé.
- Nam TJ. 2011. Biological Effects of Proteins Extracted from Marine Algae. In *Handbook of Marine Macroalgae: Biotechnology and Applied Phycology*, pp. 387–397, John Wiley & Sons.
- NCBI Resource Coordinators. 2017. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **45**: D12–D17.
- Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, *et al.* 2017. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature*.
- Necas J, Bartosikova L. 2013. Carrageenan : a review. *Vet Med (Praha)* **58**: 187–205.

- Nedashkovskaya OI, Suzuki M, Vancanneyt M, Cleenwerck I, Lysenko AM, Mikhailov V V., Swings J. 2004. *Zobellia amurskyensis* sp. nov., *Zobellia laminariae* sp. nov. and *Zobellia russellii* sp. nov., novel marine bacteria of the family *Flavobacteriaceae*. *Int J Syst Evol Microbiol* **54**: 1643–1648.
- Nisizawa K, Noda H, Kikuchi R, Watanabe T. 1987. The main seaweed foods in Japan. *Hydrobiologia* **151–152**: 5–29.
- O'Dowd C, Jimenez JL, Bahreini R, Flagan RC, Seinfeld JH, Hameri K, Pirjola L, Kulmala M, Jennings SG, Hoffmann T. 2002. Marine aerosol formation from biogenic iodine emissions. *Nature* **417**: 1–5.
- O'Sullivan L, Murphy B, McLoughlin P, Duggan P, Lawlor PG, Hughes H, Gardiner GE. 2010. Prebiotics from marine macroalgae for human and animal health applications. *Mar Drugs* **8**: 2038–2064.
- Olsen JL, Rouzé P, Verhelst B, Lin Y, Bayer T, Collen J, Dattolo E, De Paoli E, Dittami S, Maumus F, *et al.* 2016. The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature* **530**: 331–335.
- Osawa T, Matsubara Y, Muramatsu T, Kimura M, Kakuta Y. 2005. Crystal structure of the alginate (Poly α -L-guluronate) lyase from *Corynebacterium* sp. at 1.2 Å resolution. *J Mol Biol* **345**: 1111–1118.
- Östling S, Virtama P. 1946a. A modified preparation of the universal buffer described by Teorell and Stenhagen. *Acta Physiol Scand* **11**: 289–293.
- Östling S, Virtama P. 1946b. A Modified Preparation of the Universal Buffer Described by Teorell and Stenhagen. *Acta Physiol Scand* **11**: 289–293.
- Payne CM, Babans J, Horn SJ, Backe PH, Arvai AS, Dalhus B, Bjørns M, Eijsink VGH, Sørli M, Beckham GT, *et al.* 2012. Hallmarks of processivity in glycoside hydrolases from crystallographic and computational studies of the *Serratia marcescens* chitinases. *J Biol Chem* **287**: 36322–36330.
- Pêches et Océans Canada . 2013. L'Aquaculture au Canada: Aquaculture Multi-trophique intégrée (AMTI).
- Philippsen A, Wild P, Rowe A. 2014. Energy input, carbon intensity and cost for ethanol produced from farmed seaweed. *Renew Sustain Energy Rev* **38**: 609–623.
- Philpott J, Bradford M. 2006. Seaweed : Nature's Secret for a Long and Healthy Life ? *Nutr Pract* Winter 1-21.
- Planas A. 2000. Bacterial 1,3-1,4- β -glucanases : structure, function and protein engineering. *Biochim Biophys Acta* **1543**: 361–382.
- Potin P, Richard C, Barbeyron T, Henrissat B, Gey C, Petillot Y, Forest E, Dideberg O, Rochas C, Kloareg B. 1995. Processing and Hydrolytic Mechanism of the cgkA-Encoded κ -Carrageenase of *Alteromonas carrageenovora*. *Eur J Biochem* **228**: 971–975.
- Potin P, Sanseau A, Le Gall Y, Rochas C, Kloareg B. 1991. Purification and characterization of a new κ -carrageenase from a marine *Cytophaga*-like bacterium. *Eur J Biochem* **201**: 241–7.
- Préchoux A. 2012. Modification des propriétés physico-chimiques des carraghénanes par la purification et caractérisation de carraghénane-sulfatases.
- Préchoux A, Genicot S, Rogniaux H, Helbert W. 2013. Controlling carrageenan structure using

- a novel formylglycine-dependent sulfatase, an endo-4S-iota-carrageenan sulfatase. *Mar Biotechnol (NY)* **15**: 265–74.
- Préchoux A, Genicot S, Rogniaux H, Helbert W. 2016. Enzyme-Assisted Preparation of Furcellaran-Like κ - β -Carrageenan. *Mar Biotechnol* **18**: 133–143.
- Purich DL. 2010. *Enzyme Kinetics: Catalysis & Control: A Reference of Theory and Best-Practice Methods*. Elsevier.
- Qi H, Zhang Q, Zhao T, Hu R, Zhang K, Li Z. 2006. In vitro antioxidant activity of acetylated and benzoylated derivatives of polysaccharide extracted from *Ulva pertusa* (Chlorophyta). *Bioorganic Med Chem Lett* **16**: 2441–2445.
- Quemener B, Lahaye M, Bobin-Dubigeon C. 1997. Sugar determination in ulvans by a chemical-enzymatic method coupled to high performance anion exchange chromatography. *J Appl Phycol* **9**: 179–188.
- Rebuffet E, Barbeyron T, Jeudy A, Jam M, Czjzek M, Michel G. 2010. Identification of catalytic residues and mechanistic analysis of family GH82 ι -carrageenases. *Biochemistry* **49**: 7590–7599.
- Rebuffet E, Groisillier A, Thompson A, Jeudy A, Barbeyron T, Czjzek M, Michel G. 2011. Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ Microbiol* **13**: 1253–1270.
- Rees DA, Steele IW, Williamson FB, Road WM. 1969. Conformational analysis of polysaccharides. III. The relation between stereochemistry and properties of some natural polysaccharide sulfates (1). *J Polym Sci* **276**: 261–276.
- Renn DW, Santos GA, Dumont LE, Parent CA, Stanley NF, Stancioff DJ, Guiseley KB. 1993. β -Carrageenan : isolation and characterization. *Carbohydr Polym* **22**: 247–252.
- Rhein-Knudsen N, Ale MT, Meyer AS. 2015. Seaweed hydrocolloid production: An update on enzyme assisted extraction and modification technologies. *Mar Drugs* **13**: 3340–3359.
- Riordan JF, Hayashida H. 1970. Chemical evidence for a functional carboxyl group in carboxypeptidase A. *Biochem Biophys Res Commun* **41**: 122–127.
- Riordan JF, McElvany KD, Borders CL. 1977. Arginyl residues: anion recognition sites in enzymes. *Science* **195**: 884–886.
- Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**: 320–324.
- Robic A, Gaillard C, Sassi JF, Leral Y, Lahaye M. 2009a. Ultrastructure of Ulvan: A polysaccharide from green seaweeds. *Biopolymers* **91**: 652–664.
- Robic A, Sassi JF, Dion P, Lerat Y, Lahaye M. 2009b. Seasonal variability of physicochemical and rheological properties of ulvan in two *Ulva* species (chlorophyta) from the Brittany coast. *J Phycol* **45**: 962–973.
- Ropartz D, Giuliani A, Fanuel M, Hervé C, Czjzek M, Rogniaux H. 2016. Online coupling of high-resolution chromatography with extreme UV photon activation tandem mass spectrometry: Application to the structural investigation of complex glycans by dissociative photoionization. *Anal Chim Acta* **933**: 1–9.
- Sakurama H, Fushinobu S, Hidaka M, Yoshida E, Honda Y, Ashida H, Kitaoka M, Kumagai H, Yamamoto K, Katayama T. 2012a. 1,3-1,4- α -L-Fucosynthase that specifically introduces Lewis a/x antigens into type-1/2 chains. *J Biol Chem* **287**: 16709–16719.

- Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. 2012b. Differences in the Substrate Specificities and Active-Site Structures of Two α -L-Fucosidases (Glycoside Hydrolase Family 29) from *Bacteroides thetaiotaomicron*. *Biosc Biotech Biochem* **76**: 1022–1024.
- Salmeán AA, Guillouzo A, Duffieux D, Jam M, Matard-Mann M, Larocque R, Pedersen HL, Michel G, Czjzek M, William GT. 2017. Double blind microarray-based polysaccharide profiling enables parallel identification of uncharacterized polysaccharides and carbohydrate-binding proteins with unknown specificities. *Submitted to Nature*.
- Sánchez B, Delgado S, Blanco-Míguez A, Lourenço A, Gueimonde M, Margolles A. 2017. Probiotics, gut microbiota, and their influence on host health and disease. *Mol Nutr Food Res* **61**: 1–15.
- Schroeder BO, Bäckhed F. 2016. Signals from the gut microbiota to distant organs in physiology and disease. *Nat Med* **22**: 1079–1089.
- Schultz-johansen M, Glaring MA, Bech PK, Stougaard P. 2016. Draft Genome Sequence of a Novel Marine Bacterium, *Paraglaciecola* sp . Strain S66, with Hydrolytic Activity against Seaweed. **4**: 4–5.
- Seisun D, Al-Mosawi AJ, Imeson AP, Helgerud T, Gaserød O, Fjæreide T, Imeson AP, Corporation FMC, Cash MJ, Caputo SJ, *et al.* 2010. *Food Stabilisers, Thickeners and Gelling Agents*. ed. A. Imeson. Wiley-Blackwell.
- Senni K, Gueniche F, Foucault-Bertaud A, Igondjo-Tchen S, Fioretti F, Collic-Jouault S, Durand P, Guezennec J, Godeau G, Letourneur D. 2006. Fucoidan a sulfated polysaccharide from brown algae is a potent modulator of connective tissue proteolysis. *Arch Biochem Biophys* **445**: 56–64.
- Senni K, Pereira J, Gueniche F, Delbarre-Ladrat C, Siquin C, Ratiskol J, Godeau G, Fischer AM, Helley D, Collic-Jouault S. 2011. Marine polysaccharides: A source of bioactive molecules for cell therapy and tissue engineering. *Mar Drugs* **9**: 1664–1681.
- Seydel A, Gounon P, Pugsley AP. 1999. Testing the “+2 rule” for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Mol Microbiol* **34**: 810–821.
- Shaikh FA, Müllegger J, He S, Withers SG. 2007. Identification of the catalytic nucleophile in Family 42 β -galactosidases by intermediate trapping and peptide mapping: YesZ from *Bacillus subtilis*. *FEBS Lett* **581**: 2441–2446.
- Shubin NH, Daeschler EB, Jenkins FA. 2014. Pelvic girdle and fin of *Tiktaalik roseae*. *Proc Natl Acad Sci* **111**: 893–899.
- Sinnott ML. 1990. Catalytic mechanisms of enzymatic glycosyl transfer. *Chem Rev* **90**: 1171–1202.
- Smith SP, Bayer E. 2013. Insights into cellulosome assembly and dynamics: from dissection to reconstruction of the supramolecular enzyme complex. *Curr Opin Struct Biol* **23**: 686–94.
- Spezio M, Wilson DB, Karplus PA. 1993. Crystal Structure of the Catalytic Domain of a Thermophilic Endocellulase. *Biochemistry* **32**: 9906–9916.
- Stam MR, Danchin EGJ, Rancurel C, Coutinho PM, Henrissat B. 2006. Dividing the large glycoside hydrolase family 13 into subfamilies: Towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng Des Sel* **19**: 555–562.
- Stein JR, Borden CA. 1984. Causative and beneficial algae in human disease conditions: a review. *Phycologia* **23**: 485–501.

- Stern R, Jedrzejewski MJ. 2008. Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity. *Chem Rev* **108**: 5061–5085.
- Strohmeier M, Hrmova M, Fischer M, Harvey AJ, Fincher GB, Pleiss J. 2004. Molecular modeling of family GH16 glycoside hydrolases: potential roles for xyloglucan transglucosylases/hydrolases in cell wall modification in the poaceae. *Protein Sci* **13**: 3200–3213.
- Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* **41**: 207–234.
- Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y. 2004. Crystal structure of *Thermotoga maritima* α -L-fucosidase: Insights into the catalytic mechanism and the molecular basis for fucosidosis. *J Biol Chem* **279**: 13119–13128.
- Swain MR, Natarajan V, Krishnan C. 2017. Marine Enzymes and Microorganisms for Bioethanol Production. *Adv Food Nutr Res* **80**.
- Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J, et al. 2012. Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. *Science* **336**: 608–611.
- Teeling H, Fuchs BM, Bennke CM, Kru K, Chafee M, Kappelmann L, Reintjes G, Waldmann J, Lucas J, Wichels A, et al. 2016. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *Elife* **5**: 1–31.
- Terrapon N, Lombard V, Gilbert HJ, Henrissat B. 2015. Automatic prediction of polysaccharide utilization loci in *Bacteroidetes* species. *Bioinformatics* **31**: 647–655.
- Thomas F. 2011. Identification et caractérisation du système alginolytique de la bactérie marine *Zobellia galactanivorans*.
- Thomas F, Barbeyron T, Tonon T, Génicot S, Czjzek M, Michel G. 2012. Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine *Flavobacteriia* to their independent transfers to marine *Proteobacteria* and human gut *Bacteroides*. *Environ Microbiol* **14**: 2379–94.
- Thomas F, Bordron P, Eveillard D, Michel G. 2017. Gene Expression Analysis of *Zobellia galactanivorans* during the Degradation of Algal Polysaccharides Reveals both Substrate-Specific and Shared Transcriptome-Wide Responses. *Front Microbiol* **8**: 1808.
- Thomas F, Cosse A, Le Panse S, Kloareg B, Potin P, Leblanc C. 2014. Kelps feature systemic defense responses: Insights into the evolution of innate immunity in multicellular eukaryotes. *New Phytol* **204**: 567–576.
- Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut *Bacteroidetes*: The food connection. *Front Microbiol* **2**: 1–16.
- Tobacman JK. 2001. Review of harmful gastrointestinal effects of carrageenan in animal experiments. *Environ Health Perspect* **109**: 983–994.
- Tomme P, van Tilbeurgh H, Pettersson G, van Damme J, Vandekerckhove J, Knowles J, Teeri T, Claeysens M. 1988. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. *Eur J Biochem* **170**: 575–581.
- Torode TA, Siméon A, Marcus SE, Jam M, Le Moigne M-A, Duffieux D, Knox JP, Hervé C. 2016. Dynamics of cell wall assembly during early embryogenesis in the brown alga *Fucus*. *J Exp Bot* **67**: 6089–6100.

- Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, Thies C. 2005. Landscape perspectives on agricultural intensification and biodiversity - Ecosystem service management. *Ecol Lett* **8**: 857–874.
- Turner NJ. 2003. The Ethnobotany of edible seaweed (*Porphyra abbottae* and related species; Rhodophyta: Bangiales) and its use by First Nations on the Pacific Coast of Canada. *Can J Bot* **81**: 283–293.
- Usov AI. 2011. *Polysaccharides of the red algae*. 1st ed. Elsevier Inc.
- Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN. 2004. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr Sect D Biol Crystallogr* **60**: 2184–2195.
- Vagin A, Teplyakov A. 1997. MOLREP: an Automated Program for Molecular Replacement. *J Appl Cryst* **30**: 1022–1025.
- Van De Velde F, Rollema HS, Grinberg N V, Burova T V, Grinberg VY, Hans Tromp R. 2002. Coil-helix transition of ι -carrageenan as a function of chain regularity. *Biopolymers* **65**: 299–312.
- Varki A, Lowe JB. 2009. *Biological Roles of Glycans*. Cold Spring Harbor Laboratory Press.
- Varrot A, Frandsen TP, von Ossowski I, Boyer V, Cottaz S, Driguez H, Schülein M, Davies GJ. 2003. Structural Basis for Ligand Binding and Processivity in Cellobiohydrolase Cel6A from *Humicola insolens*. *Structure* **11**: 855–864.
- Violle N, Rozan P, Demais H, Nyvall Collen P, Bisson J-F. 2017. Evaluation of the antidepressant- and anxiolytic-like effects of a hydrophilic extract from the green seaweed *Ulva* sp. in rats. *Nutr Neurosci* **30**: 1–9.
- Von Ossowski I, Ståhlberg J, Koivula A, Piens K, Becker D, Boer H, Harle R, Harris M, Divne C, Mahdi S, *et al.* 2003. Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D. *J Mol Biol* **333**: 817–829.
- von Sonntag C. 1980. Free-radical reactions of carbohydrates as studied by radiation techniques. *Adv Carbohydr Chem Biochem* **37**: 7–77.
- Wang L, Li S, Zhang S, Li J, Yu W, Gong Q. 2015. A new κ -carrageenase CgkS from marine bacterium *Shewanella* sp. Kz7. *J Ocean Univ China* **14**: 759–763.
- Wang L, Zhang Y, Gao P. 2008. A novel function for the cellulose binding module of cellobiohydrolase I. *Sci China, Ser C Life Sci* **51**: 620–629.
- Watanabe T, Ariga Y, Sato U, Toratani T, Hashimoto M, Nikaidou N, Kezuka Y, Nonaka T, Sugiyama J. 2003. Aromatic residues within the substrate-binding cleft of *Bacillus circulans* chitinase A1 are essential for hydrolysis of crystalline chitin. *Biochem J* **376**: 237–244.
- Weigl J, Yaphe W. 1966. The enzymatic hydrolysis of carrageenan by *Pseudomonas carrageenovora*: purification of a κ -carrageenase. *Can J Microbiol* **12**: 939–947.
- Wiesmann C, Beste G, Hengstenberg W, Schulz GE. 1995. The three-dimensional structure of 6-phospho- β -galactosidase from *Lactococcus lactis*. *Structure* **3**: 961–968.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, *et al.* 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* **67**: 235–242.

- Wolfe MS. 2000. Crop strength through diversity. *Nature* **406**: 681–682.
- Wolfenden R, Lu X, Young G. 1998. Spontaneous hydrolysis of glycosides. *J Am Chem Soc* **120**: 6814–6815.
- Wong KF, Craigie JS. 1978. Sulfohydrolase Activity and Carrageenan Biosynthesis in *Chondrus crispus* (Rhodophyceae). *Plant Physiol* **61**: 663–666.
- Wood CG. 1974. Seaweed extracts: A unique ocean resource. *J Chem Educ* **51**: 449.
- Wuitner E. 1921. *Les algues marines de côtes de France*. ed. P. Lechevalier.
- Xie G, Bruce DC, Challacombe JF, Chertkov O, Detter JC, Gilna P, Han CS, Lucas S, Misra M, Myers GL, *et al.* 2007. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. *Appl Environ Microbiol* **73**: 3536–3546.
- Xu H, Qin S, Lan Y, Liu M, Cao X, Qiao D, Cao Y, Cao Y. 2017. Comparative genomic analysis of *Paenibacillus* sp. SSG-1 and its closely related strains reveals the effect of glycometabolism on environmental adaptation. *Sci Rep* **7**: 5720.
- Xu X, Li S, Yang X, Yu W, Han F. 2015. Cloning and characterization of a new κ -carrageenase gene from marine bacterium *Pseudoalteromonas* sp. QY203. *J Ocean Univ China* **14**: 1082–1086.
- Yang B, Bhattacharyya S, Linhardt R, Tobacman J. 2012. Exposure to common food additive carrageenan leads to reduced sulfatase activity and increase in sulfated glycosaminoglycans in human epithelial cells. *Biochimie* **94**: 1309–1316.
- Yang B, Yu G, Zhao X, Ren W, Jiao G, Fang L, Wang Y, Du G, Tiller C, Girouard G, *et al.* 2011. Structural characterisation and bioactivities of hybrid carrageenan-like sulphated galactan from red alga *Furcellaria lumbricalis*. *Food Chem* **124**: 50–57.
- Yaphe W, Baxter B. 1955. The enzymic hydrolysis of carrageenin. *Appl Microbiol* **3**: 380–383.
- Yermak IM, Kim YH, Titlynov EA, Isakov VV, Solov'eva TF. 1999. Chemical structure and gel properties of carrageenans from algae belonging to the Gigartinales and Tichocarpaceae, collected from the Russian Pacific Coast. *J Appl Phycol* **11**: 41–48.
- Yew WS, Fedorov AA, Fedorov EV, Rakus JF, Pierce RW, Almo SC, Gerlt JA. 2006. Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry* **45**: 14582–14597.
- Zablackis E, Vreeland V, Kloareg B. 2009. Isolation of Protoplasts from *Kappaphycus alvarezii* var. *tambalang* (Rhodophyta) and Secretion of ι -Carrageenan Fragments by Cultured Cells. *Carbohydr Res* **44**: 1515–1522.
- Zechel DL, Withers SG, And DLZ, Withers SG. 2000. Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc Chem Res* **33**: 11–18.
- Zeuner B, Jers C, Mikkelsen JD, Meyer AS. 2014. Methods for improving enzymatic trans-glycosylation for synthesis of human milk oligosaccharide biomimetics. *J Agric Food Chem* **62**: 9615–9631.
- Zhao T, Zhang Q, Qi H, Zhang H, Niu X, Xu Z, Li Z. 2006. Degradation of porphyran from *Porphyra haitanensis* and the antioxidant activities of the degraded porphyrans with different molecular weight. *Int J Biol Macromol* **38**: 45–50.
- Zhou G, Sun YP, Xin H, Zhang Y, Li Z, Xu Z. 2004. In vivo antitumor and immunomodulation activities of different molecular weight lambda-carrageenans from *Chondrus ocellatus*. *Pharmacol Res* **50**: 47–53.

- Zhu Y, Thomas F, Larocque R, Li N, Duffieux D, Cladière L, Souchaud F, Michel G, McBride MJ. 2017. Genetic analyses unravel the crucial role of a horizontally acquired alginate lyase for brown algal biomass degradation by *Zobellia galactanivorans*. *Environ Microbiol* **19**: 2164–2181.
- Zinoun M, Diouris M, Potin P, Floch JY, Deslandes E. 1997. Evidence of sulfhydrolase activity in the red alga *Calliblepharis jubata*. *Bot Mar* **40**: 49–53.

Bibliographie II

- Allouch J, Helbert W, Henrissat B, Czjzek M. 2004. Parallel substrate binding sites in a β -agarase suggest a novel mode of action on double-helical agarose. *Structure* **12**: 623–632.
- Allouch J, Jam M, Helbert W, Barbeyron T, Kloareg B, Henrissat B, Czjzek M. 2003. The Three-dimensional Structures of Two β -Agarases. *J Biol Chem* **278**: 47171–47180.
- Aoki Y, Kamei Y. 2006. Preparation of recombinant polysaccharide-degrading enzymes from the marine bacterium, *Pseudomonas* sp. ND137 for the production of protoplasts of *Porphyra yezoensis*. *Eur J Phycol* **41**: 321–328.
- Araki T, Hayakawa M, Lu Z, Karita S, Morishita T. 1998. Purification and characterization of agarases from a marine bacterium, *Vibrio* sp. PO-303. *J Mar Biotechnol* **6**: 260–265.
- Badur AH, Jagtap SS, Yalamanchili G, Lee JK, Zhao H, Rao C V. 2015. Alginate lyases from alginate-degrading *Vibrio splendidus* 12B01 are endolytic. *Appl Environ Microbiol* **81**: 1865–1873.
- Barbeyron T, Gerard A, Potin P, Henrissat B, Kloareg B. 1998. The Kappa-Carrageenase of the Marine Bacterium *Cytophaga drobachiensis*. Structural and Phylogenetic Relationships Within Family-16 Glycoside Hydrolases. *Mol Biol Evol* **15**: 528–537.
- Barbeyron T, Henrissat B, Kloareg B. 1994. The gene encoding the kappa-carrageenase of *Alteromonas carrageenovora* is related to β -1,3-1,4-glucanases. *Gene* **139**: 105–109.
- Barbeyron T, Michel G, Potin P, Henrissat B, Kloareg B. 2000. ι -Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of κ -carrageenases. *J Biol Chem* **275**: 35499–35505.
- Belas R. 1989. Sequence Analysis of the agrA Gene Encoding β -Agarase from *Pseudomonas atlantica*. *J Bacteriol* **171**: 602–605.
- Bleicher L, Prates ET, Gomes TCF, Silveira RL, Nascimento AS, Rojas AL, Golubev A, Martínez L, Skaf MS, Polikarpov I. 2011. Molecular basis of the thermostability and thermophilicity of laminarinases: X-ray structure of the hyperthermostable laminarinase from *Rhodothermus marinus* and molecular dynamics simulations. *J Phys Chem B* **115**: 7940–7949.
- Chi WJ, Park DY, Seo Y Bin, Chang YK, Lee SY, Hong SK. 2014. Cloning, expression, and biochemical characterization of a novel GH16 β -agarase AgaG1 from *Alteromonas* sp. GNUM-1. *Appl Microbiol Biotechnol* **98**: 4545–4555.
- Chhabra SR, Shockley KR, Ward DE, Kelly RM. 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. *Appl. Environ. Microbiol.*, **68**: 545-554.
- Cobucci-Ponzano B, Conte F, Bedini E, Corsaro MM, Parrilli M, Sulzenbacher G, Lipski A, Dal Piaz F, Lepore L, Rossi M, et al. 2009. β -Glycosyl Azides as Substrates for α -Glycosynthases: Preparation of Efficient α -L-Fucosynthases. *Chem Biol* **16**: 1097–1108.
- Colin S, Deniaud E, Jam M, Descamps V, Chevotot Y, Kervarec N, Yvin JC, Barbeyron T, Michel G, Kloareg B. 2006. Cloning and biochemical characterization of the fucanase FcnA: Definition of a novel glycoside hydrolase family specific for sulfated fucans. *Glycobiology* **16**: 1021–1032.

- Cui F, Dong S, Shi X, Zhao X, Zhang X-H. 2014. Overexpression and Characterization of a Novel Thermostable β -Agarase YM01-3, from Marine Bacterium *Catenovulum agarivorans* YM01T. *Mar Drugs* **12**: 2731–2747.
- Dong J, Hashikawa S, Konishi T, Tamaru Y, Araki T. 2006. Cloning of the novel gene encoding β -agarase C from a marine bacterium, *Vibrio* sp. strain PO-303, and characterization of the gene product. *Appl Environ Microbiol* **72**: 6399–6401.
- Dong J, Tamaru Y, Araki T. 2007a. A unique β -agarase, AgaA, from a marine bacterium, *Vibrio* sp. strain PO-303. *Appl Microbiol Biotechnol* **74**: 1248–1255.
- Dong J, Tamaru Y, Araki T. 2007b. Molecular Cloning, Expression, and Characterization of a β -Agarase Gene, agaD, from a Marine Bacterium, *Vibrio* sp. Strain PO-303. *Biosci Biotechnol Biochem* **71**: 38–46.
- Dong Q, Ruan L, Shi H. 2016. A β -agarase with high pH stability from *Flammeovirga* sp. SJP92. *Carbohydr Res* **432**: 1–8.
- Dong S, Wei T Di, Chen XL, Li CY, Wang P, Xie B Bin, Qin QL, Zhang XY, Pang XH, Zhou BC, et al. 2014. Molecular insight into the role of the N-terminal extension in the maturation, substrate recognition, and catalysis of a bacterial alginate lyase from polysaccharide lyase family 18. *J Biol Chem* **289**: 29558–29569.
- Du Z, Wang J, Yang L, Chen G. 2011. Identification of a marine agarolytic bacterium *Agarivorans albus* QM38 and cloning and sequencing its beta-agarase genes. *Acta Oceanol Sin* **30**: 118–124.
- Ekborg NA, Taylor LE, Longmire AG, Henrissat B, Weiner RM, Steven W, Hutcheson SW. 2006. Genomic and Proteomic Analyses of the Agarolytic System Expressed by *Saccharophagus degradans* 2-40. *Appl Environ Microbiol* **72**: 3396–3405.
- Flament D, Barbeyron T, Jam M, Potin P, Czjzek M, Kloareg B, Michel G. 2007. Alpha-agarases define a new family of glycoside hydrolases, distinct from beta-agarase families. *Appl Environ Microbiol* **73**: 4691–4694.
- Fu XT, Lin H, Kim SM. 2008. Purification and characterization of a novel β -agarase, AgaA34, from *Agarivorans albus* YKW-34. *Appl Microbiol Biotechnol* **78**: 265–273.
- Fu XT, Pan C-H, Lin H, Kim SM. 2009. Gene cloning, expression, and characterization of a beta-agarase, agaB34, from *Agarivorans albus* YKW-34. *J Microbiol Biotechnol* **19**: 257–264.
- Guibet M, Colin S, Barbeyron T, Genicot S, Kloareg B, Michel G, Helbert W. 2007. Degradation of λ -carrageenan by *Pseudoalteromonas carrageenovora* λ -carrageenase: a new family of glycoside hydrolases unrelated to κ - and ι -carrageenases. *Biochem J* **404**: 105–114.
- Ha SC, Lee S, Lee J, Kim HT, Ko HJ, Kim KH, Choi IG. 2011. Crystal structure of a key enzyme in the agarolytic pathway, α -neoagarobiose hydrolase from *Saccharophagus degradans* 2-40. *Biochem Biophys Res Commun* **412**: 238–244.
- Han WJ, Gu JY, Liu HH, Li FC, Wu ZH, Li YZ. 2013. An extra peptide within the catalytic module of a β -agarase affects the agarose degradation pattern. *J Biol Chem* **288**: 9519–31.
- Hatada Y, Mizuno M, Li Z, Ohta Y. 2011. Hyper-production and Characterization of the ι -Carrageenase Useful for ι -Carrageenan Oligosaccharide Production from a Deep-sea Bacterium, *Microbulbifer thermotolerans* JAMB-A94 T, and Insight into the Unusual Catalytic Mechanism. *Mar Biotechnol* **13**: 411–422.

- Hatada Y, Ohta Y, Horikoshi K. 2006. Hyperproduction and application of α -agarase to enzymatic enhancement of antioxidant activity of porphyran. *J Agric Food Chem* **54**: 9895–9900.
- Hehemann JH, 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.
- Hehemann JH, Correc G, Thomas F, Bernard T, Barbeyron T, Jam M, Helbert W, Michel G, Czjzek M. 2012a. Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium *Zobellia galactanivorans*. *J Biol Chem* **287**: 30571–84.
- Hehemann JH, Kelly AG, Pudlo NA, Martens EC, Boraston AB. 2012b. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc Natl Acad Sci* **109**: 19786–19791.
- Hu Z, Lin BK, Xu Y, Zhong MQ, Liu GM. 2009. Production and purification of agarase from a marine agarolytic bacterium *Agarivorans* sp. HZ105. *J Appl Microbiol* **106**: 181–190.
- Ilari A, Fiorillo A, Angelaccio S, Florio R, Chiaraluce R, Van Der Oost J, Consalvi V. 2009. Crystal structure of a family 16 endoglucanase from the hyperthermophile *Pyrococcus furiosus* - Structural basis of substrate recognition. *FEBS J* **276**: 1048–1058.
- Jam M, Flament D, Allouch J, Potin P, Thion L, Kloareg B, Czjzek M, Helbert W, Michel G, Barbeyron T. 2005. The endo- β -agarases AgaA and AgaB from the marine bacterium *Zobellia galactanivorans*: two paralogue enzymes with different molecular organizations and catalytic behaviours. *Biochem J* **385**: 703–713.
- Jeng WY, Wang NC, Lin CT, Shyur LF, Wang AHJ. 2011. Crystal structures of the laminarinase catalytic domain from *Thermotoga maritima* MSB8 in complex with inhibitors: Essential residues for β -1,3- and β -1,4-glucan selection. *J Biol Chem* **286**: 45030–45040.
- Jeon MJ, Kim AR, Lee DG, Lee SH. 2012. Cloning, Expression, and Characterization of a Novel GH-16 β -Agarase from *Agarivorans* sp. JA-1. *J Life Sci* **22**: 1545–1551.
- Kin J, Hong SK. 2012. Isolation and characterization of an agarase-producing bacterial strain, *Alteromonas* sp. GNUM-1, from the west sea, Korea. *J Microbiol Biotechnol* **22**: 1621–1628.
- Kirimura K, Masuda N, Iwasaki Y, Nakagawa H, Kobayashi R, Usami S. 1999. Purification and characterization of a novel β -agarase from an alkalophilic bacterium, *Alteromonas* sp. E-1. *J Biosci Bioeng* **87**: 436–441.
- Kitamura E, Kamei Y. 2006. Molecular cloning of the gene encoding β -1,3(4)-glucanase a from a marine bacterium, *Pseudomonas* sp. PE2, an essential enzyme for the degradation of *Pythium porphyrae* cell walls. *Appl Microbiol Biotechnol* **71**: 630–637.
- Kobayashi T, Uchimura K, Koide O, Deguchi S, Horikoshi K. 2012. Genetic and Biochemical Characterization of the *Pseudoalteromonas tetraodonis* Alkaline κ -Carrageenase. *Biosci Biotechnol Biochem* **76**: 506–511.
- Kusaykin MI, Belik AA, Kovalchuk SN, Dmitrenok PS, Rasskazov VA, Isakov VV, Zvyagintseva TN. 2017. A new recombinant endo-1,3- β -d-glucanase from the marine bacterium *Formosa algae* KMM 3553: enzyme characteristics and transglycosylation products analysis. *World J Microbiol Biotechnol* **33**: 1–12.

- Labourel A, Jam M, Jeudy A, Hehemann JH, Czjzek M, Michel G. 2014. The β -glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site adapted for efficient degradation of algal laminarin. *J Biol Chem* **289**: 2027–2042.
- Lee DG, Park GT, Kim NY, Lee EJ, Jang MK, Shin YG, Park GS, Kim TM, Lee JH, Lee JH, et al. 2006. Cloning, expression, and characterization of a glycoside hydrolase family 50 β -agarase from a marine *Agarivorans* isolate. *Biotechnol Lett* **28**: 1925–1932.
- Lee Y, Jeon S-M, Ye B-R, Jang J, Kim J, Heo S-J, Kang D-H. 2013a. A Novel GH-50 Family β -agarase from *Saccharophagus* sp. AG21: Molecular Cloning and Enzymatic Characterization. *Int Meet Microbiol Soc Korea* **5**: 282–282.
- Lee Y, Oh C, De Zoysa M, Kim H, Don Niroshana Wickramaarachchi W, Whang I, Kang DH, Lee J. 2013b. Molecular cloning, overexpression, and enzymatic characterization of glycosyl hydrolase family 16 β -agarase from marine bacterium *Saccharophagus* sp. AG21 in *Escherichia coli*. *J Microbiol Biotechnol* **23**: 913–922.
- Li J, Hu Q, Li Y, Xu Y. 2015. Purification and characterization of cold-adapted beta-agarase from an antarctic psychrophilic strain. *Brazilian J Microbiol* **46**: 683–690.
- Li JW, Dong S, Song J, Li CB, Chen XL, Xie B Bin, Zhang YZ. 2011. Purification and characterization of a bifunctional alginate lyase from *Pseudoalteromonas* sp. SM0524. *Mar Drugs* **9**: 109–123.
- Li S, Hao J, Sun M. 2017. Macromolecules Cloning and characterization of a new cold-adapted and thermo-tolerant ι -carrageenase from marine bacterium *Flavobacterium* sp. YS-80-122. *Int J Biol Macromol* **102**: 1059–1065.
- Liang SS, Chen YP, Chen YH, Chiu SH, Liaw LL. 2014. Characterization and overexpression of a novel β -agarase from *Thalassomonas agarivorans*. *J Appl Microbiol* **116**: 563–572.
- Liao L, Xu XW, Jiang XW, Cao Y, Yi N, Huo YY, Wu YH, Zhu XF, Zhang XQ, Wu M. 2011. Cloning, expression, and characterization of a new beta-agarase from *Vibrio* sp. strain CN41. *Appl Environ Microbiol* **77**: 7077–9.
- Lin B, Lu G, Zheng Y, Xie W, Li S, Hu Z. 2012. Gene cloning, expression and characterization of a neoagarotetraose-producing β -agarase from the marine bacterium *Agarivorans* sp. HZ105. *World J Microbiol Biotechnol* **28**: 1691–1697.
- Liu GL, Li Y, Chi Z, Chi ZM. 2011. Purification and characterization of kappa-carrageenase from the marine bacterium *Pseudoalteromonas porphyrae* for hydrolysis of kappa-carrageenan. *Process Biochem* **46**: 265–271.
- Liu N, Mao X, Du Z, Mu B, Wei D. 2014a. Cloning and characterisation of a novel neoagarotetraose-forming- β -agarase, AgWH50A from *Agarivorans gilvus* WH0801. *Carbohydr Res* **388**: 147–151.
- Liu N, Mao X, Yang M, Mu B, Wei D. 2014b. Gene cloning, expression and characterisation of a new β -agarase, AgWH50C, producing neoagarobiose from *Agarivorans gilvus* WH0801. *World J Microbiol Biotechnol* **30**: 1691–1698.
- Liu Z, Li G, Mo Z, Mou H. 2013. Molecular cloning, characterization, and heterologous expression of a new κ -carrageenase gene from marine bacterium *Zobellia* sp. ZM-2. *Appl Microbiol Biotechnol* **97**: 10057–10067.
- Long M, Yu Z, Xu X. 2010. A novel β -agarase with high pH stability from marine *Agarivorans* sp. LQ48. *Mar Biotechnol* **12**: 62–69.
- Lu X, Chu Y, Wu Q, Gu Y, Han F, Yu W. 2009. Cloning, expression and characterization of a

- new agarase-encoding gene from marine *Pseudoalteromonas* sp. *Biotechnol Lett* **31**: 1565–1570.
- Ma C, Lu X, Shi C, Li J, Gu Y, Ma Y, Chu Y, Han F, Gong Q, Yu W. 2007. Molecular cloning and characterization of a novel β -agarase, AgaB, from marine *Pseudoalteromonas* sp. CY24. *J Biol Chem* **282**: 3747–3754.
- Ma S, Duan G, Chai W, Geng C, Tan Y, Wang L, Le Sourd F, Michel G, Yu W, Han F. 2013. Purification, Cloning, Characterization and Essential Amino Acid Residues Analysis of a New ι -Carrageenase from *Cellulophaga* sp. QY3. *PLoS One* **8**.
- McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B, Rhodes RG, Goltsman E, Wang W, Xu J, Hunnicutt DW, et al. 2009. Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl Environ Microbiol* **75**: 6864–6875.
- McLean MW, Williamson FB. 1981. Neocarratetraose 4-O-monosulfate beta-hydrolase from *Pseudomonas carrageenovora*. *Eur J Biochem* **113**: 443–456.
- Michel G, Chantalat L, Duee E, Barbeyron T, Henrissat B, Kloareg B, Dideberg O. 2001a. The κ -carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* **9**: 513–25.
- Michel G, Chantalat L, Fanchon E, Henrissat B, Kloareg B, Dideberg O. 2001b. The ι -carrageenase of *Alteromonas fortis*. A beta-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J Biol Chem* **276**: 40202–9.
- Michel G, Helbert W, Kahn R, Dideberg O, Kloareg B. 2003. The structural bases of the processive degradation of ι -carrageenan, a main cell wall polysaccharide of red algae. *J Mol Biol* **334**: 421–433.
- Morrice LM, McLean MW, Long WF, Williamson FB. 1983. Beta-agarases I and II from *Pseudomonas atlantica*. Substrate specificities. *Eur J Biochem* **137**: 149–54.
- Nikapitiya C, Oh CH, Lee YD, Lee SK, Whang IS, Lee JH. 2010. Characterization of a Glycoside Hydrolase Family 50 Thermostable β -agarase AgrA from Marine Bacteria *Agarivorans* sp. AG17. *Fish Aquat Sci* **13**: 36–48.
- Nyvall Collén P, Jeudy A, Sassi JF, Groisillier A, Czjzek M, Coutinho PM, Helbert W. 2014. A novel unsaturated β -glucuronyl hydrolase involved in ulvan degradation unveils the versatility of stereochemistry requirements in family GH105. *J Biol Chem* **289**: 6199–6211.
- Nyvall Collén P, Sassi JF, Rogniaux H, Marfaing H, Helbert W. 2011. Ulvan lyases isolated from the flavobacteria *Persicivirga ulvanivorans* are the first members of a new polysaccharide lyase family. *J Biol Chem* **286**: 42063–42071.
- Oh C, Nikapitiya C, Lee Y, Whang I, Kang DH, Heo SJ, Choi YU, Lee J. 2010a. Molecular Cloning, Characterization and Enzymatic Properties of a Novel Beta-Agarase From a Marine Isolate *Pseudoalteromonas* Sp. Ag52. *Brazilian J Microbiol* **41**: 876–889.
- Oh C, Nikapitiya C, Lee Y, Whang I, Kim SJ, Kang DH, Lee J. 2010b. Cloning, purification and biochemical characterization of beta agarase from the marine bacterium *Pseudoalteromonas* sp. AG4. *J Ind Microbiol Biotechnol* **37**: 483–494.
- Ohta Y, Hatada Y. 2006. A novel enzyme, β -carrageenase, isolated from a deep-sea bacterium. *J Biochem* **140**: 475–481.
- Ohta Y, Hatada Y, Ito S, Horikoshi K. 2005. High-level expression of a neoagarobiose-

- producing β -agarase gene from *Agarivorans* sp. JAMB-A11 in *Bacillus subtilis* and enzymic properties of the recombinant enzyme. *Biotechnol Appl Biochem* **41**: 183.
- Ohta Y, Hatada Y, Nogi Y, Li Z, Ito S, Horikoshi K. 2004a. Cloning, expression, and characterization of a glycoside hydrolase family 86 β -agarase from a deep-sea *Microbulbifer*-like isolate. *Appl Microbiol Biotechnol* **66**: 266–275.
- Ohta Y, Hatada Y, Nogi Y, Miyazaki M, Li Z, Akita M, Hidaka Y, Goda S, Ito S, Horikoshi K. 2004b. Enzymatic properties and nucleotide and amino acid sequences of a thermostable β -agarase from a novel species of deep-sea *Microbulbifer*. *Appl Microbiol Biotechnol* **64**: 505–514.
- Osawa T, Matsubara Y, Muramatsu T, Kimura M, Kakuta Y. 2005. Crystal structure of the alginate (Poly α -L-guluronate) lyase from *Corynebacterium* sp. at 1.2 Å resolution. *J Mol Biol* **345**: 1111–1118.
- Park D, Jagtap S, Nair SK. 2014a. Structure of a PL17 family alginate lyase demonstrates functional similarities among exotype depolymerases. *J Biol Chem* **289**: 8645–8655.
- Park DY, Chi WJ, Park JS, Chang YK, Hong SK. 2014b. Cloning, Expression, and Biochemical Characterization of a GH16 β -Agarase AgaH71 from *Pseudoalteromonas hodoensis* H7. *Appl Biochem Biotechnol* **175**: 733–747.
- Pluvinaige B, Hehemann JH, Boraston AB. 2013. Substrate recognition and hydrolysis by a family 50 exo-beta-agarase, aga50D, from the marine bacterium *Saccharophagus degradans*. *J Biol Chem* **288**: 28078–28088.
- Potin P, Richard C, Rochas C, Kloareg B. 1993. Purification and characterization of the α -agarase from *Atteromonas agarlyticus* (Cataldi) comb. nov., strain G J1B. *Eur J Biochem* **599**–607.
- Potin P, Sanseau A, Le Gall Y, Rochas C, Kloareg B. 1991. Purification and characterization of a new κ -carrageenase from a marine *Cytophaga*-like bacterium. *Eur J Biochem* **201**: 241–7.
- Rebuffet E, Barbeyron T, Jeudy A, Jam M, Czjzek M, Michel G. 2010. Identification of catalytic residues and mechanistic analysis of family GH82 ι -carrageenases. *Biochemistry* **49**: 7590–7599.
- Rebuffet E, Groisillier A, Thompson A, Jeudy A, Barbeyron T, Czjzek M, Michel G. 2011. Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ Microbiol* **13**: 1253–1270.
- Schroeder DC, Jaffer MA, Coyne VE. 2003. Investigation of the role of a β (1-4) agarase produced by *Pseudoalteromonas gracilis* B9 in eliciting disease symptoms in the red alga *Gracilaria gracilis*. *Microbiology* **149**: 2919–2929.
- Shi YL, Lu XZ, Yu WG. 2008. A new β -agarase from marine bacterium *Janthinobacterium* sp. SY12. *World J Microbiol Biotechnol* **24**: 2659–2664.
- Sugano Y, Matsumoto T, Noma M. 1994. Sequence analysis of the agaB gene encoding a new β -agarase from *Vibrio* sp. strain JT0107. *Biochim Biophys Acta* **1218**: 105–108.
- Sugano Y, Terada I, Arita M, Noma M, Matsumoto T. 1993. Purification and Characterization of a New Agarase from Marine Bacterium, *Vibrio* sp. Strain JT0107. *Med Mycol* **59**: 1549–1554.
- Sulzenbacher G, Bignon C, Nishimura T, Tarling CA, Withers SG, Henrissat B, Bourne Y. 2004. Crystal structure of *Thermotoga maritima* α -L-fucosidase: Insights into the catalytic

- mechanism and the molecular basis for fucosidosis. *J Biol Chem* **279**: 13119–13128.
- Takagi E, Hatada Y, Akita M, Ohta Y, Yokoi G, Miyazaki T, Nishikawa A, Tonozuka T. 2015. Crystal structure of the catalytic domain of a GH16 β -agarase from a deep-sea bacterium, *Microbulbifer thermotolerans* JAMB-A94. *Biosci Biotechnol Biochem* **79**: 625–632.
- Tao X, Wu J, Kim KS, Jing W, Lee YC. 2008. Cloning, Expression and Characterization of β -agarase Gene from a Marine Bacterium, *Pseudoalteromonas* sp. JT-6. *J Life Sci* **18**: 625–630.
- Thomas F, Barbeyron T, Tonon T, Génicot S, Czjzek M, Michel G. 2012. Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine Flavobacteriia to their independent transfers to marine Proteobacteria and human gut Bacteroides. *Environ Microbiol* **14**: 2379–94.
- Thomas F, Lundqvist LCE, Jam M, Jeudy A, Barbeyron T, Sandström C, Michel G, Czjzek M. 2013. Comparative characterization of two marine alginate lyases from *Zobellia galactanivorans* reveals distinct modes of action and exquisite adaptation to their natural substrate. *J Biol Chem* **288**: 23021–23037.
- Wang L, Li S, Zhang S, Li J, Yu W, Gong Q. 2015. A new κ -carrageenase CgkS from marine bacterium *Shewanella* sp. Kz7. *J Ocean Univ China* **14**: 759–763.
- Wu HJ, Ho CW, Ko TP, Papat SD, Lin CH, Wang AHJ. 2010. Structural basis of α -fucosidase inhibition by iminocyclitols with K_i values in the micro- to picomolar range. *Angew Chemie - Int Ed* **49**: 337–340.
- Xie W, Lin B, Zhou Z, Lu G, Lun J, Xia C, Li S, Hu Z. 2012. Characterization of a novel β -agarase from an agar-degrading bacterium *Catenovulum* sp. X3. *Appl Microbiol Biotechnol* **97**: 4907–4915.
- Xu F, Dong F, Wang P, Cao HY, Li CY, Li PY, Pang XH, Zhang YZ, Chen XL. 2017. Novel molecular insights into the catalytic mechanism of marine bacterial alginate lyase AlyGC from polysaccharide lyase family 6. *J Biol Chem* **292**: 4457–4468.
- Xu X, Li S, Yang X, Yu W, Han F. 2015. Cloning and characterization of a new κ -carrageenase gene from marine bacterium *Pseudoalteromonas* sp. QY203. *J Ocean Univ China* **14**: 1082–1086.
- Yang JI, Chen LC, Shih YY, Hsieh C, Chen CY, Chen WM, Chen CC. 2011. Cloning and characterization of β -agarase AgaYT from *Flammeovirga yaeyamensis* strain YT. *J Biosci Bioeng* **112**: 225–232.
- Zhang WW, Sun L. 2007. Cloning, characterization, and molecular application of a beta-agarase gene from *Vibrio* sp. strain V134. *Appl Environ Microbiol* **73**: 2825–2831.
- Zverlov VV, Volkov IY, Velikodvorskaya T V, Schwarz WH. 1997. *Thermotoga neapolitana* bglB gene, upstream of lamA, encodes a highly thermostable β -glucosidase that is a laminariobiase. *Microbiology* **143**: 3537–3542.

Résumé

Les polysaccharides algaux sont une ressource marine dont la valorisation est limitée par le manque d'outils adéquats à leur modification chimique. Grâce aux enzymes qu'elles synthétisent, les bactéries marines, qui vivent au contact des algues, représentent un enjeu majeur pour le développement des biotechnologies bleues. Cette thèse a permis de cloner et d'obtenir sous forme soluble plusieurs dizaines de nouvelles enzymes bactériennes marines actives sur des polysaccharides algaux. Quatre d'entre elles ont fait l'objet d'une étude plus spécifique. Les deux premières sont impliquées dans les voies de dégradation des carraghénanes. La première, ZgCgkA, est une κ -carraghénase de la famille 16 des glycosides hydrolases (GH), synthétisée par la bactérie *Zobellia galactanivorans*. Son étude biochimique et structurale, par cristallographie des rayons X, a permis de corréliser certaines différences structurales à des modes d'interaction différents avec le substrat, au sein de cette sous-famille des GH16. La seconde enzyme étudiée, une β -carrabiose hydrolase de *Pseudoalteromonas carrageenovora*, agit sur des oligosaccharides hybrides β/κ . Son étude biochimique et phylogénétique a permis de proposer la création d'une nouvelle famille de GH apparentée aux GH42. Enfin, les deux dernières enzymes étudiées, une GH29 et une GH non-classée, sont codées dans un locus de *Z. galactanivorans* qui semble impliqué dans la dégradation de substrats enrichis en fucose sulfaté. La GH29 a fait l'objet d'une caractérisation biochimique sur substrat synthétique, et son analyse structurale est en cours. Ces différents résultats permettent d'envisager l'utilisation de ces enzymes comme outils de valorisation des polysaccharides algaux, en particulier des carraghénanes.

Abstract

Algal polysaccharides are marine resources valorization of which is hindered by the lack of proper tools to modify their structure. Marine bacteria living associated to macroalgae synthesize enzymes acting on these polysaccharides. They represent a great opportunity for the development of blue biotechnology. This thesis project resulted in the successful cloning and soluble protein production of several dozen new bacterial enzymes active on algal polysaccharides. Four of them have been studied in detail. The first two are involved in carrageenan degradation pathways. The first one, ZgCgkA, is a κ -carrageenase from family 16 of glycoside hydrolases (GH), synthesized by *Zobellia galactanivorans*. Its biochemical and structural study, by X-ray crystallography, provided a link between structural features and different interaction modes with the substrate in this GH16 sub-family. The second enzyme, a β -carrabiose hydrolase from *Pseudoalteromonas carrageenovora*, is active on hybrid oligosaccharides of β/κ -carrageenan. Its biochemical and phylogenetic study suggests the creation of a new GH family, distantly related to the GH42 family. The last two enzymes, a GH29 and a non-classified GH, are encoded in a locus of *Z. galactanivorans* probably dedicated to the degradation of sulfated fucans. GH29 was biochemically characterized on synthetic substrate, and its structural study is ongoing. These results raise the possibility to use these enzymes as tools for the valorization of algal polysaccharides, particularly carrageenans.