Catabolisme des agars chez Zobellia galactanivorans

Anaïs Naretto

▶ **To cite this version:**

Anaïs Naretto. Catabolisme des agars chez Zobellia galactanivorans. Biochimie [q-bio.BM]. Sorbonne Université, 2018. Français. NNT : 2018SORUS498 . tel-03018586

HAL Id: tel-03018586
https://theses.hal.science/tel-03018586

Submitted on 23 Nov 2020

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.

CNRS • SORBONNE UNIVERSITÉ
Station Biologique de Roscoff

Sorbonne Université

Ecole doctorale 515 : Complexité du vivant

Laboratoire de Biologie Intégrative des Modèles marins / Equipe Glycobiologie Marine
Unité Fonctionnalité et Ingénierie des Protéines /
Equipe Ingénierie moléculaire et Glycobiologie

Catabolisme des agars chez Zobellia galactanivorans

Par Anaïs Naretto

Thèse de doctorat de Biochimie et Biologie structurale

Dirigée par Dr. Gurvan Michel et le Pr. Charles Tellier

Présentée et soutenue publiquement le 21 Novembre 2018

Devant un jury composé de :

Dr. Solange MORERA (DR CNRS) Rapportrice
Dr. Jean-Guy BERRIN (DR INRA) Rapporteur
Dr. Hélène ROGNIAUX (IR INRA) HDR Examinatrice
Pr. François LALLIER (Pr. SU) Examinateur
Pr. Charles TELLIER (Pr. Université de Nantes) Co-directeur de thèse
Dr. Gurvan MICHEL (DR CNRS) Directeur de thèse

Remerciements

Et bien voilà le moment que j'ai le plus redouté, celui des remerciements. En effet, il s'agit de l'exercice le plus inconfortable pour moi, puisqu'il m'incombe de n'oublier personnes (et pourtant que nous sommes nombreux à la station et je n'ai guère de mémoire, n'est-ce pas Sonia ?), de plus verbaliser mes sentiments et ce que je pense ne m'est pas aisé. Quelle pression !!! Et oui, j'ai bien conscience que cette partie est surement l'une des plus attendue…
Pour toutes ces raisons, je demande à mes lecteurs d'être indulgents.
Commençons par le plus facile, les remerciements « officielles » :

Je tiens à remercier Catherine Boyen de m'avoir accueillie au sein de son UMR. Je tiens à exprimer mon admiration pour ta mémoire des noms et pour le fait que tu prends toujours le temps de dire un petit mot à tout le monde.

Un grand merci aux membres du jury pour avoir accepté de juger ce travail, ce qu'ils ont fait avec bienveillance et compétence (non je ne fayote pas). Merci : à François Lallier de s'être libéré au dernier moment pour présider le jury, à Jean-Guy Berrin pour avoir accepté le rôle de rapporteur (qui dans mon cas ne fut pas de tous repos), à Hélène Rogniaux pour avoir fait parti du jury mais aussi pour le fantastique travail qu'elle et tous les membres de la plateforme de spectrométrie de masse de Nantes ont accompli (oui je finis de rédiger les remerciements et je te contacte pour l'article ou les articles … on peut rêver n'est-ce pas ? ;-)), et j'ai gardé Solange Moréra pour la fin car elle occupe une place un peu a part. En effet, elle a acceptée d'être rapportrice de la thèse et à ce titre elle a souffert, comme Jean-Guy, de ma grammaire calamiteuse, mais c'est aussi celle qui m'a initiée à la cristallographie des protéines et donc qui m'a enseigné toutes les bases pratiques de ce domaine (et de comment ne pas faire de bulle dans mes gouttes) et pour tout ceci ainsi que pour ta bienveillance et ta gentillesse encore merci. J'espère te revoir souvent (cette dernière phrase vaut pour les autres membres du jury).

Passons au chefs (ffes) direct :

Merci à Gurvan et à Mirjam de m'avoir choisie pour cette thèse et acceptée au sein de l'équipe Glycobiologie Marine.
Gurvan tu m'as laissé en autonomie, tu m'as fait confiance pour la mise au point de nombreux protocoles et pour quasiment gérer le projet. Merci de m'avoir laissé lancer et animer la formation enzymologie. Je me souviendrais toujours de ton très grand enthousiasme à chaque fois que je te ramenais un résultats ce qui rachetais mon scepticisme naturel. Tu m'as écrit l'article *Zg*AgaC à partir d'un draft que je qualifierais au mieux de préliminaires… Merci de m'avoir toujours gentiment écouté lorsque je te parlais de mes ambitions post-doctorales. Ainsi merci de m'avoir fait confiance et de ton implication dans le projet.

Mirjam, je ne sais quoi te dire… Tu as supportée avec un stoïcisme et une grande bonté mes crises de stress, de peur et mon ras le bol de fin de thèse. Merci pour tout ce que tu m'as appris (oui je suis très fière d'affirmer savoir générer des self Patterson

maintenant. On se vante de ce qu'on peut, n'est-ce pas ?) Merci d'avoir été là, de m'avoir écouté, remonté le morale. Merci d'être resté tard le soir quand j'ai cru que le jeu de donnée de *Zg*3515 sélénié avait disparu. Merci pour tous les week-ends, soirées, vacances que tu as sacrifié à mon profit à la fin de la rédaction de ma thèse alors que tu n'étais pas obligée de le faire. Merci d'avoir pensé à acheter des mouchoirs en grande quantité et de m'avoir laissé humidifier ton bureau…

Charles (ou devrais-je dire piou-piou émérite ou chefchef ?), je tiens à te dire que j'ai adoré ta présence au labo pour manipuler avec moi, nos discussions et ton humour pince sans rire… Je suppose que tu le sais mais tu es redoutable quand il s'agit de plaisanterie à froid. J'ai adoré découvrir et mettre au point des techniques avec toi, mais aussi de discuter de plein de choses diverses, tu es un véritable puit de science, et ça va me manquer de ne plus pouvoir y puiser, ou comme je l'ai dit devant ton équipe de façon fort peu élégante une tronche (Si, si c'est un compliment je t'assure). Me vient à l'esprit pour te qualifier une vieille expression de l'ère victorienne : une brique. Je n'ai qu'un regret te concernant, ne pas avoir pu passer encore plus de temps avec toi. En fait j'en ai un deuxième : ne pas avoir discuté plus tôt d'escalade avec toi, nous aurions pu nous faire une petite sortie grimpette ;-).

Et maintenant il s'agit des gens de l'équipe, par où commencer ? Disons dans l'ordre chronologique :

Les deux personnes que j'ai rencontrées étaient Sabine et Agnès (et oui j'ai partagé leur bureau pendant quelques jours…). J'étais un peu perdue et pas très à l'aise dans ce nouvel environnement, merci d'avoir été là pour ces débuts.
Merci à Agnès et Sabine pour le chocolat, les biscuits (et les fantastiques gaufres de Sabine) mais aussi pour tous les coups de main, les discussions et les conseils que vous m'avez prodigués. Sabine je n'oublierais jamais notre voyage au pays de l'oncle Sam et je suis heureuse que ce soit bien fini, cette expérience en ta compagnie a été géniale. Merci de m'avoir aidé cette dernière année, notamment pour le mat et meth du présent manuscrit mais aussi pour le déménagement et pour toutes tes petites attentions.
Agnès tu m'as beaucoup manqué après ton départ, et je te remercie de m'avoir accueillie à Nantes et soutenue pendant cette période difficile. J'espère que vous serez heureux à Nantes.

Murielle ou Mumu… Que dire ? Merci de m'avoir présenté tout le labo, de m'avoir formé sur plusieurs appareils mais aussi éclairé sur les usages du labo et de m'avoir initié à cette technique fantastique que sont les FACE (oui il y a une pointe de sarcasme). Merci de toujours gérer les stocks pour qu'il n'y ait jamais de rupture. Merci de m'avoir accompagné lors de cette sortie mémorable sur l'estran où nous avons trouvé cette algue de 12 mètre de long et où tu as perdu la peau du visage (et moi la peau des oreilles) à cause d'un vilain coup de soleil. Finalement, le sacrifice de nos épidermes a permis de trouver des activités enzymologiques…

Alex (alias poulette, maman ou petit berserker adoré), notre première rencontre fut comment dire… abrupte. En effet, tu as une franchise folle et un caractère fort (ou de cochon comme tu préfère, c'est surement pour ça que l'on s'entend bien) mais surtout un cœur en or. Tu m'as montré plein de choses en cristallogenèse (et tu m'as même laissé ton labo en garde, une grande marque de confiance de ta part), tu m'as aidé à me

faire mon trou à la station. Tu as été comme une grande sœur pour moi et je ne l'oublierais jamais, je n'en dirais pas plus cela risque de devenir gênant. J'ai hâte de venir te voir toi et ta famille à Norfolk.

En réalité, j'ai rencontré Elizabeth avant Mumu, mais soyons honnête nous n'avons pas discuté lors de cette première rencontre. J'était derrière une porte que Liz a ouvert de façon énergique et enthousiaste (deux adjectifs qualifiant très bien Liz) et j'ai finis encastré dans un bureau (au sens propre) avec un tympan en moins Elizabeth venant de nous annoncé l'arrivée de sa sœur en France. En dehors des portes, ce qui frappe chez Liz, c'est son sourire et son rire si communicatif. Liz tu es un poème d'une gentillesse folle, d'une énergie débordante et d'une excentricité rafraichissante. Merci d'avoir été là, de m'avoir corrigée, écoutée et soutenue. Je suis très heureuse pour ton recrutement et je te souhaite tout le bonheur possible.

Je ne sais plus qui de François ou de Cécile j'ai rencontré en premier alors je vais faire un tir groupé. Vous avez tous les deux étés d'un secours fou lors de cette fin de thèse. Cécile a essuyée, comme Mirjam, quelques crises de larmes et a été d'une écoute et d'un réconfort admirable. Merci de m'avoir éclairé sur les parois algales qui était *terra incognita* pour moi. Merci pour tes conseils en tout poil. Si ma soutenance a été aussi appréciée, c'est en grande partie grâce à François qui m'a donné des conseils très précieux. Merci pour l'opération impression de la thèse en urgence.

Celui qui m'a transmis le flambeau des sulfatases, c'est Tristan. Merci d'avoir corrigé ma thèse, de m'avoir éclairé sur la phylogénie (même si pour être honnête je trouve toujours cela fort obscur), de m'avoir prêter ta paillasse pendant trois ans.

Merci à notre « responsable informatique » : Lionel. Merci pour la correction du manuscrit, pour tes conseils multiples (surtout sur l'Akta Avant) et pour m'avoir débloquer plusieurs fois mon ordi et aussi pour ton aide active pour la recherche de post-doc. Et aussi pour la vidéo de thèse.

Merci à Robert pour tous les clonages, à Ludo, à Diane, à Thomas, à Kévin et à tous ceux que j'ai pu oubliés. Il est vrai que je ne vous écrit pas un grand paragraphe mais cela ne diminue en rien le fait que vous avez tous joué un rôle dans ma thèse mais aussi dans mes trois dernières années et pour cela je vous remercie.

Voici le tour des « autres » :

Anne-Catherine tu vas me manquer. Que de discussions, de repas, verres, cafés partagés !! De quoi n'avons nous pas discuté, débattu ? Je ne saurais le dire. Arrivée un peu après moi à la station, nous nous sommes trouvés par esprit grégaire (et oui les structuralistes se sont quelque peu regroupés). Tes anecdotes, ta gentillesse, ton écoute et aussi tes conseils avisés vont me manquer… Brisons là, tout ne se dit pas. (PS : et merci pour la cafetière qui à permis à la junky que je suis d'avoir suffisamment de dose pour être fonctionnelle).

Passons à mon binôme, et mentor, de TD : Jean. Anecdote : une thésarde naïve veut aider ses collègues à se rafraichir la mémoire en enzymologie en faisant une présentation d'une heure ou deux. Les collègues en parlent et la thèsarde se retrouvent avec une

trentaine de personne intéressées par suffisamment de sujet pour occupés deux jours de formation. Et là tel un chevalier de la connaissance accompagné de son fidèle cartable marron est apparut Jean pour sauver la thèsarde dépassée. C'est comme ça que nous avons commencé a donné des cours ensemble. Merci, de m'avoir appris à faire un cours et à le gérer. Merci de m'avoir aidé dans cette entreprise folle et de m'avoir aidé à m'épanouir dans le rôle ô combien difficile du professeur. Mais aussi d'avoir toujours été là pour discuter de mes résultats et de tout et rien. Monsieur vous êtes le genre de professeur à susciter des vocations et je vous admire pour ça.

Merci à Philippe pour toutes les fois où tu m'as accompagné sur l'estran chercher des algues, où tu les as identifiées pour moi, j'aurais été complétement perdue sans ton aide. Merci également pour ta gentillesse et ton écoute.

Thomas Robert, mon acolyte de screen et de mécanisation. Merci de m'avoir fait suffisamment confiance pour accepter de te lancer dans cette aventure du crible sulfatase. Merci pour tes logiciels, plasmides, machines… Merci pour ce que tu m'as dit avant mon départ. Je ne t'oublierais pas.

Durant ma thèse, j'ai investi de nombreux espaces… Ainsi, j'ai été adopté par l'équipe dites des « oursins ». Je vous remercie de m'avoir accueillie et permis d'utiliser leurs équipements. Merci à Virginie de m'avoir appris le clonage, à Julia pour nos blagues (au fait tu me dois toujours des cafés ;-)), à Patrick pour nos conversations du midi, à Agnès pour nos échanges et Sandrine notre conteuse d'histoire préférées. J'ai également passé un temps certains à l'USR, où Blandine, Stéphane, Natalie, Béatrice… m'ont accueillie, formée (pauvre Nat qui a du me faire découvrir les western blot, je crains de ne pas avoir été très douée…) et écoutée, merci à vous. Pour finir, j'ai investi les locaux de l'UMR7144 et comme seul leur autoclave me permettait de faire mes extractions je leur présente ici mes excuses pour les avoir enfumé de façon nauséabonde), merci à Domi, Natalie, Estelle, Fabienne et tous les autres…

Je tiens aussi à remercier toutes les personnes qui permettent au laboratoire de fonctionner, je pense aux services d'entretien (Marie-Noelle, Martine…), de restauration(Barbara, Frank, et tous les membres du Gulf Stream) et à l'administration. Un grand merci à Maryvonne et Sophie pour toutes les fois où elles m'ont aidé avec les papiers, les demandes de papier… Vous êtes indispensables, merci pour tout.

Voici venues le temps des amies, ou plutôt les « jeunes », bien que les autres ne soit pas vieux !!!

Merci à Laure qui est la première à m'avoir accueillie, à Amandine (la prochaine fois nous irons voir l'éléphant ;-)), à Hervé (bon courage pour ta nouvelle vie), à Simon (j'espère que tu auras le job à Kansas City), à Maria (quel dommage que je n'ai pas discuté plus tôt avec toi), à Margot (ou devrais-je dire Plankton belette), à Sonia (ce coup ci, je ne pense pas que tu me suivras ;-)), à Omid (je regrette que nous n'ayons pas passée plus de temps ensemble, embrasse tes femmes pour moi), à Jung (j'envie ta capacité d'être toujours souriante), à Jocelyn (bébé Jojo), à Magda, à Claire (j'espère que tu seras heureuse et que tu auras un beau bébé en pleine santé).
A tous ceux qui sont partis avant moi et qui me manque : Aline, Bertille, Marie, Emilie, Céline, Agnès et Léa. Vous avez tous compté pour moi, et vous allez tous me manquer. A

tous merci, pour les verres, les soirées, les conversations, les blagues (non, je n'ai jamais cousu de blouse, je le jure), les cafés et tous les moments partagés. Il serait trop long ici, de décrire tous ce que nous avons vécu sachez juste que sans vous ces années auraient été très longues et auraient singulièrement manqué de joie et d'amusement.

Je n'oublie pas mes amies connues bien avant tout ceci : Camille (petit paté), Ornella (ma petite Louise), Claire, Dafné, Mélu, JC (pourrais-tu m'écrire la recette de ta perfection ?)… Je ne vous le dis pas souvent mais je vous adore tous et vous êtes très important pour moi même si donner des nouvelles ne fait pas parties de mes capacités.

Je me dois aussi de remercier mes parents qui malgré les grandes difficultés rencontrés ont toujours tout sacrifié pour me permettre de faire des études. Heureusement, vous n'avez jamais écouté les personnes (professeur, éducateur et génétiquement proches) qui à cause de mes singularités prédisaient que je n'aurais pas d'avenir. Vous m'avez poussée à continuer, merci.
A ma sœur, je ne sais que dire si ce n'est qu'il est temps que tu vives pour toi et pour ce tu aime et pas en cherchant à prouver quoi que ce soit à qui que ce soit.
Merci à ma belle-famille de m'avoir accueillie et d'avoir été là, merci à Cid et Alex. J'espère que Titi fera bientôt ces nuits.

Je crois que je n'oublie person… Ah et bien sur John. Nous n'avons pas toujours eu des moments faciles, surtout ces derniers temps, mais si tout était facile où serait le plaisir ? La vie est une aventure autant la vivre jusqu'au bout, je pense que nous aurons bien d'autres moments de joie, d'aventures, de difficultés, de découvertes mais nous serons deux dans cette aventure.

Table des matières

Table des abréviations 5

Chapitre I : Introduction 7

I. 1 Les algues rouges 8
I.1.A Qu'est-ce qu'une algue ? 8
I.1.B Caractéristiques morphologiques des Rhodophytes 9
I.1.B. a. Polysaccharides de réserve 11
I.1.B. b. Pigments 11
I.1.B. c. Organisation des chloroplastes 12
I.1.B. d. Synapses 12
I.1.B. e. Absence de système de propulsion 13
I.1.B. f. Cycle de reproduction 13
I.1.B. g. Les polysaccharides mucilagineux 14
I.1.B. h. Conclusion sur la systémique classique des Rhodophytes 14
I.1.C Taxonomie et classification des Rhodophytes 15
I.1.D Présentation des algues modèles 16
I.1.D. a Porphyra C. Agardh 17
I.1.D. b Polysiphonia elongata (Hudson) Sprengel 18
I.1.D. c Osmundea pinnatifida (Hudson) Stackhouse. 19
I.1.D. d Chondria dasyphylla (Woodward) C. Agardh 20
I.1.D. e Delesseria sanguinea (L.) Lamouroux 21
I.1.D. f Palmaria palmata (L) Kuntz 22
I. 2 Les agars : de l'algue à la caractérisation 24
I. 2. A. Biogénèse et organisation des parois 24
I. 2. A. a Présentation de la paroi algale 24
I. 2. A. b Fonctions de la paroi 29
I. 2. A. c Biosynthèse des polysaccharides matriciels chez les algues rouges : cas particuliers des agars 30
I. 2. B. Diversité des agars 41
I. 2. B. a. Historique des agars. 41
I. 2. B. b. Nomenclature des agars. 42
I. 2. B. c Structure chimique et tridimensionnelle de l'agarose. 45
I. 2. B. d Quel intérêt à l'étude systématique des agars modifiés ? 48
I. 2. B. e Introduction à la chémotaxonomie 49
I. 2. B. g Agars modifiés : défis et perspectives 52
I. 3 Les enzymes de modification des agars 54
I. 3. A. Les glycosides hydrolases 54
I. 3. A. a. Zobellia galactanivorans : Un organisme modèle pour l'étude du catabolisme des polysaccharides marins. 54
I. 3. A. b. La classification CAZy et généralités sur les glycoside hydrolases 55
I. 3. A. c. Les loci d'utilisation de polysaccharides 59
I. 3. A. d. Les glycosides hydrolases impliquées dans la dégradation des agars. 61
I. 3. B. Les sulfatases actives sur les polymères marins 69
I. 3. B. a. Famille S1 ou sulfatase formylglycine dépendante. 70
I. 3. B. b Activation des sulfatases par la Formylglycine Generating Enzyme (FGE) 70
I. 3. B. c Les sulfatases : mécanismes enzymatiques et structures 72
I. 4. Objectifs de la thèse 78

Chapitre II : Matériels et méthodes 81

II.1 Clonage des différents gènes cibles 82
II.1. A Vecteurs d'expression 82
II. 1. B Analyse bio-informatique et définition des séquences à cloner 84

II.1.C ADN Matrice ... *84*
II. 1. D Conception des amorces ... *84*
II.1. E Clonage ... *88*
II.1.F Transformation des souches d'E. coli ... *90*
II.2 Mutagénèse dirigée ... 90
II.3 Expression des différentes protéines cibles ... 90
II.4 Co-expression cytoplasmique de la *Zg*FGE et de *Zg*4660 chez *E.coli* ... 91
II.5 Co-expression périplasmique de *Zg*FGE et de *Zg*4660 chez *E.coli* ... 92
II.6 Purification des différentes protéines ... 92
II.6. A Purification de la sulfatase Zg4660. ... *92*
II.6. B Purification de Zg3515 et Zg3516 ... *93*
II.6. C Etude de la thermostabilité des protéines. ... *94*
II.7 Cinétiques des sulfatases sur substrats artificiels ... 94
II.8 Gels de polyacrylamide FACE ou « *Fluorophore Assited Carbohydrate electrophoresis* » ... 96
II.9 Criblage des substrats naturels des sulfatases (*Zg*4660 et *Zg*3515) ... 97
II. 10 Criblage de l'activité enzymatique de la GH30 (*Zg*3516) sur substrats artificiels et naturels ... 97
II. 11 Tests de *Zg*3515 et *Zg*3516 sur la fraction oligoporphyrane F50 ... 99
II. 12 Chromatographie sur couche mince (CCM) ... 99
II.13 Extraction et digestion des polysaccharides ... 100
II. 13. A Extraction ... *100*
II. 13. B Digestion enzymatique ... *101*
II.14 Tests des différentes chromatographies échangeuses d'anions. ... 101
II.15 Purification des oligosaccharides de *Polysiphonia elongata* et *Chondria dasyphylla* par chromatographie de type échangeuse d'anion ... 102
II.16 Cristallisation ... 103
II.17 Collecte des données et traitement ... 103
II.18 Résolution du problème de phase et affinement ... 104
II.19 Tests d'activation/inhibition de *Zg*3515 par différents oses. ... 104
II.20 Essai de *trans*-sulfatation avec *Zg*3515 ... 105

Chapitre III : Caractérisation d'une nouvelle sous-famille de GH16 ... 107

Chapitre IV : Caractérisation de la sulfatase *Zg*4660 ... 113
IV. 1 Introduction ... 114
IV. 2 Résultats ... 116
IV. 2. A Contexte génétique de la sulfatase Zg4660 ... *116*
IV. 2. B Production chez Escherichia coli d'une sulfatase recombinante active et caractérisation cinétique avec un substrat artificielle ... *116*
IV. 2. C Criblage de l'activité sulfatase sur des oligosaccharides issus de diverses algues rouges agarophytes ... *119*
IV. 2. D Extraction de l'agar de Polysiphonia elongata ... *123*
IV. 2. E Purification des oligosaccharides issus de Polysiphonia elongata et tests d'activité sulfatases ... *123*
IV. 2. F Tests pour reproduire l'enzyme Zg4660 sous forme active ... *129*
IV. 2. G Analyse des biais de codons présents au sein des séquences codant pour les sulfatases ... *130*
IV. 2. H Cristallisation de Zg4660 ... *131*
IV. 3 Conclusions et Perspectives ... 134

Chapitre V : Couple *Zg*3515 (sulfatase S1_20) et *Zg*3516 (GH30) ... 137
V. 1 Introduction ... 138
V. 1. A La famille 30 des glycoside hydrolases (GH30) ... *140*
V. 1. B Les divers polysaccharides retrouvées chez les algues brunes et vertes ... *142*
V. 2 Résultats ... 146
V. 2. A Etude de la glycoside hydrolase Zg3516 ... *146*

V. 2. A. a Purification et contexte génétique de la GH30_4 Zg3516 ..... *146*
V. 2. A. b Recherche des substrats naturels de Zg3516 ..... *148*
V. 2. A. c Caractérisation de l'activité de Zg3516. ..... *150*
V. 2. A. d Structure de Zg3516 ..... *153*
V. 2. A. e Obtention d'un complexe Zg3516-oligoporphyranes *159*
V. 2. B. Etude de la sulfatase Zg3515 ..... *161*
V. 2. B. a Purification de la sulfatase Zg3515 et contexte génomique *161*
V. 2. B. b Caractérisation cinétique de Zg3515 ..... *162*
V. 2. B. c Première évidence de l'activité de Zg3515 sur substrats naturels *163*
V. 2. B. d Purification et caractérisation des oligo-porphyranes, désulfatés par Zg3515 ..... *164*
V. 2. B. e Détection et caractérisation de l'activité sulfatase de Zg3515 sur des oligosaccharides provenant de diverses algues agarophytes. *167*
V. 2. B. f Cristallisation et résolution de la structure de Zg3515. ..... *170*
V. 2. B. h Premier test de trans-sulfatation avec Zg3515 ..... *182*
V. 3 Conclusions et perspectives 183

VI Conclusions et perspectives générales. 187
VI. 1 Principaux résultats 188
VI. 2 Perspectives générales. 191

VII. Bibliographie 195

Table des abréviations

AMAC : 2-aminoacridone

ANTS : acide 8-aminonaphthalene-1,3,6-trisulfonique

CAZymes : enzymes actives sur les carbohydrates

CBM : *Carbohydrate Binding Module* (module de liaisons aux sucres)

chl : chlorophylle

CHST : carbohydrate sulfo-transférase

CMT : carbohydrate méthyl-transférases

CPT : carbohydrate pyruvyl-transférase

EDTA : acide éthylène diamine tétracétique

EGTA : acide egtazique

FACE : électrophorèse de carbohydrates couplés à un fluorophore

FGE : *Formylglycine Generating Enzyme*

F6P : fructose-6-phosphate

GAG : glucan-1,4-α-glucosidase

GALT : galactose-1-phosphate urydylyl-transférase

GFP : *Green Fluorescent Protein*

GH : Glycoside hydrolase

GK : glucokinase

GLA : α-galactosidase

GME : GDP-mannose 3', 5'-épimérase

GMP : GDP-mannose phosphorylase

GP : α-1,4-glucane/glycogène phosphorylase

GPI : glucose-6-phosphate isomérase

GST : glutathion-S--transférase

GT : glucosyl--transférase

HPLC : chromatographie en phase liquide à haute performance

IPTG : isopropyl b-D-1 thiogalactopyranoside

LB : *Lysogeny broth* ou Luria-Bertani

MBP : Maltose binding protein

MPI : mannose-6-phosphate isomérase

MUF: méthylumbelliférone

MUS : méthyllumnelliféryl-sulfate

NST : *nucleotides sugar transporters*

PDB : *Protein data bank*

PGM : phosphoglucomutase

PMM : posphomannose mutase

PT : pyruvyl-transférases

RMN : résonnace magnétique nucléaire

SULT : sulfo-transférases

Sus : Starch utilization system

TBDR : récepteur TonB dépendant

UGE : UDP-glucose-4-épimérase

Zg : Zobellia galactanivorans

Chapitre I : Introduction

I. 1 Les algues rouges

I.1.A Qu'est-ce qu'une algue ?

La réponse de Bruno de Reviers, Professeur au Muséum National d'Histoire Naturelle:

Le terme 'algue' couvre un ensemble d'organismes photosynthétiques très divers (**Figure 1**). Certaines algues sont des bactéries (les algues bleues ou cyanobactéries) ; d'autres comme les algues rouges et vertes ou les glaucophytes sont des eucaryotes proches des plantes terrestres ; ou encore les algues brunes, les euglénophycées, les cryptophytes, les haptophytes, les chlorarachniophycées, les algues brun-doré et les dinophytes sont des eucaryotes sans lien de parenté direct avec les plantes. L'étude de l'ensemble de ces organismes a donné naissance à une discipline : la phycologie (du grec phycos, algue).

Si les algues ne constituent pas un taxon, elles possèdent malgré tout une caractéristique commune : ce sont soit des cyanobactéries, soit des organismes contenant des plastes (i.e. compartiments cellulaires sièges des réactions photosynthétiques) dérivant d'une cyanobactérie ancestrale. En effet, chez les algues eucaryotes, le plaste est soit dérivé d'une cyanobactérie (endosymbiose primaire), soit dérivé d'une algue eucaryotique unicellulaire dont le plaste est lui-même d'origine cyanobactérienne (endosymbioses secondaires et tertiaires) (Reyes-Prieto et al., 2007). La photosynthèse oxygénique basée sur l'utilisation de la chlorophylle *a* (pigment commun à toutes les algues et aux plantes) est une caractéristique propre aux cyanobactéries. Sa présence dans le plaste des algues eucaryotiques et celui des embryophytes est le résultat de ce phénomène d'endosymbiose. Les embryophytes ont divergés d'une algue verte charophyte et se sont adaptées au milieu terrestre (Vries and Archibald, 2018). Cependant, dans le langage courant, les algues sont définies comme étant des organismes photosynthétiques à chlorophylle *a*, autres que les embryophytes et généralement associés aux milieux aquatiques.

Le terme d'algue ne renvoie pas qu'aux macroalgues mais également à tout un ensemble d'organismes microscopiques extrêmement variés qui sont qualifiés de microalgues. Trois grandes lignées d'algues eucaryotes sont particulièrement étudiées : les algues brunes (Phaeophyceae), les algues vertes (Chlorophytes) et les algues rouges

(Rhodophytes). Les Glaucophytes, les Chlorophytes (y compris les plantes) et les algues rouges constituent le phylum des Archaeplastida (Hori and Osawa, 1987) (Adl et al., 2005).

Les algues rouges, également appellées Rhodophytes, sont très probablement l'une des plus anciennes lignées d'eucaryotes : elles seraient apparues il y a 1400 millions d'années (Reyes-Prieto et al., 2007), de plus la majorité des espèces de macroalgues sont des Rhodophytes. En effet, nous dénombrons plus d'espèces d'algues rouges (6400 espèces décrite http://www.algaebase.org) que d'espèces d'algues brunes et vertes confondues à l'échelle mondiale.

Les algues rouges sont très présentes dans les zones équatoriales et tempérées. A l'inverse, dans les milieux polaires et subpolaires les algues brunes prédominent. Une autre particularité des Rhodophytes est leur capacité à coloniser les niveaux les plus profonds du plancher océanique, jusqu'à 200 mètres de profondeur. Cette capacité provient directement de leur composition pigmentaire.

L'objet de cette thèse était l'étude de plusieurs enzymes de dégradation de polysaccharides issues des Rhodophytes, c'est pourquoi seule cette lignée sera présentée dans l'introduction.

I.1.B Caractéristiques morphologiques des Rhodophytes

Les macro-algues se présentent couramment sous la forme d'un thalle (i.e. appareil végétatif sans système vasculaire et où ne sont identifiés ni feuilles, ni racine, ni tige) (**Figure 1**). Mais de nombreuse exceptions existes à quant aux formes que peut prendre le thalle.

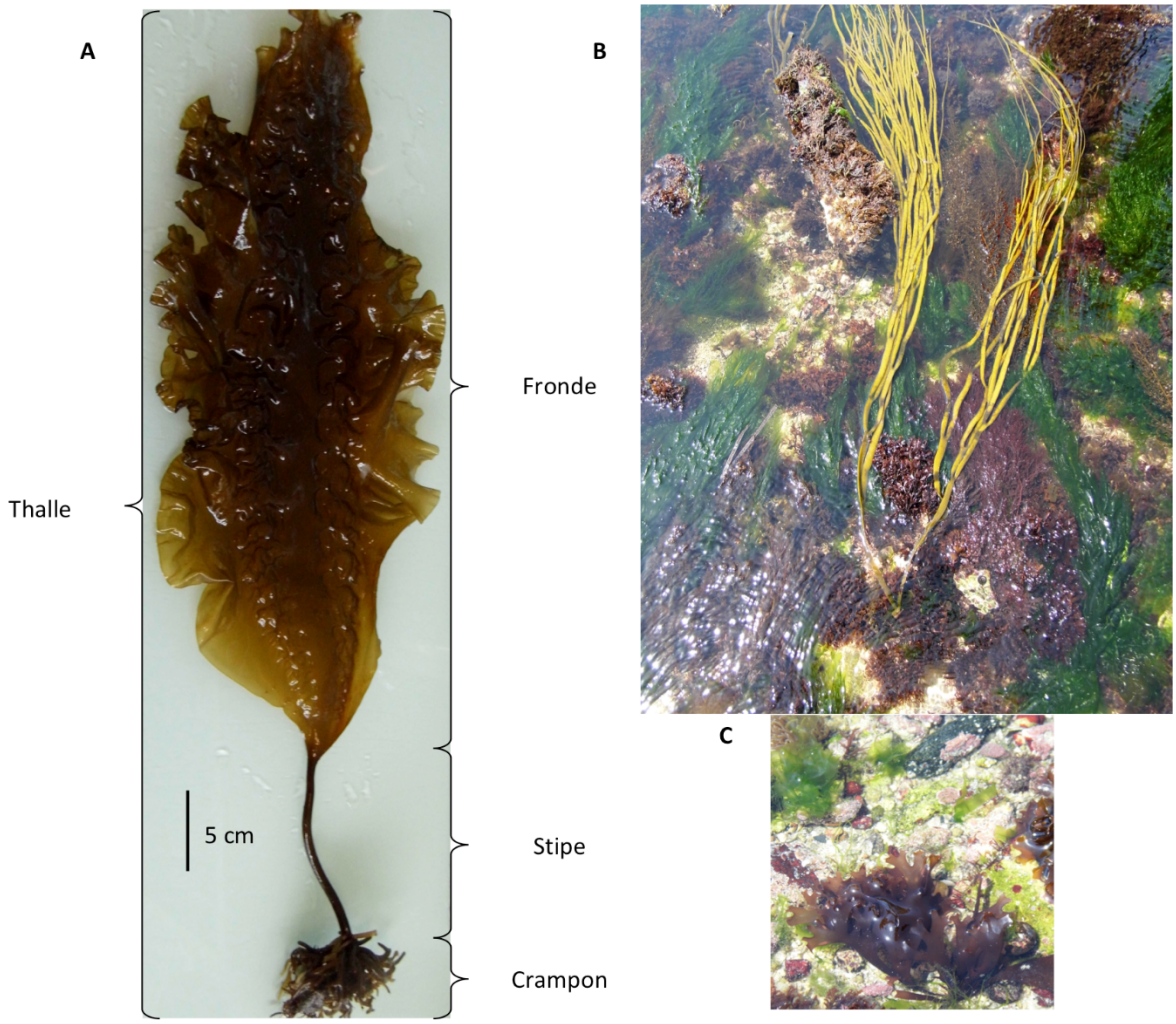


Figure 1 : Diversité morphologique des macro-algues.
La photo A est une algue brune *Saccharina latissima* sur laquelle sont mentionnées les différentes parties composant le thalle (stipe, crampon et fronde). Cependant, les algues comportent un très grand nombre de formes et de couleurs (photo B), il est donc difficile de retrouver ces éléments sur toutes les algues. Ainsi, l'algue rouge *Hildenbrandia rubra* se déposent sous forme de pellicules sur les pierres de la photo C, ce qui leur donnent cette coloration rouge, et pour ces dernières aucun stipe ou fronde n'est retrouvé (Photo : J. Dorival et A. Naretto)

En 1813, pour réaliser une première classification de ces organismes si divers en termes de formes que sont les algues, Lamouroux a proposé de les classer en fonction de leurs couleurs, ce classement est toujours en usage. En effet, les termes d'algues brunes, vertes et rouges restent largement utilisés. L'apparition de nouvelles technologies, comme le microscope, ont permis de définir d'autres caractéristiques, en sus de la pigmentation, pour différencier les Rhodophytes. Ces caractéristiques qui permettent la classification systémique des algues rouges sont présentées dans les paragraphes suivants. Toutes les informations des paragraphes I.1. B sont tirées de trois ouvrages de

références différents : *Biology of Red Algae* (Cole and Sheath, 1990), *Phycology* (Lee, 2008) et Biologie et phylogénie des algues (de Reviers, 2002).

I.1.B. a. Polysaccharides de réserve

L'amidon floridéen est le composé de stockage des Rhodophytes, ce dernier ressemble à l'amylopectine des plantes supérieures. Il s'agit d'un polymère de glucose lié en α-1, 4 présentant ponctuellement des branchements liés en α-1, 6. Contrairement aux plantes et aux algues vertes, l'amidon floridéen est synthétisé dans le cytoplasme et non dans les chloroplastes, où il se présente sous forme de grains ou d'écailles. L'autre différence est que son précurseur est l'UDP-glucose et non l'ADP-glucose comme chez les plantes (Steven and Matthew, 2003).

I.1.B. b. Pigments

La couleur rouge caractéristique de nombreuses Rhodophytes provient d'une combinaison de pigments intrinsèques à cette lignée : la chlorophylle *a* (*chl a*) et les phycobilosomes (complexes protéines-pigments). Bien que généralement seule la chlorophylle a (*chl a*) soit présente chez les algues rouges, une autre chlorophylle, la chlorophylle *d,* a également été retrouvée chez plusieurs espèces (Landráu and Welschmeyer, 2000).

Les phycobiliprotéines sont des protéines liées covalemment à un chromophore (biline) Un complexe de phycobiliprotéines constitue un phycobilosome, complexe qui est retrouvé en surfaces des thylakoides chez les Rhodophytes (Lee, 2008). Les chromophores associés aux phycobiliprotéines sont classés en fonction de leur spectre d'absorption, ainsi dans le cas des pigments bleu nous parlerons de phycocyanine et pour les pigments rouges de phycoerythrines (de Reviers, 2002).

Ces complexes permettent de transférer l'énergie lumineuse au centre réactionnel des photosystèmes II, elles remplacent les chlorophylles *b et c* retrouvées respectivement chez les algues vertes et brunes.

Cette composition en pigments apporte un avantage aux algues rouges en eaux profondes. En effet, au sein de la colonne d'eau la lumière devient essentiellement bleue / verte avec la profondeur or ce sont les longueurs d'onde qui sont absorbées par les phycobiliprotéines.

I.1.B. c. Organisation des chloroplastes

Les chloroplastes ne présentent pas de thylakoïdes agrégés, ni de réticulum endoplasmique situé sur la membrane externe. Les phycobilosomes se situent en surface des thyllakoides, ce qui empêche leur agrégation. Cette organisation est proche de ce qui est retrouver chez les cyanobactéries (**Figure 2**)(R. E. Lee, 2008).

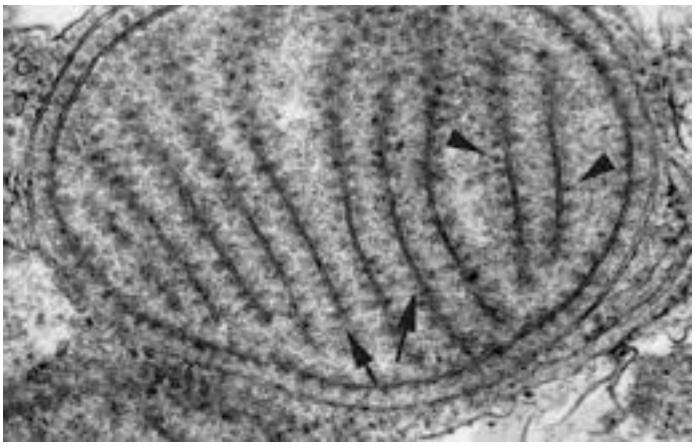


Figure 2 : Photo obtenue par microscopie électronique d'un chloroplaste de *Griffithsia pacifica*

Les thylakoïdes ne sont pas agrégés (flèches) et sur leurs membranes sont fixés des phycobilosomes (petits points noires) (tête de flèche). (Image de C. M. Pueschel tiré de tolweb.org/Rhodophyta)

I.1.B. d. Synapses

Les synapses (« *pit connection* » en anglais) sont présentes chez tous les Rhodophytes sauf exception. Il ne s'agit pas d'une connexion à proprement parler mais d'une séparation incomplète entre deux cellules filles du thalle. Cette brèche est comblée par le bouchon synaptique constitué d'un mélange complexe de protéines et de polysaccharides acides mis en place par le réticulum endoplasmique. Cette structure permet chez de nombreuses espèces d'algues un échange direct de métabolites entre les cellules (Cole and Sheath, 1990).

I.1.B. e. Absence de système de propulsion

Quel que soit le stade de vie des Rhodophytes, ces dernières ne possèdent ni centrioles ni flagelles (Cole and Sheath, 1990) privant les gamètes d'une propulsion directionnelle.

I.1.B. f. Cycle de reproduction

La plupart des algues rouges ont un cycle reproductif en trois étapes (ou trigénétique) : carposporophyte, tétrasporophyte et gamétophyte (de Reviers, 2002) (**Figure 3**).

Le tétrasporophyte va produire des spores haploïdes par divisions méiotiques: les tétraspores. Ces dernières vont se développer pour former un thalle haploïde qui va produire des gamètes mâles ou femelles, il s'agit du gamétophyte. Les spermaties (gamètes mâles sans flagelles) produits par le gamétophyte mâle (dans le cas où les gamétophytes mâles et femelles sont des individus différents) vont ainsi se fixer sur le trichogyne (i. e. structure surmontant la base de la cellule femelle) et féconder le carpogone (gamète femelle) qui est resté fixé sur la femelle gamétophyte. Comme les événements suivant la fécondation sont extrêmement divergents selon les espèces d'algues rouges, ils ont servis à l'établissement de la classification des Rhodophytes.

Le zygote diploïde résultant de la fécondation va se développer en un carposporophyte sur la femelle gamétophyte puis produire et relâcher des carpospores (diploïdes). Ces derniers vont germer pour former un thalle diploïde qui mènera à la formation de tétraspores après méiose (Saunders and Hommersand, 2004).

Le cycle de vie est un point crucial dans notre étude, car, en fonction du stade de l'algue, la composition polysaccharidique des parois peut varier.

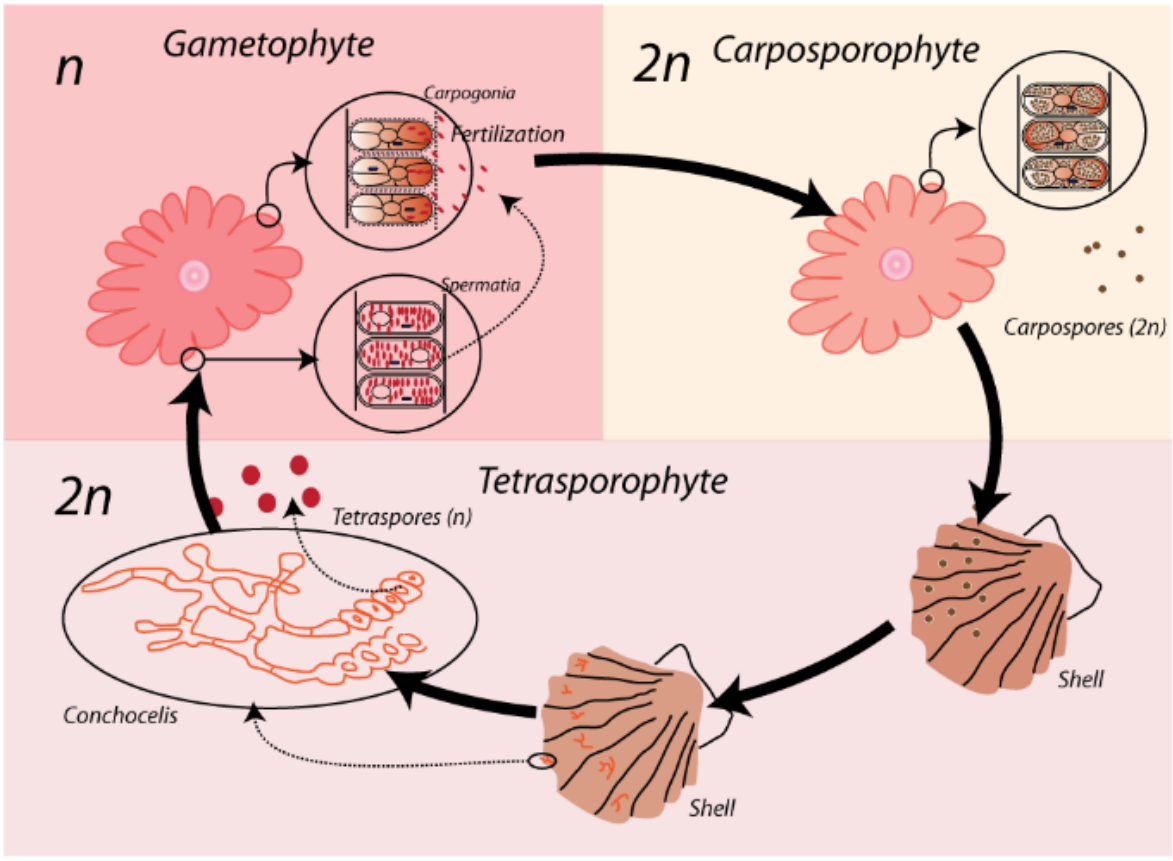


Figure 3 : Exemple de cycle de vie trigénétique avec *Porphyra sp*. (Extrait de (Bast, 2014)).

Le gamétophyte chez *Porphyra* correspond à un seul individu avec des zones mâles et des zones femelles et non pas des individus distinct. Le carpospore relâché va se fixer sur des coquillages (tels que les huitres) en perforant la coquille, puis se diviser par mitose pour former le tétrasporophyte. La particularité de *Porphyra* sp. est que le tétrasporophyte a longtemps été considéré comme une algue à part entière et est nommé conchocelis. Ce dernier va produire des tétraspores par méiose, qui vont ensuite donner le gamétophyte.

I.1.B. g. Les polysaccharides mucilagineux

Comme caractéristique propre aux Rhodophytes, nous pouvons ajouter que ce sont les seules algues à pouvoir produire les polysaccharides suivants : les carraghénanes et les agars. Néanmoins, toutes les algues rouges ne produisent pas ces galactanes sulfatés (Kloareg and Quatrano, 1988) (Popper et al., 2011).

I.1.B. h. Conclusion sur la systémique classique des Rhodophytes

Il est certain que ces éléments pris séparément ne sont pas tous spécifiques des Rhodophytes et que nous pourrions rajouter d'autres caractéristiques pour les

différencier des autres groupes d'algues. Néanmoins, la combinaison des caractéristiques précitées est unique à cette lignée, même s'il faut toujours se rappeler que de nombreuses exceptions existent. Ces difficultés et ces exceptions ont motivées la conclusion de Ragan et Gutell en 1995 « *in practice there is little difficulty in distinguishing what is or is not a red alga* » (Ragan and Gutell, 1995). Plus récemment, l'apparition de la génomique a révolutionné la classification et la taxonomie des algues en générale.

I.1.C Taxonomie et classification des Rhodophytes

Les Rhodophytes ne comportent qu'une seule classe les Rhodophyceae. La classification des algues rouges reposait historiquement sur l'appareil reproducteur du gamétophyte femelle et sur les événements suivant la fécondation (Schmitz, 1892) (Kylin, 1956). En 1982, Pueschel et Cole ont mis en évidence l'intérêt de la structure des synapses comme marqueur, ce qui a été confirmé par des analyses génétiques (Saunders and Kraft, 1997).

Depuis le début du 20ème siècle et jusqu'à récemment, ce phylum a été subdivisé en deux classes sur la bases de l'anatomie, de la morphologie et du cycle de vie des Rhodophytes : les Bangiophyceae et les Florideophyceae (Dixon, 1973).

Les Bangiophycées étaient supposées n'avoir ni synapses ni croissance apicale. Leur morphologie est généralement plutôt simple et elles sont considérées comme étant les formes les plus primitives d'algues rouges (Müller et al., 2001). Au contraire, les Florideophycées étaient décrites comme possédant des synapses, une croissance apicale ainsi qu'un cycle reproductif triphasique. Cette subdivision a été remise en cause dans les années 80 suite à la découverte d'exceptions à ces règles et à la mise en évidence du caractère non monophylétique des Bangiophycées (Gabrielson et al., 1985) (Saunders and Hommersand, 2004).

Désormais, la taxonomie repose sur les données phylogénomiques (comme l'analyse des données du 16S ribosomiques) et s'est complexifiée avec le temps. Dans la **Figure 4**, nous présentons la phylogénie la plus récente des Rhodophytes, et c'est cette dernière qui nous a servi de base pour décrire la diversité chimique des agars présenté dans la partie **I. 2. B. f**.

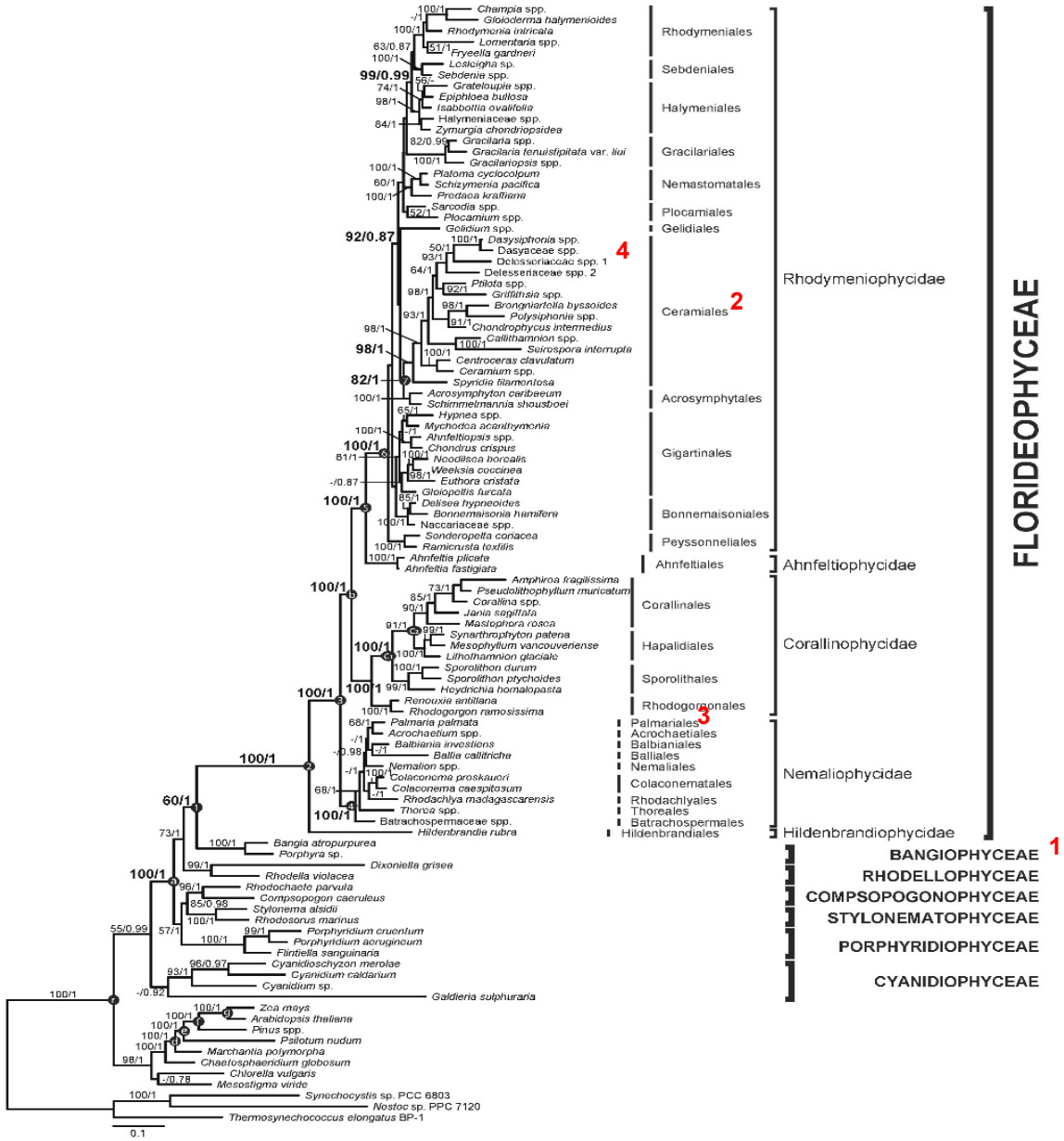


Figure 4 : Phylogénie des Rhodophytes (d'après (Yang et al., 2016))

La classe des Florophycées est toujours présente dans cette nouvelle classification et est racinée par la classe des Bangiophycées. Cette dernière a été redéfinie et de nouvelles classes sont apparues. En effet, cinq autres classes ont été définies en plus des deux historiques : les Rhodellophycées, les Compsopogonophycées, les Stylonematophycées, les Porphyridiophycées et les Cyanidiophycées. Au sein des Bangiophycées se trouvent les algues du genre *Porphyra* (signalé par 1). Les Florophycées contiennent les ordres des Palmariales (signalés par 3), des Ceramiales (signalés par 2) et la famille des Delesseriaceae (signalée par un 4) ce qui permet de repérer les algues utilisées lors de la thèse.

I.1.D Présentation des algues modèles

Dans la suite de ce manuscrit, nous parlerons en particulier des espèces suivantes : *Porphyra sp.* (Bangiales, Bangiaceae), *Polysiphonia elongata* (Ceramiales,

Rhodomélaceae), *Osmundea pinnatifida* (Ceramiales, Rhodomélaceae), *Chondria dasyphylla* (Ceramiales, Rhodomélaceae) et *Delesseria sanguinea* (Ceramiales, Delesseriaceae), espèces qui seront succinctement décrites dans les sous-parties suivantes.

Les algues rouges se présentent sous de très nombreuses formes et leurs couleurs sont extrêmement variables. En effet, en fonction de l'éclairement, elles peuvent paraître noirâtres, bleutées, jaunes… Leur recherche et leur identification n'en sont que plus complexes. Ainsi, les parties suivantes présentent les algues utilisées lors de ce travail ainsi que des clés d'identification illustrées.

I.1.D. a *Porphyra* C. Agardh

Le genre *Porphyra* est le genre le plus connu le plus connu et le plus consommé au monde, mais pas sous ce nom, qui a d'ailleurs changé au cours de la dernière décennie pour créer le genre *Pyropia* (Sutherland et al., 2011), mais sous la dénomination de *nori* (nom japonais). Il s'agit en effet de l'algue permettant la préparation des sushi makis. Il est difficile de distinguer visuellement les différentes espèces de *Porphyra*, quatre espèces sont référencées à Roscoff (*Porphyra umbilicalis*, *P. linearis* ; *P. dioica* ; *P. purpurea*), et deux espèces rattachées à présent au genre *Pyropia (Pyropia leucosticta et P. drachii*). Les thalles retrouvés dans la nature sont des gamétophytes, où les parties mâles et femelles se reconnaissent grâce à des zones décolorées ou rougeâtres respectivement. Elles se présentent sous la forme d'une lame composée d'une couche d'une cellule en épaisseur. (**Figure 5**) Ce genre produit l'un des agars sulfatés le mieux décrit à ce jour : le porphyrane.

Figure 5 : Photo de *Porphyra* sp.

A gauche : Photo de *Porphyra umbilicalis*, sur la partie supérieure, il est très facile de discerner l'épaisseur de l'algue. Au centre se trouve un ombilic central qui est la cause des replis. A droite *: Porphyra sp.* Sur l'estran, une fois émergée, elle est souvent retrouvée sur les pierres, recouvrant les autres algues, il s'agit d'une espèce épiphyte. (Photo : J. Dorival et A. Naretto)

I.1.D. b *Polysiphonia elongata* (Hudson) Sprengel

Cette algue rouge a une très large répartition géographique : en Europe, elle est trouvée de la Mer du Nord à la Mer Méditerranée. Elle vit fixée sur des rochers et possède une couleur variant du rouge vif au rouge brun caractéristique (**Figure 6**). Bien que pérennes, seuls les axes principaux sont visibles en hiver, alors qu'au printemps, la présence de fins filaments latéraux lui donnent un aspect en pinceaux, qui la rend très reconnaissable. Elle peut atteindre 30 cm de long. A ce jour plusieurs agars provenant du genre *Polysiphonia* ont été caractérisés mais pas celui de *P. elongata*.

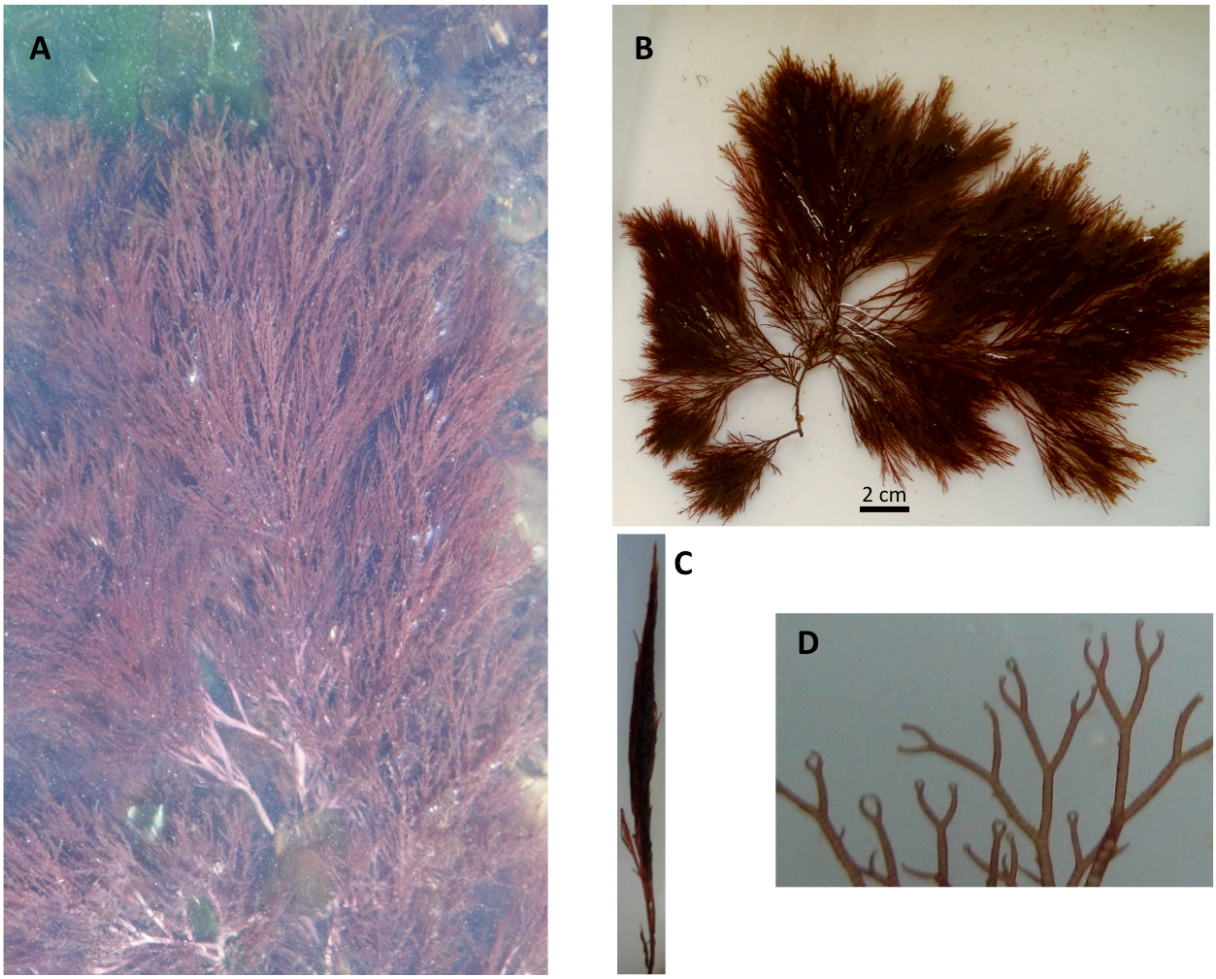


Figure 6 : *Polysiphonia elongata* et *Ceramium sp.*
La couleur de cette algue est très reconnaissable surtout dans l'eau où elle prend une teinte plus vive (A). Elle possède un axe principale cylindrique avec de multiples ramures qui se terminent eux même en de nombreuses petites ramures (B). Hors de l'eau, elles forment un pinceau légèrement cartilagineux (C). Cette algue peut facilement se confondre avec *Ceramium rubrum*, mais cette dernière possède de petits crochets à l'extrémité de ses filaments ce qui permettent de les différencier (D). (Photos J. Dorival et A. Naretto)

I.1.D. c *Osmundea pinnatifida* (Hudson) Stackhouse.

Appelée communément *pepper dulse*, cette espèce est connue des gastronomes pour son goût poivré. Elle se présente sous forme d'un coussin très dense fixé sur les rochers. Bien qu'étant une algue rouge, sa couleur varie beaucoup en fonction des conditions environnementales, et en particulier de la luminosité. Elle peut ainsi être brun-rouge très foncée, mais devenir jaune après photoinhibition. Ceci fait que la couleur ne saurait être un critère d'identification pour cette espèce (**Figure 7**). Elle est présente en Atlantique nord-est (de la Scandinavie à la Mauritanie), en Mer Méditerranée et en Mer Noire.

L'agar d'*Osmundea pinnatifida* (*Laurencia pinnatifida* selon l'ancienne classification) a été largement caractérisé par des techniques chimiques et par RMN, voir **I. 2. B. f**.

Figure 7 : *Osmundea pinnatifida*

Osmundea pinnatifida se fixe sur les rochers (espèce lithophyte) grâce à des stolons et forme alors des « coussins » (A). Les thalles de couleur rouge foncé (B) peuvent paraître jaune-orangé en cas de surexposition à la lumière (A). Les thalles sont aplatis et épais, et se présentent sous la forme d'un axe très ramifié (B). Les ramifications sont elles aussi aplaties et dentelées, et toutes situées dans un seul plan (C). (Photos J. Dorival et A. Naretto)

I.1.D. d *Chondria dasyphylla* (Woodward) C. Agardh

Chondria dasyphylla se présente sous forme de touffes rouges foncées de 10 à 20 cm de haut. Cette espèce est peu présentes sur l'estran de Roscoff, mais est représentée en Atlantique nord-est (des Iles Britanniques à la Mauritanie), en Atlantique nord-ouest et en Méditerranée.

L'agar de cette espèce n'a pas été caractérisé, alors que celui de *Chondria macrocarpa* a déjà fait l'objet d'études (voir **I. 2. B. f**). Notons que cette algue peut être confondue avec *Chondria scintillans* qui, elle, présente des reflets bleu iridescent (**Figure 8**).

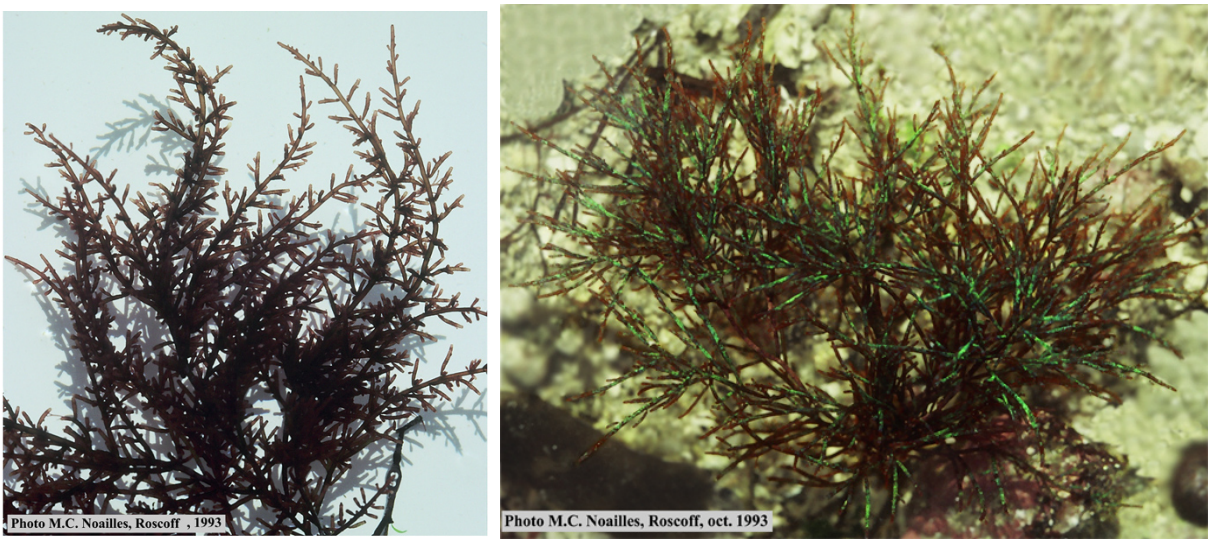


Figure 8 : *Chondria dasyphilla* et *C. scintillans*

Chondria dasyphylla, à droite, comporte un à deux axes cylindriques principaux, très ramifiés dans un seul plan. *Chondria scintillans* à gauche se distingue par sa taille plus réduite et surtout par ses reflets iridescents. (Photos M. C. Noailles)

I.1.D. e *Delesseria sanguinea* (L.) Lamouroux

D'une couleur rouge très vive ces algues sont malheureusement très rarement accessibles à pied et leur récolte nécessite l'intervention de plongeurs. En effet ces dernières se trouvent jusqu'à 30 mètres de profondeur où elles peuvent se trouver à l'ombre de grandes algues comme les laminaires. Elles possèdent un stipe cylindrique d'où partent plusieurs lames ovales lancéolées. Les lames possèdent une nervure centrale assez épaisse et ramifiée par de nombreuses petites veines latérales. A la fin du printemps les lames âgées sont ondulées (**Figure 9**).

Les oligosaccharides extraits de *Deleisseria sanguinea* possèdent des propriétés pharmaceutiques (hémostatiques, anti-inflammatoires…) (Grunewald (Grünewald and Alban, 2009) (Grünewald et al., 2009). De plus, c'est sur cette espèce d'algue qu'a été isolée la flavobactérie modèle : *Zobellia galactanivorans* dont sont issues toutes les enzymes de cette étude (cf. partie I.3.A.a).

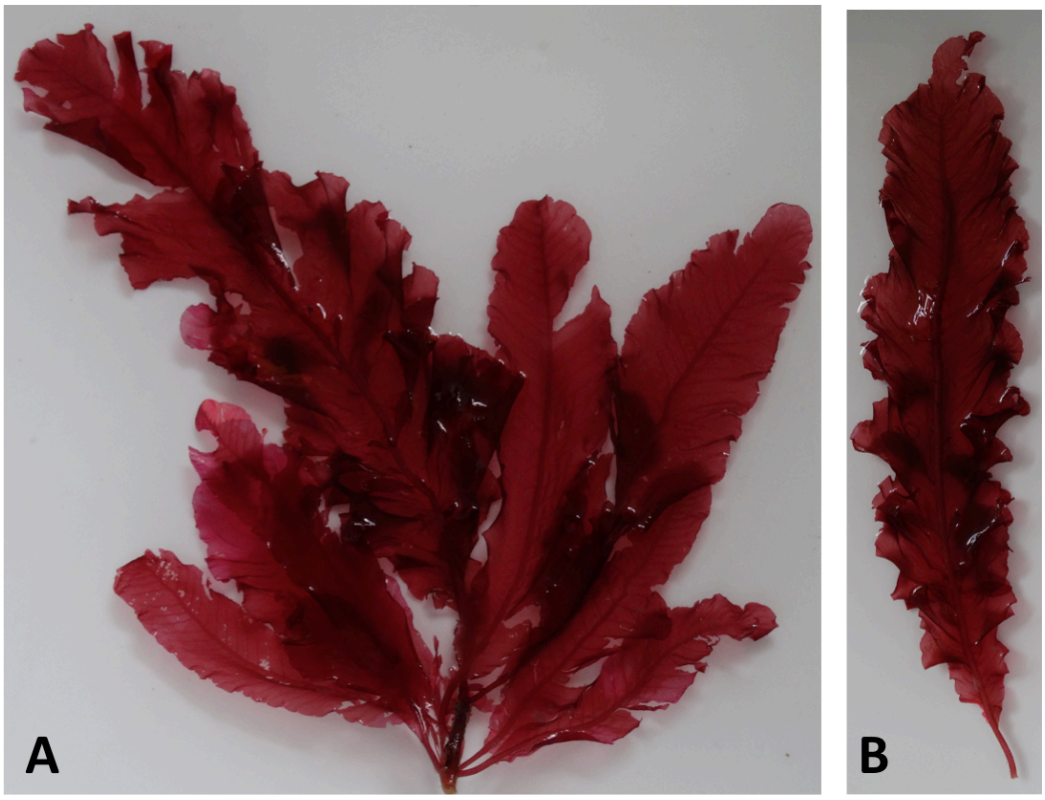


Figure 9 : *Delesseria sanguinea*
Les jeunes lames ont des bords réguliers (A) qui s'onduleront avec l'âge (B). En été, les parties membraneuses (couche monocellules) disparaissent peu à peu. Il ne restera plus que le stipe et les nervures centrales des lames. Ces algues peuvent être confondues avec *Phycodris rubens* (épiphytes strictes des *Laminaria*) dont les lames ont la forme de feuilles de chênes. (Photos J. Dorival)

I.1.D. f *Palmaria palmata* (L) Kuntz

Appelée communément goémon à vache, *Palmaria palmata* est reconnaissable à son thalle souvent ramifié, lobé et de couleur rouge (rosé à pourpre). Les lobes sont plats sans nervure centrale et peuvent atteindre 50 cm de long et 3 à 5 cm de larges. Le crampon est de forme discoïde, le stipe très court et d'un rouge très soutenu (**Figure 10**). Le crampon lui permet de se fixer aux rochers mais elle est également retrouvée en tant qu'épiphyte sur d'autres algues (e.g. *Fucus serratus*, *Laminaria hyperborea*). Elle est consommée dans différents pays (dont la France) et est référencée sous l'appellation *dulse*. La composition de sa paroi a été le sujet d'une étude assez poussée (cf I.2.A.a.3) (Bowker and Turvey, 1968a, b).

Figure 10: *Palmaria palmata* (photo provenant de https://www.pinterest.fr/pin/364580532316437894/)

I. 2 Les agars : de l'algue à la caractérisation

I. 2. A. Biogénèse et organisation des parois

I. 2. A. a Présentation de la paroi algale

La paroi des algues est le compartiment le plus externe et présente des similitudes avec celles des plantes. Les cellules sont ainsi entourées d'une paroi semi-rigide constituée de plusieurs polysaccharides bien distincts. Néanmoins, les parois des algues et des plantes sont différentes, en particulier du fait de la nature et de la répartition de leurs composants. Ainsi, comme pour les plantes, des fibres insolubles (comme la cellulose) sont présentes, mais surtout les parois des algues contiennent de nombreux polysaccharides solubles différents de ceux des plantes, notamment au niveau de la sulfatation.

I. 2. A. a. 1 Les polysaccharides anioniques

Les polysaccharides anioniques constituent la fraction majoritaire de la paroi des algues contrairement à ce qui peut être retrouvé chez les plantes (Kloareg and Quatrano, 1988). En effet, chez les plantes la majorité de la biomasse est issue de la cellulose (35 à 50%) et peut représenter jusqu'à 90% de la paroi, alors que cette dernière représente en général moins de 10 % de la paroi des algues et peut même être complétement absente (Usov, 2011).

Chez les Rhodophytes, les polysaccharides majoritaires sont des galactanes sulfatés. Ces derniers sont constitués par des D-galactoses et des D- (ou L-) galactoses liés alternativement par des liaisons α-1,3 et β-1,4. En fonction de la configuration (D ou L) du galactose lié en 4, deux familles sont définies (**Figure 11**):

- les carraghénanes, lorsque le galactose lié en 4 est de série D.
- les agars, lorsque le galactose lié en 4 est de série L.

A cause des groupements sulfate, les polysaccharides algaux sont fortement anioniques comparativement à ceux des plantes. L'origine de cette différence serait une adaptation des algues à l'eau de mer, car elle a une teneur en sulfate plus élevée (28 mM) que l'eau douce ou les eaux interstitielles plus acides (10 µM à 1,4 mM) (Schmidt-

Nielsen, 1997). Ainsi, les algues ont conservé cette capacité à sulfater leurs polysaccharides alors que les plantes l'ont perdu lors de la colonisation des milieux terrestres (Michel et al., 2010) (Collén et al., 2013).

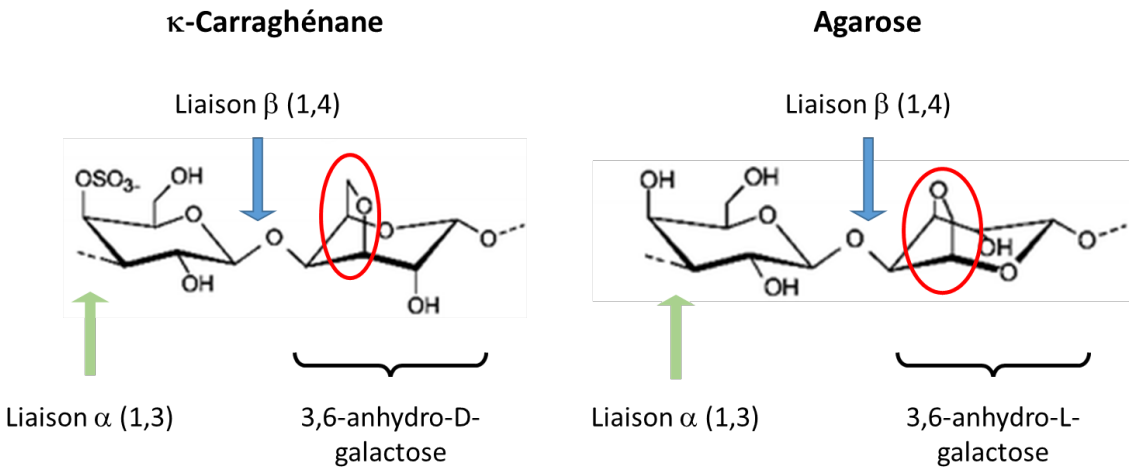


Figure 11 : Représentation des unités dissaccharidiques répétitives pour l'agarose et le κ-carraghénane

La différence entre les agars et les carraghénanes réside dans la configuration du galactose lié en 4 : D pour le carraghénane et L pour l'agar. Les carraghénanes comme les agars contiennent des modifications comme le pont 3,6-anhydro (ellipse rouge) ou bien des sulfatations. Concernant les carraghénanes, une nomenclature est en place et repose sur le profil de modification. Le disaccharide représenté ici présente un pont 3,6-anhydro et un sulfate en C4 sur le D-galactose lié en 3 ce qui fait que nous avons un motif κ-carrabiose. Le motif présenté ici pour l'agarose est nommé agarobiose.

Dans la littérature, il est également fait mention d'une troisième famille de polysaccharides : les carragars ou « hybride DL » (Usov, 2011). Dans cette famille, la chaine polysaccharidique présenterait une alternance dans la configuration des galactoses liés en β-1,4. Des réserves ont tout de même été émises sur leur existence. En effet, les auteurs n'apportent pas de preuve indiscutable qu'il s'agisse bien d'un polymère mixte et non d'un mélange d'agar et de carraghénane (Usov, 2011) (Lahaye, 2001).

Il est à noter que le squelette des agars et des carraghénanes est souvent compliqué par de nombreuses modifications et substitutions. L'une de ces modifications est la formation d'un pont 3,6-anhydro sur le galactose lié en 4. Les autres modifications sont : méthylations, sulfatations, branchements de xyloses ou de glucoses voir plus rarement des pyruvilations (Lahaye, 2001). Cependant, bien que les modifications des galactanes varient en fonction de la saison et du cycle de vie de l'algue, la nature du polysaccharide (agar ou carraghénane) est, quant à elle, invariable, *a contrario* des fibres insolubles (par

exemple la cellulose qui peut être remplacée, durant le cycle de vie, par du mannanes dans la paroi de certaines algues)(Usov, 2011).

I. 2. A. a. 2 Les polysaccharides neutres

Cette fraction polysaccharidique est composée de polysaccharides linéaires, neutres et généralement insolubles (cellulose, mannanes et xylanes). La cellulose reste très minoritaire au sein des algues ou elle représente moins de 10 % de la paroi (Usov, 2011) (Kloareg and Quatrano, 1988) (Popper et al., 2011).

Les compositions polysaccharidiques peuvent varier en fonction du stade de vie, comme par exemple dans le cas de *Porphyra tenera,* où lors d'une des phases diploïdes (*conchocelis*) la paroi contient de la cellulose et peu de mannanes, alors qu'au cours de la phase de développement du thalle (haploïde), la paroi ne contient plus que du mannane (Mukai et al., 1981).

En plus d'être très peu présente chez les algues, la cellulose issue de ces dernières présente également une structure différente de celle retrouvée chez les plantes. Cela a été révélée par des analyses de RMN des solides et des études de diffraction des rayons X : la cellulose chez les algues rouges, surtout des Bangiophycées, présente une organisation différente qui la rend moins cristalline (Stiger-Pouvreau et al., 2016). En effet, les microfibrilles sont plus fines (1 à 1,5 nm d'épaisseur) et ont une structure de type Iα (composée d'une seul chaine) au lieu du type Iβ majoritaire chez les plantes (composée de deux chaines parallèles) (Atalla and VanderHart, 1999) (Nishiyama et al., 2003)(**Figure 12**).

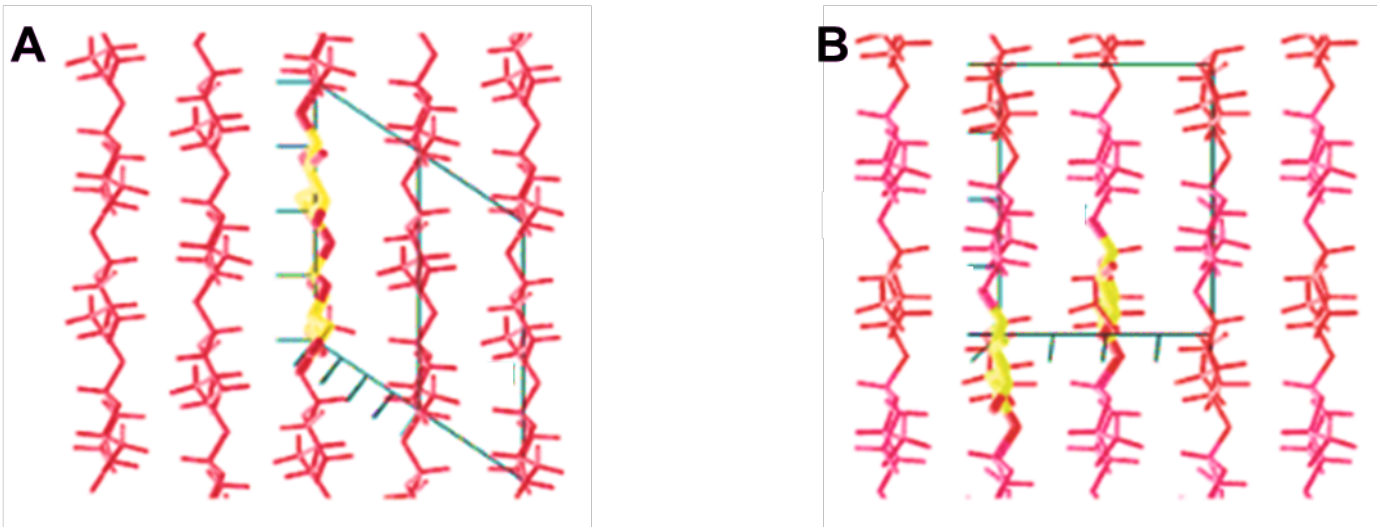


Figure 12 : Schéma de l'organisation des différents types de microfibrilles de cellulose (d'après Nishiyama et al., 2003)

Le type de fibrille de cellulose représenté en A est le type Iα et en B il s'agit du type Iβ. Cette représentation est perpendiculaire au brin et aux liaisons hydrogènes. Les brins de cellulose sont en rouges. Les fins traits noires figurent l'unité asymétriques (i.e. la plus petite région de l'espace qui quand elle est répétée, selon des opérations définies, permet de reconstruire la structure cristalline) avec le motif répété en jaune. Il apparait clairement sur ce schéma que l'unité asymétrique, et donc l'organisation, des microfibrilles et différentes selon le type.

Cette différence proviendrait du système de biogénèse propre aux algues, bien que les études sur le système de biogénèse de la cellulose chez les Bangiales révèlent des gènes similaires à ceux des plantes. Les clusters d'enzymes de synthèse de la cellulose ou TC « Terminal Complexes » se retrouvent au sein de la membrane plasmique comme chez les plantes. Une des hypothèses expliquant les différences entre les celluloses produites par les algues et les plantes seraient l'organisation de ces TC (Stiger-Pouvreau et al., 2016) : les microfibrilles ainsi produites sont différentes en tailles et en morphologies de celles produites chez les plantes.

I. 2. A. a. 3 Modèles de paroi chez les algues rouges.

Aujourd'hui il existe deux modèles d'organisation de la paroi de Rhodophytes. Le premier modèle est celui de la paroi d'une algue carraghénophyte : *Eucheuma cottonii* (Ordre des Gigartinales)(**Figure 13**) (Lechat, 1998). Sa paroi est majoritairement composée de carraghénanes (polysaccharides solubles), mais d'autres polysaccharides sont retrouvés comme du xylogalactane ou de la cellulose.

Ce modèle est très différent de celui trouvé chez *Palmaria palmata* (Ordre des Palmariales) qui est le second modèle à notre disposition. Contrairement à *Eucheuma cottonii,* il présente très peu de cellulose et de galactanes sulfatés (Deniaud Boüet, 2002) (Lahaye et al., 2003) (Deniaud et al., 2003). Ainsi dans sa paroi, les galactanes sulfatatés (comme le carraghénane) sont remplacés par des xylanes sulfatés.

Bien que nous ne disposions pas d'autres modèles de paroi, il suffit de s'intéresser à la composition en polysaccharides solubles des autres algues rouges pour être convaincu de la diversité et de la complexité de leur paroi. A titre d'exemple, nous citerons les algues de l'ordre des Sebdeniales. Les parois de ces dernières sont composées majoritairement de mannanes (Ghosh et al., 2009) et non pas de carraghénanes comme pour *E. cottonii* ou de xylanes comme pour *P. palamata*.

La conclusion essentielle de cette partie est que la composition de la paroi chez les Rhodophytes est variable selon les groupes considérés, et par conséquent, aucun modèle ne serait valable pour toutes les algues rouges. Au mieux, un modèle ne sera valide que pour un groupe donné.

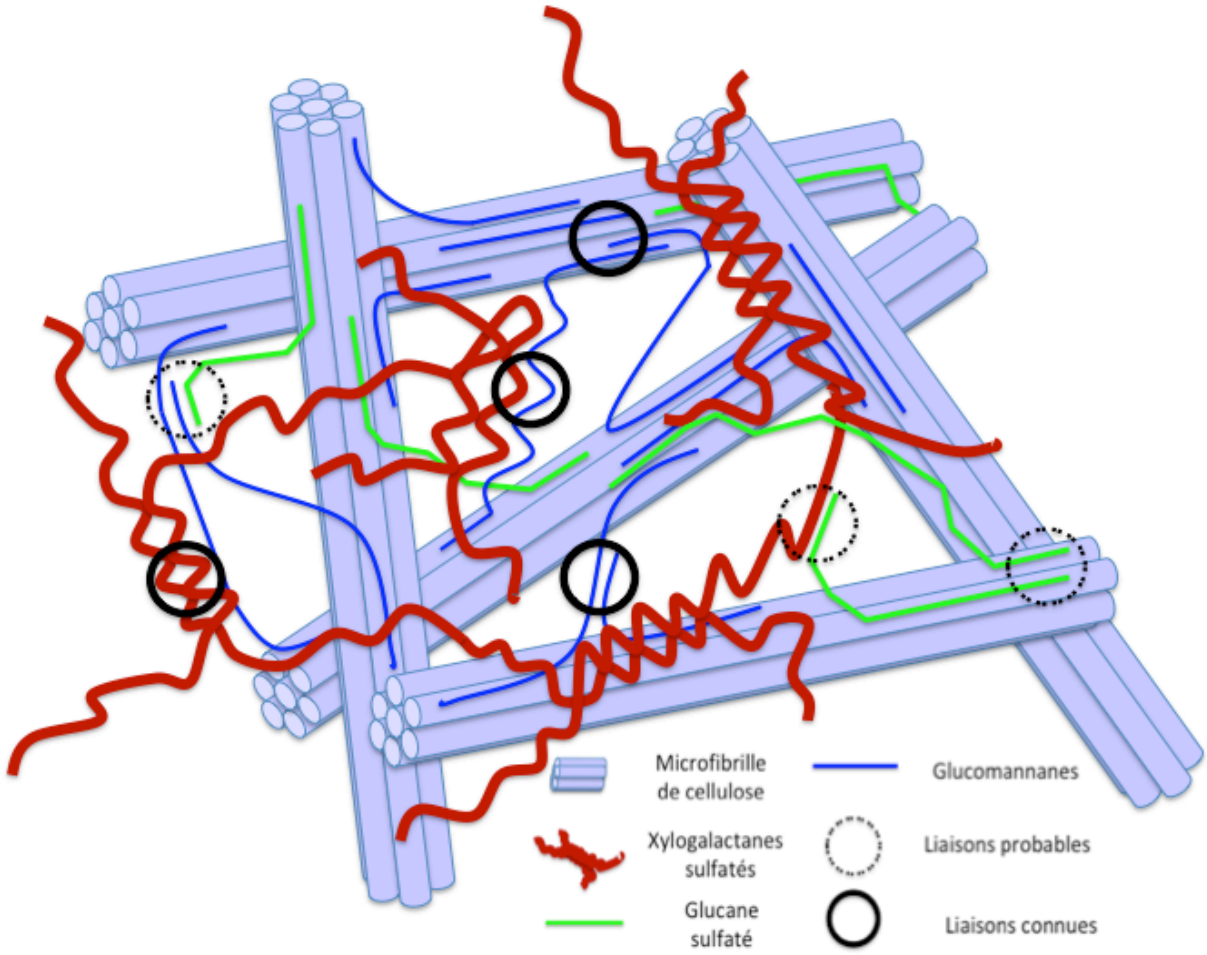


Figure 13 : Représentation schématique de la paroi de l'algue carraghénophyte *Eucheuma cottonii* (d'après Lechat, 1998)

Dans ce modèle, la cellulose est un composé minoritaire de la paroi (8%). Cette paroi, contient en effet majoritairement des polymères non-cellulosiques :80% de carraghénanes. La fraction de carraghénane solubilisée à l'eau chaude n'est pas représentée ici. Une seconde fraction est résistante à l'extraction aqueuse contenant des xylogalactanes sulfatés. Le β-glucane (Glucane sulfaté en vert) présent chez cette algue comporte des liaisons (1,3).

I. 2. A. b Fonctions de la paroi

La paroi des algues se différencie de celles des plantes par le ratio polysaccharides insolubles/solubles, ainsi que par la nature très acide de ces derniers (due à la présence de sulfatation). Ces différences s'expliquent par le milieu de vie extrême des algues : une forte pression osmotique, des courants pouvant être importants et l'alternance des périodes d'immersion et d'émersion (Kloareg and Quatrano, 1988). Si nous corrélons ces données avec les propriétés des polysaccharides pariétaux exposés dans les paragraphes précédents, nous percevons l'importance de leur originalité. En effet, ces composés confèrent aux thalles une flexibilité permettant de résister aux forces de

cisaillement dues aux courants. Le caractère acide de la matrice est un élément important jouant sur la régulation des échanges cationiques, qui permettent de diminuer la pression osmotique exercée par l'eau de mer sur la cellule. De plus, les galactanes sulfatés sont fortement hydratés permettant ainsi aux algues de résister à la dessiccation lorsqu'elles sont émergées (Kloareg and Quatrano, 1988).

La paroi contient également des protéines, qui, bien que quantitativement minoritaires, assurent des rôles primordiaux pour la cellule tels que le transport d'ions, la détoxification, l'adhésion cellulaire... Elles peuvent également avoir un rôle de support aux polysaccharides comme cela a été proposé par E. Deniaud dans sa thèse. D'autre part, certaines de ces protéines sont aussi probablement impliquées dans la modification de la paroi (Craigie, 1990).

I. 2. A. c Biosynthèse des polysaccharides matriciels chez les algues rouges : cas particuliers des agars

Actuellement, la biosynthèse des agars est considérée comme étant une voie terminale du cycle du carbone chez les agarophytes (Lee et al., 2017b) (Macler, 1986). *A contrario* des plantes, la biosynthèse des composants matriciels algaux est peu connue. Les voies de biosynthèse exposées ici représentent donc des hypothèses qui reposent largement sur les connaissances acquises sur les voies de synthèses des plantes. La synthèse des agars se déroulerait en quatre grandes étapes, premièrement les précurseurs activés seraient produits dans le cytosol. Deuxièmement, des galactosyl-transférases catalyseraient l'élongation de la chaine polysaccharidique, chaines qui seraient ensuite sulfatées grâce à l'action de sulfo-transférases. Finalement des galactose-6-sulfurylases convertiraient les galactoses sulfatés en 6 en résidus 3,6-anhydro-galactoses (Rees, 1961) (Hemmingson et al., 1996b) (Genicot-Joncour et al., 2009) **(Figure 14**). Concernant, le compartiment cellulaire et le moment où les chaines polysaccharidiques serait modifiées par les enzymes autres que les L-galatoses-6-sulfotransférases (nécessaire pour permettre la formation du pont 3,6 anhydro-galactose) nous n'avons aucune preuve. Mettre toutes les enzymes de modification au même point de la biosynthèse est un choix de notre part et ne reflète en aucun cas ce qui se passe réellement chez les Rhodophytes.

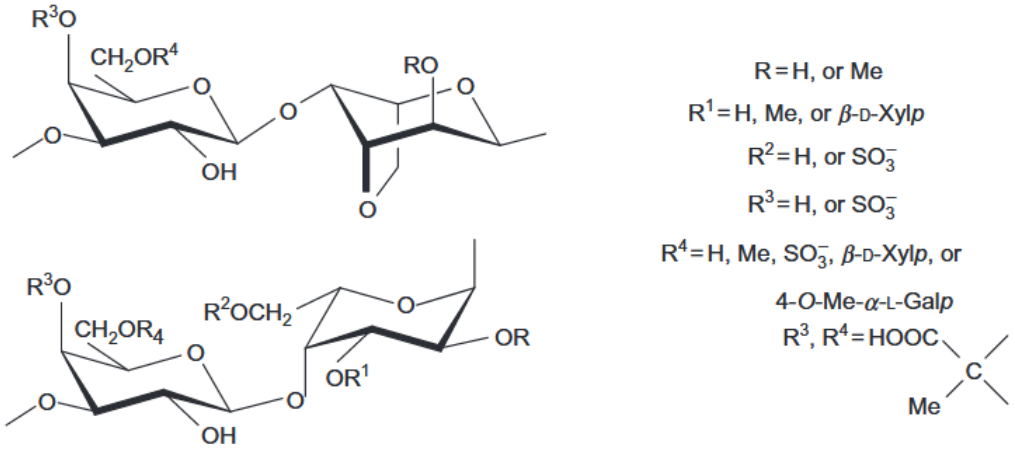


Figure 14 : Exemple de modification retrouvée chez les agars (d'après d'Usov, 2011).

En haut : un exemple d'agars modifiés avec le pont 3,6 anhydro. En bas : un exemple d'agars modifiés en absence du pont 3,6 anhydro. En plus des méthylations et des sulfatations, les agars présentent souvent des branchements xylose et des pyruvylations.

I. 2. A. c. 1 Biosynthèse des précurseurs

Afin de synthétiser un polymère de galactose, la cellule doit déjà disposer d'une forme activée du galactose. Dans le cas des agars, les deux énantiomères activés du galactose sont l'uridine diphosphate (UDP)-D-galactose et le guanosine diphosphate (GDP)-L-galactose.

Ces deux composés proviendraient de deux voies de biosynthèse différentes : celle du glucose et celle du mannose (Major et al., 2005b) (Frey, 1996) (Lee et al., 2017) avec comme point de départ commun, le fructose-6-phosphate (F6P) issu du cycle de Calvin. En fonction de l'isomérase prenant en charge le F6P, la Glucose-6-phosphate isomérase (GPI) ou la Mannose-6-phosphate isomérase (MPI), la cellule produit soit de l'UDP-D-galactose soit du GDP-L-galactose (Lee et al., 2017)(Frey, 1996). Une autre voie est également possible, voire complémentaire, avec pour origine l'amidon floridien. Ce polysaccharide est le principal produit de la photosynthèse chez les Rhodophytes. Il est dégradé soit par une glucane-1,4-α-glucosidase (GAG, EC 3.2.1.3) couplée à une glucokinase (GK, EC 2.7.1.2), soit par une α-1,4-glucane/glycogène phosphorylase (GP, EC 2.4.1.1) pour produire respectivement du glucose-6-phophate ou du glucose-1-phosphate (Lee et al., 2017) (Hemmingson et al., 1996a).

Biosynthèse de l'UDP-D-galactose

L'UDP-D-galactose serait synthétisé *via* la voie de Leloir (Leloir, 1951) (Frey, 1996). Il peut soit être produit par la galactose-1-phosphate urydylyl-transférase (GALT, EC 2.7.7.12) qui va transférer l'UDP de l'UDP-D-glucose au galactose-1-phosphate pour générer de l'UDP-D-galactose. Le galactose-1-phosphate (G1P) quant à lui proviendrait de la dégradation du floridoside (hétéroside spécifique des algues rouges, un des principaux produits de la photosynthèse et précurseur de nombreuses macromolécules) (Li et al., 2001) (Li et al., 2002). A partir de ce dernier, une α-galactosidase (GLA) produit du G1P.

L'autre voie possible pour produire l'UDP-D-galactose serait l'intervention d'une épimérase : l'UDP-glucose-4-épimérase (UGE, EC 5.1.3.2). Cette dernière convertirait l'UDP-D-glucose en UDP-D-galactose (Frey, 1996).

Biosynthèse du GDP-L-galactose

La voie de biosynthèse du GDP-L-galactose telle que décrite aujourd'hui, ressemble au début de la voie de production de l'acide ascorbique chez les plantes terrestres : la voie de Wheeler-Smirnoff (Wheeler et al., 1998) (Smirnoff and Wheeler, 2000).

Le GDP-L-galactose proviendrait de la voie du D-mannose/L-galactose (Major et al., 2005a). Tout d'abord le mannose doit être sous forme GDP-D-mannose, qui est synthétisé par la GDP-mannose phosphorylase (GMP EC 2.7.7.22). Cette réaction réversible consiste en la conversion de mannose-1-phosphate et de GTP en GDP-D-mannose et pyrophosphate. La production de GDP-L-galactose serait réalisée par la GDP-mannose-3', 5'-épimérase (GME, EC 5.1.3.18). Cette réaction d'épimérisation est réalisée sur les carbones 3 et 5 du GDP-D-mannose pour obtenir du GDP-L-galactose (Wheeler et al, 1998).

Enzymes isolée et caractérisé chez les algues rouges

Au sein de ces voies de biosynthèse putatives certaines enzymes ont pu être isolée à partir d'algues rouges et caractérisées. Leurs présences tend à confirmer la validité des mécanismes de biosynthèse présentés ci-dessus.

Concernant la synthèse de l'UDP-D-galactose, trois enzymes ont été étudiées : la GALT, la GLA et l'UGE. L'enzyme GALT a été clonée à partir de l'algue rouge agarophyte *Gracilaria changii* et caractérisée biochimiquement (Siow et al., 2012). L'alpha-galactosidase (GLA) a été extraite et caractérisée à partir de *Gracilaria sordida* (Yu and Pedersén, 1990). Concernant l'UGE, elle a été clonée et caractérisée biochimiquement, mais la nature (agarophytes ou carraghénophytes) de l'algue dont elle provient (*Galdiera sulphuraria*) n'est pas connue (Prosselkov et al., 1996).

Une seule enzyme de la voie de biosynthèse du GDP-L-galactose a été caractérisée : la GME. La GME a été clonée à partir de l'algue rouge agarophyte *Gracilaria changii* et son étude biochimique a confirmé qu'elle produit du GDP-L-galactose, et de plus, le suivi de son expression *in vivo* a démontré que son expression est corrélée à l'augmentation de production d'agar (Siow et al., 2013).

Toutes les enzymes présentées ici, ainsi que les voies métaboliques dans lesquelles elles interviennent, sont le produit de recherches effectuées sur de nombreuses Rhodophytes, y compris des non carraghénophytes et agarophytes. C'est pourquoi pour avoir une vision de la synthèse des agars, nous avons encadré sur la **Figure 15** tous les gènes présents dans le génome de l'algue agarophytes *Porphyra umbilicalis* (Brawley et al., 2017), ainsi que leur identité de séquence avec les protéines d'algues rouges caractérisées. Concernant l'α-galactosidase, nous ne disposons pas de sa séquence protéique ce qui rend impossible la recherche de cette dernière dans le génome, même si plusieurs gènes ont été annotés comme tels.

Il est intéressant de constater que nous retrouvons les enzymes de la voie de Leloir chez les algues (Leloir, 1951), mais pas chez les plantes (Lee et al., 2017). Les plantes ont probablement perdu ces gènes au cours de l'évolution, rendant l'utilisation du galactose difficile pour ces dernières. Une autre remarque, est que le L-galactose est naturellement très peu abondant chez les plantes (Baydoun and Fry, 1988) alors qu'il est extrêmement abondant dans les galactanes d'origine marine, au point que l'on retrouve la voie de Wheeler-Smirnoff modifiée pour répondre à ce besoin chez les algues. Une preuve de plus que les algues ont eu un chemin évolutif bien différent de celui des plantes.

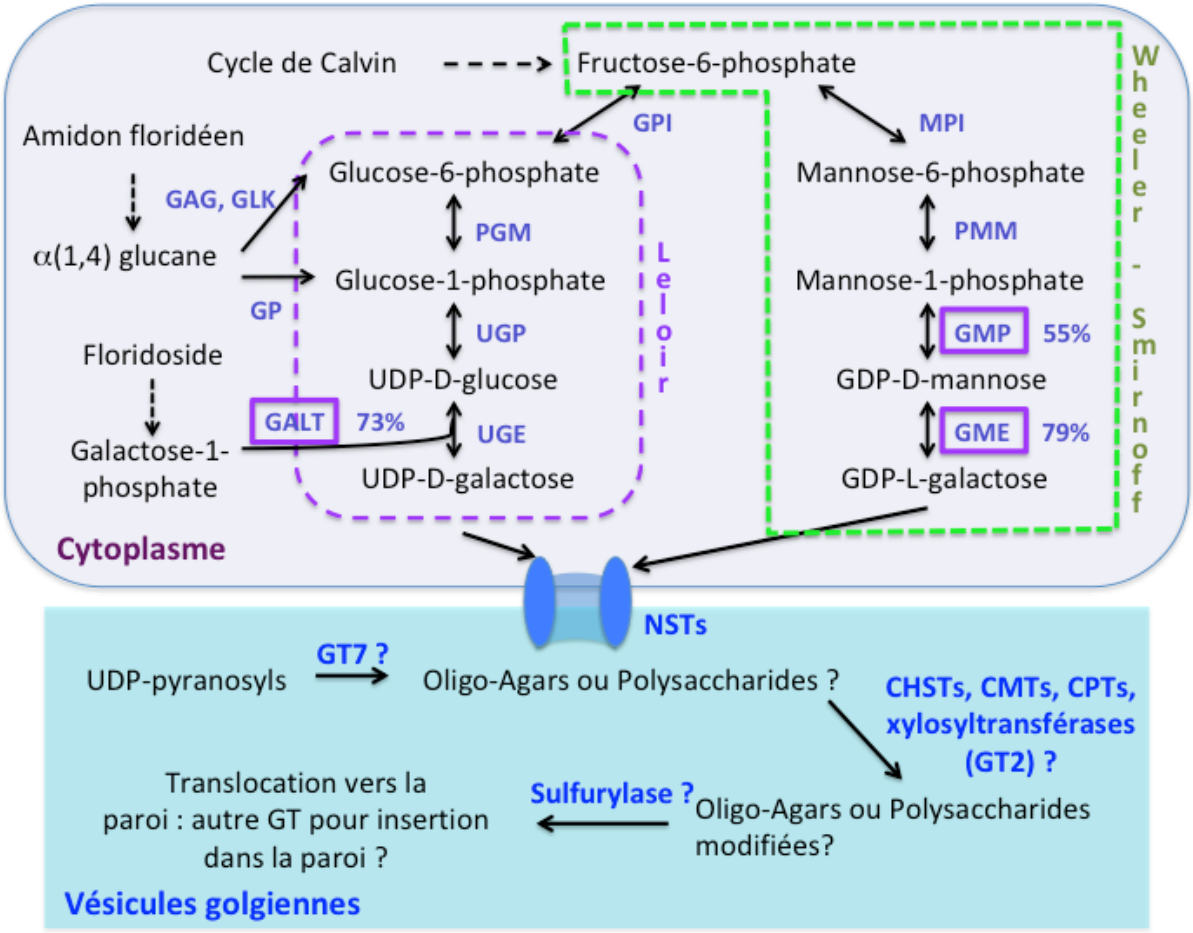


Figure 15 : Voie putative de biosynthèse des agars (d'après Lee et al, 2017)

La voie de biosynthèse commencerait avec l'UDP-D-galactose et le GDP-L-galactose, eux-mêmes synthétisés à partir de fructose-6-phosphate issu du cycle de Calvin, du floridoside ou de l'amidon floridéen. Les précurseurs activés seraient ensuite transloqués vers l'appareil de Golgi pour synthétiser des chaines polysaccharidiques. Que ce soit pour des oligosaccharides ou des polysaccharides, une étape nécessitant le concours de GTs au niveau de la paroi semble nécessaire pour assurer la continuité de cette dernière. Ces chaines seraient ensuite modifiées avant d'être envoyées au niveau de la paroi.

GT : glucosyl-transférase ; GALT : galactose-1-phosphate urydylyl-transférase ; SULT : sulfo-transférases ; CHST : carbohydrate sulfo-transférase ; CMT : carbohydrate méthyl-transférases ; GME : GDP-mannose 3',5'-épimérase ; GMP : GDP-mannose phosphorylase ; CPT : carbohydrate pyruvyl -transférase ; GLA : α-galactosidase ; UGE : UDP-glucose-4-épimérase ; F6P : fructose-6-phosphate ; GPI : glucose-6-phosphate isomérase ; MPI : mannose-6-phosphate isomérase ; GAG : glucane-1,4-α-glucosidase ; GK : glucokinase ; GP : α-1,4-glucane/glycogène phosphorylase ; PMM : phosphomannose mutase ; PGM : phosphoglucomutase ; NSTs : *nucleotides sugar transporters*.

I. 2. A. c. 2 Formation des chaines d'agar et leurs modifications

Les nucléo-monosaccharides (GDP-L-galactose et UDP-L-galactose) sont produits dans le cytosol, puis transloqués dans la lumière de l'appareil de Golgi grâce à un transporteur spécifique de nucléotides sucres (*nucleotides sugar transporters* en anglais, NSTs). Ces derniers sont inclus dans la membrane golgienne et agissent comme des

antiports : ils vont importer dans la lumière golgienne les nucléo-monosaccharides et exporter les nucléotides monophosphates (Gibeaut, 2000). Bien que n'ayant encore jamais été étudié chez les algues, les travaux menés sur le riz (*Oryza sativa*) montrent que les NSTs modulent la biosynthèse de la paroi. Ainsi, il n'est pas exclu qu'ils aient un rôle similaire chez les algues (Zhang et al., 2011).

A proximité des NSTs, des galactosyl-transférases (et des glycosyl-transférases (GTs) en général) (Colley, 1997) (Hennet, 2002) catalysent le transfert d'un monosaccharide activé (UDP-D-Gal ou GDP-L-galactose) sur un accepteur saccharidique et l'incrémentent ainsi d'une unité glycosidique. Dans le cas de l'agar (i.e. une chaine composée d'alternance de D-galactose et de L-galactose liés par des liaisons glycosidiques α-1,3 et β-1,4) au minimum deux types de GT sont donc nécessaires : les α-1,3 et les β-1,4 galactosyl-transférases. A l'heure actuelle, aucune de ces enzymes n'a été ni clonée, ni étudiée chez les algues, mais en croisant les données sur les glycosyl-transférases déjà caractérisées et les génomes d'algues rouges à notre disposition (Collen et al., 2013 ; Brawley et al., 2017) de nombreux gènes candidats peuvent être identifiés.

Concernant les algues *Porphyra umbilicalis* (agarophytes) (Brawley et al., 2017) et *Chondrus crispus* (carraghénophyte) (Collen et al., 2013), toutes deux possèdent un gène homologue à la famille GT2 (Henrissat, 1991). Ce gène est très similaire à celui que l'on retrouve chez des algues vertes et est très proche des gènes de *celluloses synthase like* (CSL) des plantes. Il est surtout proche de celui des mannanes et xyloglucane synthases (Brawley et al., 2017). Ce qui fait de cette GT2 une bonne candidate pour la synthèse du mannane.

Concernant la synthèse des galactanes, un gène étant annoté comme GT7 a été proposée chez *C. crispus* et *P. umbilicalis* (Collen et al., 2013)(Brawley et al., 2017), de plus plusieurs GTs ont été annotées mais n'ont pu être classifiées. Il est possible que ces GTs de fonction inconnue soient impliquées dans la modification de l'agar de *P. umbilicalis*. Parmi ces dernières, un gène semble très proche de ceux des GT77 (UDP-xylosyl-transférases), ce qui en fait un excellent candidat pour l'étude de la biosynthèse des galactanes xylosylés (Brawley et al., 2017) (Petersen et al., 2009).

L'analyse et la comparaison des génomes des microalgues rouges *Galdieria sulphuraria* et *Cyanidioschyzon merolae* (dont les polysaccharides pariétaux n'ont pas été caractérisés) ont révélé d'autres gènes candidats (Barbier et al., 2005). Toutes les deux possèdent un gène qui coderait putativement une α-1,3 glucosyl-transférase qui

prendrait en charge le GDP-Glucose, et *G. sulphuraria* aurait quant à elle un gène codant pour une galactosyl-transférase avec comme substrat le GDP-galactose (Barbier et al, 2005). Ainsi, ces génomes nous permettront peut-être à l'avenir d'identifier d'autres gènes responsables de la synthèse des polysaccharides algaux.

Une fois la chaine polysaccharidique allongée, les galactoses récemment ajoutés deviennent les substrats d'enzymes de modification. Cette fonctionnalisation des hydroxyles va permettre au polysaccharide d'acquérir ses propriétés finales. Ces ajouts sont la source de la grande diversité des agars.

I. 2. A. c. 3 Modification de la chaine en cours d'élongation

Sulfatation

L'hétérogénéité des agars, et plus largement des galactanes, repose en grande partie sur les esters sulfate qui sont ajoutés aux polysaccharides par des sulfo-transférases (SULTs), que nous appellerons ici des carbohydrate sulfo-transférase (CHSTs). Les positions pouvant être modifiées (car non impliquées dans une liaison glycosidique) au sein des agars sont les carbones C_2, C_4 et C_6 du D-galactose et les carbones C_2, C_3 et C_6 du L-galactose. Ces modifications sont spécifiques de chaque organisme et impliquent des gènes différents pour chacune d'entre elles. Ces enzymes, dans le cas de la biosynthèse de la matrice, sont proposées comme étant transmembranaires et situées au sein de l'appareil de Golgi, comme c'est le cas chez l'homme où de nombreuses CHSTs ont été étudiées (Tveter-Gallagher et al., 1984) (de Graffenried et al., 2004). Les SULTs ont en commun des régions très conservées, ce qui permettent de les identifier (Varin et al, 1992). C'est ainsi que le gène candidat (scaffold_2401 :1713-604) pour une CHST a été identifié dans le génome de *P. umbilicalis* et proposé comme codant une L-galactose-6-sulfo-transférase qui permettrait la sulfatation majoritaire du porphyrane (le sulfate en position 6 du L-galactose) (Brawley et al., 2017).

Bien que pour l'instant aucune SULT d'origine algale n'ait été caractérisée, nous disposons d'études enzymologiques sur des sulfo-transférases bactériennes et humaines (Chapman et al., 2004). Toutes ces enzymes nécessitent un donneur de sulfate, chez les plantes, le PAPS (3'-phosphoadenosine 5'-phosphosulphate) à partir duquel l'enzyme va pouvoir catalyser le transfert du groupe sulfuryle (SO_3) sur un hydroxyle de la chaine

polysaccharidique. Les motifs conservés chez les SULTs correspondent à la fixation du 5' et du 3' du PAPS. (**Figure 16**) Le mécanisme catalytique des sulfo-transférases commence à être bien connu. Par contre, le moment où elles agissent nous est toujours inconnu : agissent-elles sur le polymère neutre, ou bien sur le galactose juste après l'élongation ? La réponse semble toute autre. Une hypothèse est que la sulfatation aurait lieu avant la polymérisation (Cardozo et al., 2007).

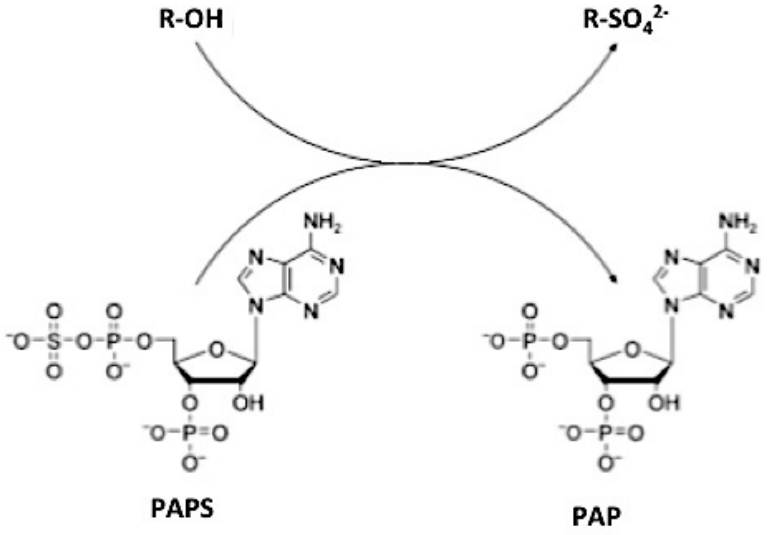


Figure 16 : Schéma générale de la réaction catalysée par les sulfo-transférases.
Le PAPS est une molécule de type adénosine tri-phosphate dans lequel le phosphate en position β est remplacé par un sulfate. Cette molécule est la forme activée du sulfate qui est utilisée par les sulfo-transférases pour sulfater divers composés.

Méthyl-transférase et pyruvate -transférase

Concernant ces deux types d'enzymes de modifications, nous avons relativement peu de données. Comme pour les sulfo-transférases, nous ne savons pas à quel moment de la synthèse elles agissent.

Les méthyl-transférases (MTs) ont comme les CHSTs besoin d'un donneur activé. La grande majorité de la littérature disponible sur les MTs à ce jour, repose sur les MTs agissant sur les protéines et les acides nucléiques (Lyko, 2017). Les données disponibles sur les carbohydrates méthyl-transférases (CMTs) ont été obtenues à partir de sources bactériennes. Par exemple, nous pouvons citer deux enzymes provenant de *Mycobacterium smegmatis,* qui permettent de méthyler le rhamnose en position 3 et 4

(Patterson et al., 2000) ou encore la CMT provenant de *Rhizobium meliloti* qui transfère un groupement méthyl en C_2 d'un résidu galacturonosyle (Ruiz and Ugalde, 1998).

Le donneur de méthyle est principalement la S-adénosylméthionine (SAM). En effet, chez les plantes, le transport de SAM dans les vésicules golgienne est corrélé à la synthèse de pectines méthylées (Reyes and Orellana, 2008). Comme pour les sulfo-transférases, il se pose également la question de savoir si les CMTs agissent après l'élongation de la chaine polysaccharidique ou bien si elles peuvent directement utiliser un ose méthylé. Lee et associé ont mené une étude sur deux algues agarophytes (*Gracilaria changii* et *G. salicorna*) en condition de privation de sulfate. Ils ont ainsi mis en évidence trois méthyl-transférases dont la transcription varie en réponse à ce stress et qui seraient impliquées dans la synthèse de la paroi (Lee et al., 2017a).

Concernant les pyruvyl-transférases (PTs), nous disposons d'encore moins d'informations que pour les CMTs. Toutes proviennent des bactéries à Gram négatif où les PTs sont essentielles à la synthèse du peptidoglycane (Brown et al., 1995) (Pleschberger et al., 2013). Bien que rares, des polysaccharides pyruvylés sont aussi retrouvés chez les eucaryotes : algues rouges et vertes, limaces de mer (*Aplysia kurodai*)(dans ce cas seul du galactose pyruvylé a été détecté) (Araki et al., 1989) et des éponges de mer (*Microciona prolifera*) (Spillmann et al., 1993). Le seul donneur de pyruvate connu des pyruvyl-transférases est le phosphoénolpyruvate (Wiederschain, 2009). Quelques PTs ont été annotées chez l'algue rouge *Gracilaria changii*. L'étude du transcriptome de cette dernière a d'ailleurs permis d'identifier des pyruvyl-transférases et des méthyl-transférases pouvant être impliquées dans la synthèse d'agar modifié (Lee, et al 2017a). Néanmoins nous n'avons pas pu effectuer les recherches chez *P.umbilicalis* car les séquences ne sont pas publiques.

I. 2. A. c. 4 La formation du pont 3,6 anhydro : dernière étape de la synthèse des agars ?

Les étapes de modification, ainsi que la biosynthèse des galactanes, seraient finalisées par la formation du pont 3,6-anhydro sur le L-galactose dans le cas de l'agar (Hemmingson et al., 1996b). Elle est réalisée par des galactose-6-sulfurylases (EC 2.5.1.5). Ces dernières catalysent la désulfatation en C_6 du L-galactose, considéré comme le précurseur du 3,6-anhydro-galactose, et la formation du pont. La première

caractérisation de ces enzymes date de 1961. L'enzyme avait été extraite à partir de *Porphyra umbilicalis* et permet la formation d'un pont 3,6-anhydro sur le L-galactose (Rees, 1961). D'autres galactose-6-sulfurylases ont été isolées et caractérisées depuis, notamment à partir d'une algue carraghénophytes *Chondrus crispus*.

Cette dernière est une galactose-6-sulfurylase qui catalyse la formation d'un pont 3,6-anhydro sur le D-galactose à partir d'un 6-sulfate-D-galactose (Genicot-Joncour et al., 2009). Comme cette étape ne peut qu'intervenir après sulfatations, il a été proposé que ce soit l'étape terminale. Mais il est tout à fait possible que d'autres modifications soient ajoutées après celle-ci. Ou encore, que le pont 3,6-anhydro soit formé sur des oligosaccharides et dans ce cas la dernière étape serait l'association des oligosaccharides pour former des polysaccharides.

I. 2. A. c. 5 Synthèse des chaines polysaccharidiques et dépôt au niveau de la paroi

Les agars sont produits dans l'appareil de Golgi, mais sont aussi trouvés dans des sacs de mucilage associés au réticulum endoplasmique (Tsekos, 1981). Il a d'ailleurs été prouvé que chez *Porphyra yezoensi*, le trafic des vésicules issues de l'appareil de Golgi est essentiel pour la synthèse de la paroi (Li et al., 2008). Ce qui accrédite l'idée qu'une fois produit dans le golgi, les néo-polysaccharides sont transloqués vers la paroi *via* les vésicules. Toujours chez la même algue, la croissance de la cellule et le dépôt de la matrice sont subordonnés à un gradient spécifique de Ca^{2+} induit par la lumière (Li et al., 2009). Le Ca^{2+} est nécessaire pour la mise en place de l'axe apical des algues et permet d'orienter le dépôt de la nouvelle matrice, ce qui explique qu'en son absence il n'y ait plus de néo-synthèse.

Le calcium semble jouer un rôle de régulation à un autre niveau. Les GTs sont cation-dépendantes et donc régulées par la disponibilité du calcium intracellulaire. Ainsi les cellules pourraient à la fois réguler le dépôt des agars neo-synthétisé ainsi que leurs synthèses en modulant la concentration intracellulaire du calcium (Reyes and Orellana, 2008). Un autre rôle des cations serait de compenser les charges négatives des polysaccharides se trouvant dans les vésicules (Reyes and Orellana, 2008).

Finalement, que ce soit au niveau de l'organisation ou de la synthèse, nous avons encore beaucoup de choses à apprendre sur les galactanes issues des Rhodophytes. C'est

d’autant plus vrai que les études de caractérisation des agars ont démontré une forte diversité ce qui implique des systèmes de synthèse et de régulation complexe.

I. 2. B. Diversité des agars

Les agars sont des galactanes méthylés et sulfatés et représentent une part importante de la matrice extracellulaire chez les Rhodophytes qui les produisent. Les agars sont des hydrocolloïdes utilisés comme agent de texture, stabilisant…

Plusieurs industries les emploient : pharmaceutiques, agroalimentaires… Nous les retrouvons aussi couramment dans les laboratoires (par exemple en microbiologie). Bien que moins utilisés que les carraghénanes (9600 tonnes vendu contre 50 000 tonnes), les agars sont les polysaccharides marins les plus chers du marché : 18$ le kilo en moyenne contre 10,5$ pour les carraghénanes (Bixler and Porse, 2011), ce qui en fait un marché relativement lucratif.

Pour l'instant, l'usage de ce polysaccharide est lié à ses propriétés rhéologiques, c'est-à-dire à sa capacité à former des gels, mais de plus en plus d'applications à plus fortes valeurs ajoutées se développent, telle que son utilisation comme agent anti viral (Carlucci et al., 2012) (Cosenza et al., 2017). De plus, l'agar issu de *Porphyra haitanensis*a révélé des propriétés anticoagulantes très intéressantes (Pierre et al., 2015). Ainsi le marché des agars pourrait connaitre un essor important, à condition d'avoir une meilleure connaissance des structures chimiques précises des agars et de leurs propriétés associées. Dans la partie qui suit, je vais présenter ce qu'est un agar, l'état actuel des connaissances concernant son origine, ses propriétés associées et comment approfondir nos connaissances à son sujet.

I. 2. B. a. Historique des agars

Les agars auraient été découverts par Minoya Tarozaemon en 1658 au Japon (Armisen and Galatas, 1987). Une légende est même rattachée à l'agar : «Un empereur japonais et sa suite royale étaient perdus dans les montagnes pendant une tempête de neige. Arrivant dans une petite auberge, ils furent cérémonieusement reçus par l'aubergiste qui leur offrit un plat de gelée aux algues. Qu'il y en eut trop ou qu'elle n'ait pas été délectable, l'histoire ne le dis pas, mais de la gelée fut jetée dehors. Elle gela alors durant la nuit puis s'effrita du fait de la décongélation et de la dessiccation, laissant une substance craquelée de faible densité. L'hôtelier prit ces résidus et

découvrit de façon surprenante qu'en la faisant bouillir avec plus d'eau la gelée pouvait être refaite. » (Armisen and Galatas, 1987).

L'isolement d'un agar formant un gel très dense (agarose) et son analyse structurale viennent quant à eux bien plus tard respectivement, en 1937 et 1956 (Araki, 1956).

I. 2. B. b. Nomenclature des agars.

Définition des termes se référant aux agars

Au départ la seule définition associée au terme 'agar' était liée à la répétition du dimère: 3-O-β-D-galactopyranose-(1→4)-3,6-α-L-anhydrogalactopyranose (ou agarobiose), c'est à dire l'unité 'idéale' de l'agarose (Lahaye, 2001). Un polymère composé quasi exclusivement par ce dimère, et ne présentant que peu de modifications telles que les sulfatations ou les méthylations, est nommée agarose et possèdent un excellent pouvoir de gélification. Cette définition a été battue en brèche par la découverte d'agars présentant de nombreuses modifications telles que méthylations, sulfatation, pyruvilation… (Araki et al., 1967)

La nomenclature actuelle des agars repose sur leur capacité à gélifier. La définition préconisée par Craigie en 1990 est la suivante : les polysaccharides provenant des algues agarophytes sont les agarocolloïdes. La famille des agarocolloïdes est sous-divisée en deux groupes, ceux qui forment des gels : les agars ; et ceux qui n'en forment pas ou seulement des gels de faible qualité : les agaroïdes. D'autres termes se sont greffés à cette définition comme agaropectine (qui renvoie aux agarocolloïdes non gélifiants) (Lahaye 2001), ou encore agarane (synonyme d'agarocolloïdes) (Ramawat and Mérillon, 2015). Cependant, ces termes sont loin d'être toujours universels et clairs. En effet, dans sa revue de 2011, Usov (Usov, 2011) définit l'agarane comme un polysaccharide constitué de 3-O-β-D-galactopyranose-(1→4)-α-L-galactopyranose et non pas comme un terme générique d'agarocolloïdes. De plus, notons que cette définition des agars est hautement subjective, car d'un côté nous avons les polysaccharides formant un gel dit « fort » et de l'autre un polysaccharide non gélifiant ou formant un gel dit « faible ».

En plus des agars formant des gels, un agar modifié ou agaroïdes a été très étudié : le porphyrane. Il est constitué par un dimère de répétition qui est 3-O-β-D-galactopyranose-(1→4)-α-L-galactopyranose-6-sulfate. Comme ce motif est retrouvé chez tous les agars extrait de algues de l'ordre des *Porphyra*, les agars comportant ce motifs ont été nommé en fonction, d'où le terme porphyrane (Craigie, 1990).

Une dernière définition est à aborder pour clore cette partie. Que considérons-nous dans ce manuscrit comme un agar *stricto sensu* ? La répétition d'un dimère de 3-O-β-D-galactopyranose-1→4-α-L-galactopyranose, ou un agarane selon Knutsen et Lahaye, ou encore un agarocolloïde selon Craigie (Knutsen et al., 1994) (Craigie, 1990) (Lahaye, 2001). Pour la suite, nous avons choisi la première de ces définitions. De ce fait, le pont 3,6 anhydro, les sulfatations, les pyruvylations (pour ne nommer que les groupements les plus fréquentes), font partie de la catégorie des modifications.

Nomenclature des unités de répétition

Les agars et les carraghénanes étant des polysaccharides complexes, Knutsen et collaborateurs ont mis au point une nomenclature permettant de décrire les constituants de la chaine de façon plus aisée (Knutsen et al., 1994). Nous allons utiliser cette nomenclature par la suite, une liste des différentes annotations ainsi que les espèces auxquelles elles se rapportent sont disponibles **Tableau 1**.

CODE	NOM IUPAC
D	→ 4) α-D-galactopyranosyl (typique des carraghénanes)
DA	→4) 3,6-anhydro-α-D-galactopyranosyl (typiques des carraghénanes)
G	→3) β-D-galactopyranosyl
L	→4) α-L-galactopyranosyl (typique des agars)
LA	→4) 3,6-anhydro-α-L-galactopyranosyl (typique des agars)
M	O-méthyle
P	4,6-O-(1-carboxy ethylidene)
S	Ester sulfate
L2M, 6S	→ 4) 2-O-méthyl-α-L-galactopyranosyl 6-sulfate
G6[L4M]	→ 3) 6-O-(4-O-méthyl-α-L-galactopyranosyl)- β-D-galactopyranosyle
-G4S-DA-G	→3) β-D-galactopyranosyle 4-sulfate (1→4) 3,6 anhydro-α-D-galactopyranosyle-(1→3) β-D-galactopyranosyle

Tableau 1 : Présentation du code mono-lettre pour les composants des galactanes sulfatés (Adapté de Knutsen et al,. 1994)

Les trois lignes inférieures représentent des exemples d'intégration d'une information complexe comme un branchement, exemple avec G6 [L4M], *via* cette convention d'écriture. Concernant le 4,6-O-(1-carboxy ethylidene), le terme plus générique sous lequel il est retrouvé est pyruvate sous forme acétale.

D'après cette nomenclature, l'agarose est constitué par la répétition de G-LA (3-O-β-D-galactopyranose-(1→4)-3,6-α-L-anhydrogalactopyranose). Cette nomenclature est idéale pour parler d'un seul disaccharide de répétition mais ne permet pas de définir une famille regroupant plusieurs polysaccharides avec un motif majoritaire commun comme c'est le cas pour les carraghénanes (**Tableau 2**) (Kutsen et al., 1994). Il manque donc un niveau de hiérarchisation dans la nomenclature des agars.

Nom grec	Lettre grec	Code du dimère de répétition selon Knutsen
Iota	ι	G4S-DA2S
Kappa	κ	G4S-DA
Lambda	λ	G2S-D2S, 6S

Tableau 2: Présentation de trois familles de carraghénanes et de la nomenclature qui leur est rattachée.

Ces trois familles sont particulièrement connues à cause de leurs importances commerciales (McHugh, 2003). Ainsi, les carraghénanes ayant les modifications décrites appartiennent à une famille. D'autres modifications peuvent faire qu'il ne s'agit pas de κ pure par exemple.

I. 2. B. c Structure chimique et tridimensionnelle de l'agarose.

Les agars sont constitués par une longue chaine de galactose, avec comme élément remarquable l'alternance de l'énantiomère D et L. Une alternance des liaisons est également typique de ces galactanes, en effet le D-galactopyranose est lié au L-galactose situé coté réducteur par une liaison de type β(1-4) et lié avec l'unité L-galactose coté non-réducteur par une liaison α(1-3) (Araki, 1966) (**Figure 17**). L'agarose a été le premier agar dont la structure a été décrite, et donc la structure des agars, définis à cette époque, impliquait l'unité L-galactose sous forme 3,6-anhydro-L-galactose (Araki, 1966) (Lahaye, 2001). En fonction de l'unité première du dimère (L ou D-galactose), nous parlerons du motif néoagarobiose ou agarobiose (**Figure 17**) (Araki, 1966). Dans tout le manuscrit, nous utiliserons deux notions : celles des motifs de répétition (composé de deux sucres) et la notion de degré de polymérisation (DP). Un motif composé de la répétition caractéristique de l'agarose représente un DP de deux (3-O-β-D-galactopyranose-(1→4)-3,6-α-L-anhydrogalactopyranose donc DP2).

Mais ce polysaccharide modèle n'existe que peu à l'état naturel et présente souvent de nombreuses modifications : sulfatation, méthylation, pyruvylation, présence d'un acide glucuronique, d'un branchement xylose et plus rarement galactose (Usov, 2011) (**Figure 17**). Par exemple, dans certaines algues rouges, telle que *Porphyra* sp., le motif de sulfatation sur le O_6 du L-galactose (qui est le précurseur biologique de l'unité contenant le pont 3,6 anhydro) peut être présent dans la chaîne jusqu'à 30%.

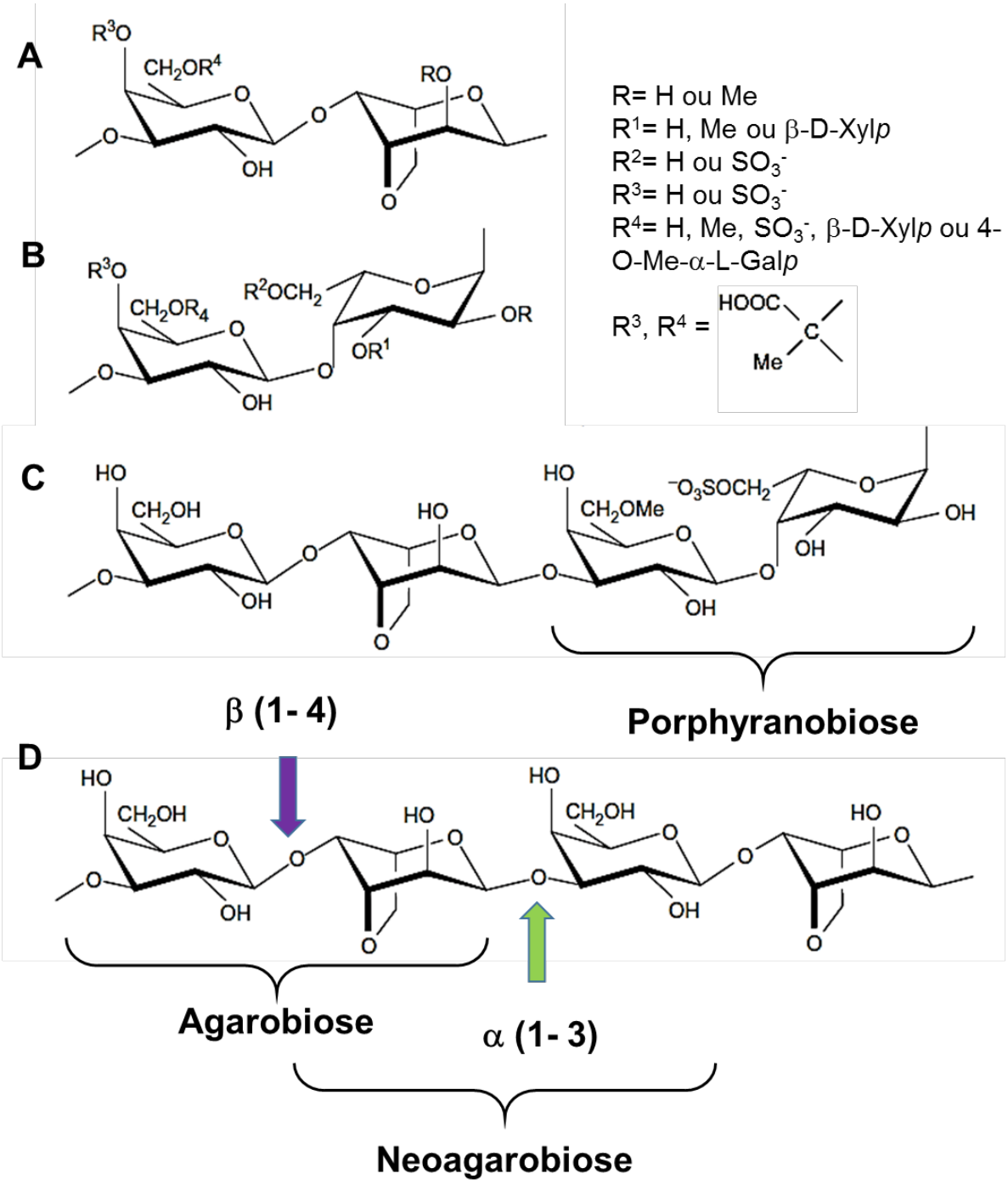


Figure 17 : Diversité des structures chimiques chez les agars (adapté de Usov, 2011).

Panel A et B : Dimères retrouvés avec les différentes modifications possibles. La modification référencée R3, R4 symbolise la présence d'un pyruvate sous forme acétal ou 4,6-O-(1-carboxy ethylidene). Panel C : Présentation d'une chaine de porphyrane avec le motif porphyranobiose. Le porphyrane est un polymère hétérogène dans le sens où les motifs porphyranobiose côtoient les motifs agarobioses (comme figuré dans le panel) dans un ratio 3 :1 environ (Correc et al., 2011). Panel D : Présentation d'une chaine idéale d'agarose avec les noms des motifs et les types de liaison.

Pourquoi l'agarose gélifie-t'il aussi bien ? Grâce à sa structure qui comprend des ponts 3,6-anhydro répétés, ce qui confèrent un caractère plus hydrophobe à la chaîne. Bien que la question reste ouverte, il semblerait que l'agarose forme des doubles hélices antiparallèles. Cette hypothèse est soutenue par différents travaux et notamment l'étude par diffraction aux rayons X de fibres d'agarose (Arnott et al., 1974), mais est aussi supportée par la structure cristallographique d'une β-agarase mutée (*Zg*AgaA) en complexe avec deux molécules d'oligo-agaroses (Allouch et al., 2004).

L'hélice d'agarose est à pas droit et a comme particularité une cavité centrale suivant l'axe de l'hélice, qui serait maintenue par des liaisons hydrogènes entre le β-D-galactose et le 3,6-anhydro-L-galactose, mais également par des liaisons hydrogènes avec des molécules d'eau situées au niveau de la cavité centrale (Arnott et al., 1974). Néanmoins, d'autres études réalisées, soit sur des films d'agar déshydraté, soit sur des agarobioses (**Figure 18**) suggèrent des structures de type simple hélice ou encore des brins β parallèle (Jimenez-Barbero et al., 1989).

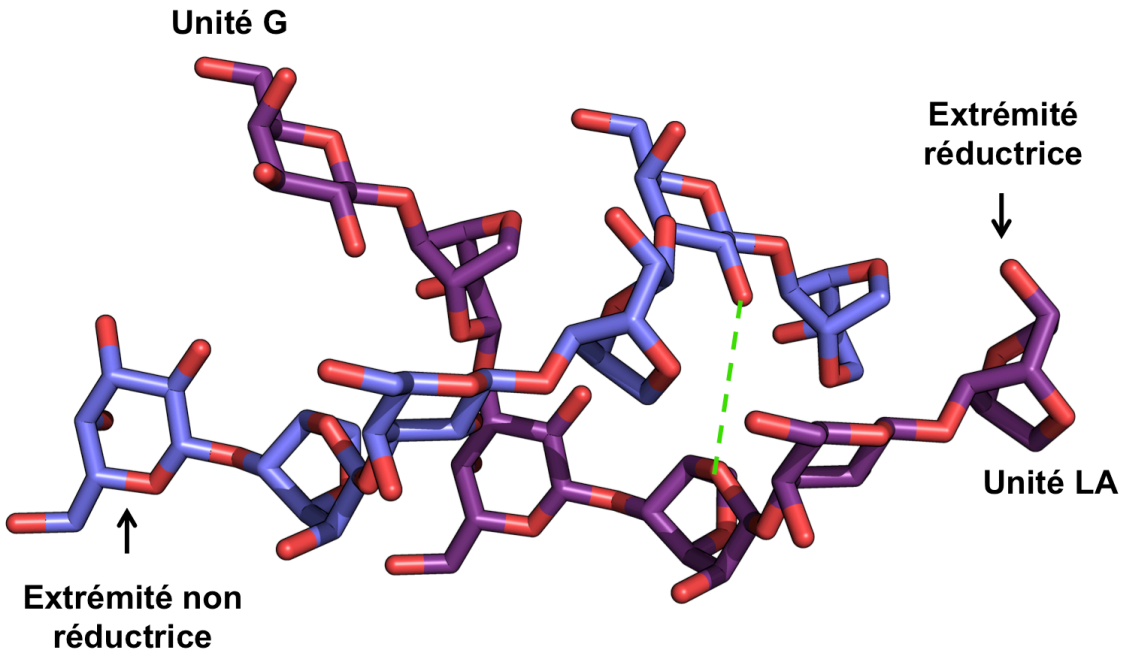


Figure 18 : Modèle de double hélice d'agarose proposé par Arnott et al, (Arnott et al, 1974) (PDB : 1AGA)

Sur ce modèle de double hélice, une liaison hydrogène entre l'O2 d'une unité G (en bleu) et l'O5 d'une unité LA (en violet) est proposée pour permettre le maintien de la structure (trait discontinu vert).

En l'absence du pont 3,6-anhydro, cette structuration ne serait plus possible. En effet, les liaisons hydrogènes impliquent l'oxygène en position 2 du D-galactose et l'oxygène en position 5 du L-galactose, mais en l'absence du pont 3,6 anhydro, l'oxygène en position 5 du L-galactose ne serait plus correctement orienté de façon à permettre cette liaison. La présence de nombreuses modifications crée une gêne stérique, et induit des coudes au sein de la structure, empêchant la formation des liaisons hydrogène (Perez and Kouwijzer, 1999). De ce fait un traitement alcalin entrainant l'élimination des groupements sulfates augmente la qualité du gel obtenu avec l'agar ainsi traité (Armisen and Galatas, 1987).

I. 2. B. d Quel intérêt à l'étude systématique des agars modifiés ?

Les algues de l'ordre des *Gelidiales* sont réputées comme produisant les agars de meilleure qualité d'un point de vue commercial (i.e. un agarose présentant peu de modifications). L'ordre des Gracillariales produit également un agarose de bonne qualité. Mais avant l'avènement des méthodes alcalines de modification de l'agar, celui extrait à partir des Gracillaires (dit agaroide car plus mou que celui du *Gelidium*) était l'autre source d'agarose pour l'industrie (Armisen and Galatas, 1987). Ainsi les études sur les agars se sont majoritairement focalisées sur les algues appartenant à ces ordres, afin de pouvoir obtenir un agarose de bonne qualité sans avoir recours à des modifications chimiques et des purifications couteuses. Mais aujourd'hui il apparait de plus en plus que les agars ont d'autres propriétés telles qu'agent anti cancéreux, anti viral… (Cosenza et al., 2017) Les agars ayant ces propriétés sont majoritairement des agars avec de nombreuses modifications. L'utilisation de ces composés à fortes valeurs ajoutées a généré de nouvelles études qui ont révélé la diversité de structure des agars. Pour ne citer qu'un exemple d'agars très modifié, celui de *Palissada flagilifera* peut avoir jusqu'à dix-huit combinaisons de modifications différentes pour un seul disaccharide de répétition (Ferreira et al., 2012).

La découverte d'agar à effet thérapeutique n'est pas la seule raison qui pousse à explorer la diversité structurelle de ces derniers. En effet, des modifications sont dues à la présence d'enzymes spécifiques au sein de chaque espèce d'algue, qui traduisent le métabolisme et l'adaptation de cette dernière à un environnement précis. On peut

également noter que les gènes codant pour ces protéines sont issus de gènes probablement hérités d'un ancêtre commun à toutes les espèces du même ordre. En partant de cette observation, l'étude des modifications des polysaccharides agaroïdes dans la paroi algale peut nous fournir de précieux indices sur l'évolution et la taxonomie de ces organismes.

I. 2. B. e Introduction à la chémotaxonomie

De premières études chémotaxonomiques ont déjà été menées sur les algues carraghénophytes. A titre d'exemple, nous pouvons citer le travail d'Ian Miller sur le galactane provenant de *Sarcodia montagneana* et *Sarcodia flabellata* (Miller, 2003). Il avait tout d'abord été décrit que ces espèces produisaient du λ-carraghénane (Liao et al., 1993), ce qui était cohérent avec l'appartenance de la famille de *Sarcodiaceae* incluse dans l'ordre des *Gigartinales*, ordre dont les membres sont supposés produire des carraghénanes.

Mais ces résultats allaient à l'encontre de la phylogénie (établie suite à l'analyse de séquence rbcL) qui concluait que les *Trematocarpus* (ordre dont les galactoses liés en 4 sont majoritairement de configuration L) et les *Sarcodia* étaient phylogénétiquement proches et aussi très proches de l'ordre des *Plocamiales* (ordre dont plusieurs espèces comportent des carragars). L'étude RMN menée par Miller sur les *Sarcodia* a démontré par la suite que ces polysaccharides étaient en réalité des carragars complexes (avec une dominance forte de L-galactose) ressemblant fortement à ceux qui sont produits par les algues de la famille des *Trematocarpus* (Miller, 2003). Ainsi, l'analyse des polysaccharides est concordante avec l'analyse génétique et confirme que les *Sarcodiaceae* n'appartiennent pas à l'ordre des *Gigartinales*. L'analyse de la structure des polysaccharides pariétaux peut donc permettre l'amélioration de la classification des algues, mais pour cela, il est nécessaire de disposer de données exhaustives sur les composants pariétaux des algues.

I. 2. B. f. Vue générale de la diversité chimique des agars.

Analyse bibliographique

Le travail de Craigie en 1990 fait référence sur la diversité des structures disponibles. Il présente en effet les travaux qui ont permis d'identifier les motifs présents chez 37 algues différentes appartenant à six ordres différents. Cette revue précurseur permet de dégager certaines conclusions.

Tout d'abord, elle indique que les méthylations sont très fréquentes, particulièrement sur l'unité G en position 6, et de façon plus anecdotique en position 2 de l'unité L. Les sulfatations sont également souvent présentes mais en proportion moindre, souvent sur l'unité L en position 6, et dans une moindre mesure en position 2.

Les motifs L6S sont trouvés chez toutes les algues du genre *Porphyra* ; ce motif (G-L6S) est en effet le motif caractéristique de l'agar issu des *Porphyra* et nommé porphyrane. Le L6S est le précurseur du pont 3,6 anhydro-galactose (Génicot-Joncour et al., 2009) mais l'absence de cette modification au sein des *Porphyra* suggère que cette modification tient également un autre rôle.

Le croisement des modifications des algues rouges avec leur phylogénie de 2016 (Yang et al., 2016) nous permet également de dire, que les algues à la base de l'arbre possèdent une forte proportion du motif G-L6S (comme les *Porphyra* sp. et les Bangiales). En revanche, plus les algues sont « évoluées », plus ce motif devient minoritaire, et plus les motifs deviennent complexes, avec des méthylations sur l'unité L et des sulfatations sur l'unité G (en position 4 et 2). L'exemple le plus marquant concerne *Laurencia pinnatifida* (ancienne nomenclature d'*Osmundea pinnatifida*, l'une de nos algues d'intérêt). Les motifs retrouvés sont en effet très variés, 6M et 2S pour l'unité G et 2M, 2M6S, LA2M et 6S pour l'unité L. Aujourd'hui, cette diversité n'est plus une surprise, car il a en effet été établi que les algues appartenant au complexe *Laurencia*, genre qui comprend les *Osmundea,* les *Laurencia, les Palisada* et les *Chondrophycus,* ont des polysaccharides pariétaux extrêmement modifiés (Cassano et al., 2009) (Ferreira et al., 2012).

D'autres travaux montrent que certains motifs sont communs à un genre. Ainsi, pour l'ordre des *Delesseriaceae*, trois espèces différentes (*Hymenena palmata*, *H.*

marginata et *Myriogramme denticulata*) partagent un motif majoritaire G4S-L2S3S, tandis que d'autres du même ordre (*Hymenena variolosa* et *Acrosorium decumbens*) présentent le motif G2S4S6S-L2S3S (Miller, 2005). Cette étude montre que pour un même ordre, des variabilités existent, mais elles se constituent autour d'un motif commun, ici le motif G4S-L2S3S. Cette étude permet aussi de remettre en question l'un des paradigmes concernant les agars, selon que « les agars sont plus méthylés que sulfatés, contrairement aux carraghénanes» (Percival and McDowell, 1990). Cette observation a été émise à une époque où la majorité des études portaient sur les *Gelidium* et les *Gracilariales*, pour lesquels cette observation est en effet confirmée. Cependant, au vu des études actuelles comme sur les *Delesseriaceae* et les *Laurencia*, il faut nuancer ce propos (Miller, 2005) (Ferreira et al., 2012)(Craigie, 1990). Certes les méthylations sont très présentes chez les agars, mais les sulfatations le sont également.

Etude bibliographique de la composition polysaccharidiques des algues modèles

Concernant les agars des algues agarophytes utilisées lors des travaux de thèse (*Porphyra sp.*, *Polysiphonia elongata*, *Chondria dasyphylla* et *Deleisseria sanguinea*) aucun n'avaient été caractérisés, mis à part les agars issus d'*Osmundea pinnatifida* et *Porphyra sp*.. Néanmoins, certains agars d'algues appartenant au même genre que nos algues d'intérêt l'ont été, et puisque au sein d'un même genre un motif peut être retrouvé chez tous les membres, les motifs caractéristiques peuvent être déduits. Dans le paragraphe suivant, ne sont présentés que les motifs autres que G-LA (agarobiose).

Chez les *Polysiphonia*, les motifs retrouvés par espèce sont pour *P. morrowii* : G6S, G6M, LA, L6S et le dimère G6S-L6S (Usov and Ivanova, 1987) ; pour *P. nigrescens* : G6M, G6S-L6S, G6S-LA et des branchement xyloses sont rapportés (Prado et al., 2008). Concernant *P. abscissoides*, deux motifs sont particulièrement retrouvés G6S-LA et G6M-LA, tandis que pour *P. strictissima* seul le motif G-L6S est observé (Miller et al., 2009). Ainsi, pour les *Polysiphonia* il est légitime de supposer que les motifs G6M/G6S et L6S puisse être retrouvé chez *P. elongata*.

La revue de Craigie, fait référence à un agar issu de *Delesseria crassifolia* qui a été caractérisé. Il s'agit, à notre connaissance, du seul agar issu d'une algue de ce genre à

avoir été chimiquement caractérisé. Les motifs décrits sont G6S, G4S et LA2M (Craigie, 1990).

Pour le genre *Chondria*, un seul polysaccharide issu de ce genre a été caractérisé il a été extrait à partir de *C. macrocarpa*. Les motifs retrouvés sont les suivants G2S, G6S, G2S6S, LA, L6S, L3X6S (Miller and Blunt, 2002). L'intérêt de cette étude n'est pas seulement d'avoir une idée des modifications présentes chez les *Chondria,* mais aussi d'illustrer un point encore non abordé : la variabilité des modifications en fonction de la saison considérée. En effet, dans cette étude les algues ont été récoltées sur deux saisons différentes. L'échantillon collecté en novembre diffère de celui collecté en mars, car il présente un plus faible taux de LA, la présence de pyruvates sous forme acétal, une augmentation de G6M et moins de G2M que l'échantillon de mars (Miller and Blunt, 2002).

Les modifications varient également en fonction des oligoéléments disponibles et de la saison. Des tests ont été menés sur des algues de l'ordre des *Graciliaria*, les plants qui sont cultivés avec peu d'azote, ont une augmentation des motifs L4M et une diminution des motifs G6M et LA par rapport aux plants cultivés dans des conditions non restrictives (Craigie et al., 1984). Ces changements en fonction de l'environnement laissent à penser que la composition de la paroi joue un rôle important pour l'adaptation au milieu. Pour la variation de la composition des agars au cours du cycle de vie, aucun exemple n'a été trouvé. Ce phénomène est, cependant, particulièrement bien documenté pour les carraghénanes. Chez de nombreuses espèces de *Gigartina*, les carraghénanes produit à l'état gamétophyte sont de la famille κ, tandis que ceux produits par le sporophyte sont de la famille λ (Parsons et al., 1977).

Ainsi, travailler sur des agars très complexes, peu documentés et dont la composition chimique peut varier grandement, se révèle être un défi.

I. 2. B. g Agars modifiés : défis et perspectives

Les agars comportent de très nombreuses modifications. Ces dernières varient en fonction de l'algue (et de l'ordre à laquelle elle appartient), de sa nutrition, de son stade de croissance et de la saison. Cette variabilité rend alors complexe toutes les études pouvant être menées sur les agars modifiés, alors même qu'ils ont des

potentiels forts en pharmacologie (De Clercq, 2000). Le problème est de les obtenir en grande quantité sans avoir recours à des processus long et couteux. Comme, expliqué juste avant, les conditions de vie influent sur la composition des agars. Espérer produire des agars avec la structure désirée en jouant sur les conditions de vie des algues est donc une idée difficilement réalisable.

D'un autre côté, la synthèse chimique ne saurait être une solution. Il n'est pas possible de synthétiser chimiquement à volonté des oligosaccharides d'une telle complexité, d'autant plus s'ils sont de grandes tailles. En effet, toutes les fonctions hydroxyles peuvent réagir au cours de la réaction d'élongation obligeant à des réactions de protection/déprotection difficiles et couteuses dès qu'un galactose doit être ajouté. Les étapes de protection/déprotection sont également nécessaires pour la sulfatation/méthylation des positions d'intérêt. Et entre chaque étapes d'élongation ou de modification une étape de purification du produit est nécessaire afin d'éviter des réactions non désirées avant une nouvelle étape d'élongation. Toutes ces étapes rendent quasiment impossible la production à haut rendement des oligosaccharides complexes.

Ainsi, pour obtenir de grande quantité d'oligosaccharides pures pour des utilisations spécifiques nécessite de se tourner vers la nature pour nous fournir les outils de production adaptés. Pour cela, il est nécessaire d'identifier et caractériser les enzymes responsables de la dégradation de l'agar, enzymes qui sont le plus souvent spécifiques des motifs portant les modifications souhaitées.

I. 3 Les enzymes de modification des agars

Pour générer des oligosaccharides à façon, il est nécessaire de comprendre et de caractériser les différentes enzymes qui permettent de les dégrader. Fort de ces connaissances nous pourrons ensuite les mettre à profit. Le propos de cette partie et de présenter deux groupes d'enzymes impliqués dans la dégradation de l'agar. Tout d'abord les glycosides hydrolases (GH) qui permettent le clivage des liaisons glycosidiques et ensuite les sulfatases qui vont éliminer les groupements sulfate.

I. 3. A. Les glycosides hydrolases

I. 3. A. a. *Zobellia galactanivorans* : Un organisme modèle pour l'étude du catabolisme des polysaccharides marins.

Zobellia galactanivorans est une bactérie d'origine marine, isolée en 1988 à Roscoff, à partir de l'algue agarophyte *Delesseria sanguinea.* Cette bactérie a été sélectionnée pour sa capacité à dégrader les galactanes sulfatés (agar et carraghénanes) (Potin et al., 1991). Tout d'abord identifiée comme appartenant au genre *Cytophaga*, et nommée *Cytophaga drobachiensis,* il est ensuite apparu que cette bactérie était phylogénétiquement et phénotypiquement différente. Suite à cette mise en évidence un nouveau genre a été créé : *Zobellia*, dont *Zobellia galactanivorans* est l'espèce type du genre (Barbeyron et al., 2001).

Les bactéries du genre *Zobellia* appartiennent au phylum des *Bacteroidetes*, classe des *Flavobacteriia*, ordre des *Flavobacteriales*, famille des *Flavobacteriaceae*. Le phylum des *Bacteroidetes* est particulièrement connu pour sa capacité à dégrader la matière de haut poids moléculaire et plus particulièrement les polysaccharides (Thomas et al., 2011).

Z. galactanivorans est une bactérie à Gram négatif se présentant sous forme de bacille (0,3-0,4 μM de large sur 3,0-8,0 μM de long). Elle est mésophile, aérobie, hétérotrophe et chimio-organotrophe, ainsi *Z. galactanivorans* a pour source d'énergie des molécules organiques comme le glucose. La couleur jaune-orange de ces colonies est due à un pigment de type flexirubine. Comme elle ne possède pas de flagelle, elle est immobile en phase liquide. *A contrario* en phase solide, elle peut se déplacer par glissement (Barbeyron et al., 2001).

Mais l'aspect le plus intéressant, pour cette étude, est son aptitude à dégrader de nombreux polysaccharides. En effet, elle est capable de se développer en milieu minimum avec comme seule source de carbone : l'alginate, la laminarine, le xylane, le

lichénane…, ce qui fait de cette bactérie un modèle pour la dégradation des polysaccharides d'algues (Barbeyron et al., 2016b).

L'annotation et le séquençage de son génome ont plus que confirmer cette observation. Au sein de son génome, 141 glycoside hydrolases différentes ont été recensées dont une majorité se retrouve dans des opérons putatifs de dégradation des glucides (appelés en anglais Polysaccharide Utilization Loci, PULs) (Barbeyron et al., 2016b).

I. 3. A. b. La classification CAZy et généralités sur les glycoside hydrolases

Comme développé précédemment, la diversité des agars est extrêmement importante, mais il ne s'agit que d'une seule famille de polymère. Il existe de nombreuses autres familles de polysaccharides et chacune d'entre elles comportent différentes modifications. Face à cette diversité, les organismes devant exploiter les glucides ont dû s'adapter et mettre au point un très grand nombre d'enzymes capable de reconnaître, modifier, synthétiser et dégrader ces glycanes.

Classification des glycosides hydrolases

La classification CAZy (Carbohydrates Actives enZymes) a été créée en 1991 (Henrissat, 1991) pour pallier aux problèmes posés par la classification des enzymes (EC) de l'IUBMB (*International Union of Biochemistry and Molecular Biology*). En effet, la classification EC repose uniquement sur la description des activités enzymatiques et ne tient pas compte des séquences protéiques ni de leur structure. Ainsi une même activité (e.g. cellulase) peut être portée par plusieurs familles d'enzymes distinctes. En revanche, les familles de la classification CAZy sont uniquement basées sur la similitude de séquence en acides aminées et donc sont constituées de protéines homologues. Au sein de chaque famille le repliement et les résidus catalytiques sont conservés (Henrissat et al., 1995).

La première classe d'enzymes à avoir été décrite dans cette classification et qui fera l'objet de ce chapitre est la classe des glycosides hydrolases (GH). Ces dernières sont capables de réaliser l'hydrolyse des liaisons glycosidiques ou des réactions de trans-glycosylation (Henrissat and Davies, 1997).

Une autre façon de classer une GH est de se référer à son mode d'action. Ces enzymes peuvent ainsi être subdivisées en *exo*- et *endo*-hydrolases, selon qu'elles réalisent l'hydrolyse des liaisons glycosidiques à l'extrémité de la chaine (mode exolytique), le plus souvent du coté non réducteur, ou au milieu de la chaine (mode endolytique). L'extrémité réductrice correspond dans ce cas à une unité d'une chaine glycosidique dont l'hydroxyle du carbone anomérique est non lié. Lors du clivage d'une liaison osidique, c'est en fonction de la position de cette extrémité que sont définis les sites de fixation et de reconnaissances des oligosaccharides. Une nomenclature des sous-sites de fixation des substrats a ainsi été mise en place (Davies et al., 1997). Les sous-sites à partir du point de clivage en direction de l'extrémité non réductrice de l'oligosaccharide portent des numéros négatifs croissants. Les sous-sites en direction de l'extrémité réductrice portent des numéros positifs (**Figure 19**).

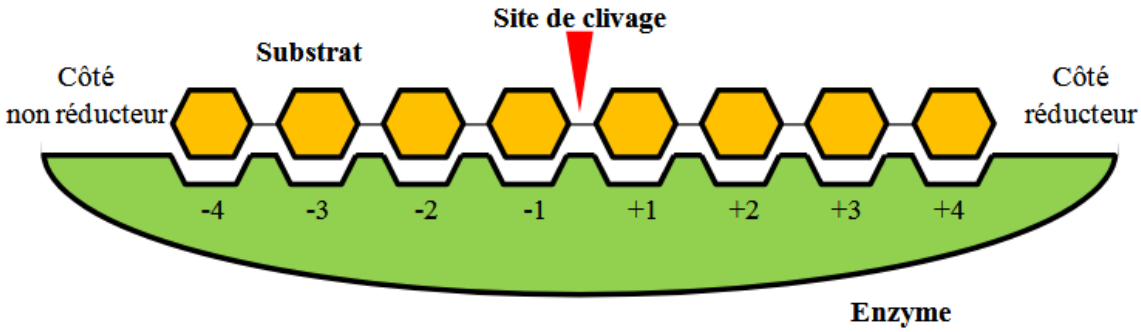


Figure 19 : Nomenclature des sous-sites des glycoside hydrolases (d'après Davies et al., 1997)

Les sous sites sont numérotés en fonction de la liaison glycosidique clivée. La numérotation est négative du côté non réducteur et positive du côté réducteur du polysaccharide substrat.

En fonction de la stéréochimie du mécanisme catalytique, les GHs sont divisées en deux grands groupes : les enzymes qui retiennent la configuration anomérique (carbone C_1) lors de l'hydrolyse ou qui l'inverse. Ainsi dans le cas où le produit à la même configuration que le substrat après la réaction, l'enzyme agit par rétention de configuration (α→α ; β→β) sinon elle agit par inversion (α→β ; β→α). Ces deux types de mécanismes ont été proposés par Koshland en 1953 (Koshland Jr, 1953). Depuis, seules quelques rares exceptions à ces mécanismes ont été découvertes, notamment un nouveau mécanisme fondamentalement différent et impliquant le NADH comme cofacteur (Yip et al., 2004) (Rajan et al., 2004). Au vu des familles de GHs étudiées lors

des travaux de ce manuscrit, nous nous limiterons à la présentation des mécanismes classiques énoncés par Koshland.

Mécanismes catalytiques

La réaction menant à une inversion de configuration est généralement réalisée en une seule étape. Ce mécanisme est qualifié de mono déplacement car à la fin de la réaction le résidu jouant le rôle de base générale se retrouve protoné comme un acide général et inversement pour le résidu jouant le rôle d'acide (**Figure 20**). Le résidu jouant le rôle de base va activer une molécule d'eau en la déprotonant, tandis que le catalyseur acide va céder un proton au groupe partant. Ce schéma réactionnel implique l'existence d'un état de transition sous forme d'ion oxocarbénium stabilisé par des molécules d'eau (catalyse électrostatique), elles-mêmes stabilisées par les résidus acide/base (Koshland Jr, 1953) (McCarter and Withers, 1994).

Le mécanisme de rétention comporte deux étapes et est donc appelé double déplacement. La première étape conduit à la formation d'un complexe glycosyl-enzyme. Pour réaliser la catalyse, l'enzyme dispose d'un résidu jouant le rôle de nucléophile et d'un résidu jouant le rôle d'acide/base générale. Lors de la première étape, le résidu nucléophile (un carboxylate des résidus Asp ou Glu) va attaquer le centre anomérique de la liaison à rompre avec pour conséquence la formation d'une liaison entre l'enzyme et l'extrémité non réductrice de l'oligosaccharide, tandis que le second résidu joue le rôle d'acide en protonant l'oxygène du groupe partant. Une fois sous forme déprotonée, le second résidu peut alors jouer le rôle de base dans la seconde étape pour activer une molécule d'eau qui va attaquer la liaison ester entre l'enzyme et l'oligosaccharide (McIntosh et al., 1996).

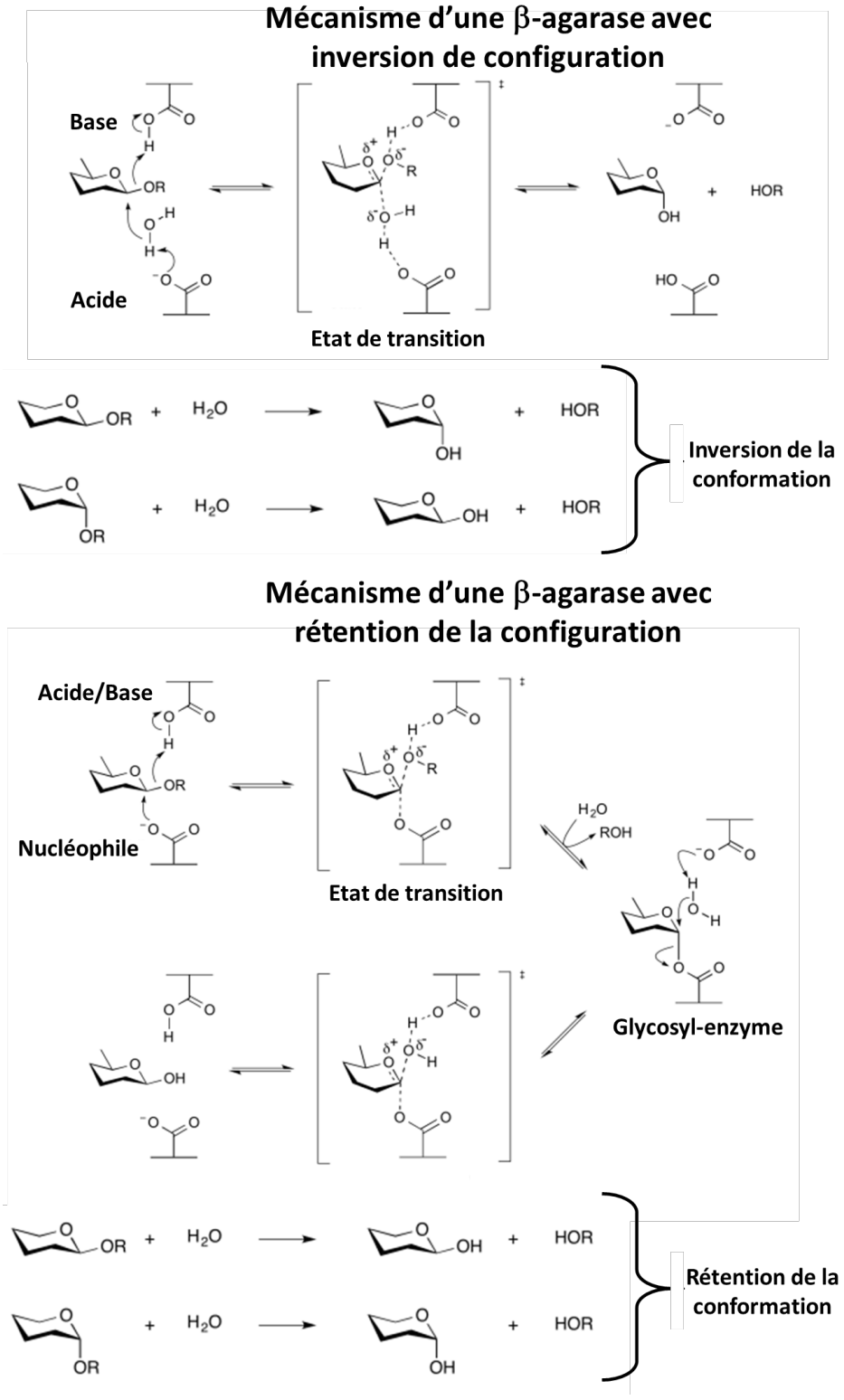


Figure 20 : Schéma réactionnel des glycosides hydrolases agissant par inversion et rétention de la configuration anomérique.

Les différents types de résidus (acide/base, nucléophiles) sont indiqués sur la figure au début de la réaction. Les mécanismes représentés ici sont ceux décrits par Koshland en 1953. L'état de transition se caractérise par l'apparition d'un ion oxocarbénium intracycle. Il s'agit de l'oxygène du cycle pyranose qui va porter une charge partielle positive. Figure réalisée à partir de www.cazypedia.org/index.php/Glycoside_hydrolases

I. 3. A. c. Les loci d'utilisation de polysaccharides

Les loci d'utilisation de polysaccharides (*Polysaccharide Utilisation Locus*, PULs) sont le produit d'une évolution de certaines bactéries pour être plus efficiente lors de la dépolymérisation et la métabolisation des polysaccharides. Cette appellation est apparue initialement au sein des bactéries du phylum des *Bacteroidetes* (Bjursell et al., 2006). Ce terme a été étendu récemment à d'autres phyla bactériens (Grondin et al., 2017) (Ficko-Blean et al., 2017). Cette stratégie implique des transporteurs et l'action synergétique de différentes CAZymes encodées dans des clusters multigéniques inductibles : les PULs. Ces derniers ont particulièrement été étudié chez des bactéries à Gram négatif : les *Bacteroidetes* (Terrapon and Henrissat, 2014). Le modèle de PUL provient du système de catabolisme de l'amidon (« *Starch Utilization System)*» qui a donné l'acronyme 'sus-like' (Anderson and Salyers, 1989a, b). Ce système est composé de huit protéines. Tout d'abord SusR, une protéine régulatrice transmembranaire, va permettre l'activation de la transcription des autres gènes. Les protéines SusD, SusE et SusF sont ancrées dans la membrane externe et interagissent avec la protéine membranaire SusC, qui fait partie de la famille des récepteurs TonB-dépendant. Toutes ces protéines sont impliquées dans la reconnaissance et la fixation du polysaccharide. L'amidon, une fois fixé, est dégradé par SusG, une α-amylase ancrée dans la membrane périplasmique. Les oligosaccharides ainsi produits sont transférés dans l'espace périplasmique *via* SusC. La dernière partie de la dégradation des oligosaccharides est ensuite réalisée par SusA et SusB (Martens et al., 2009) (**Figure 21**).

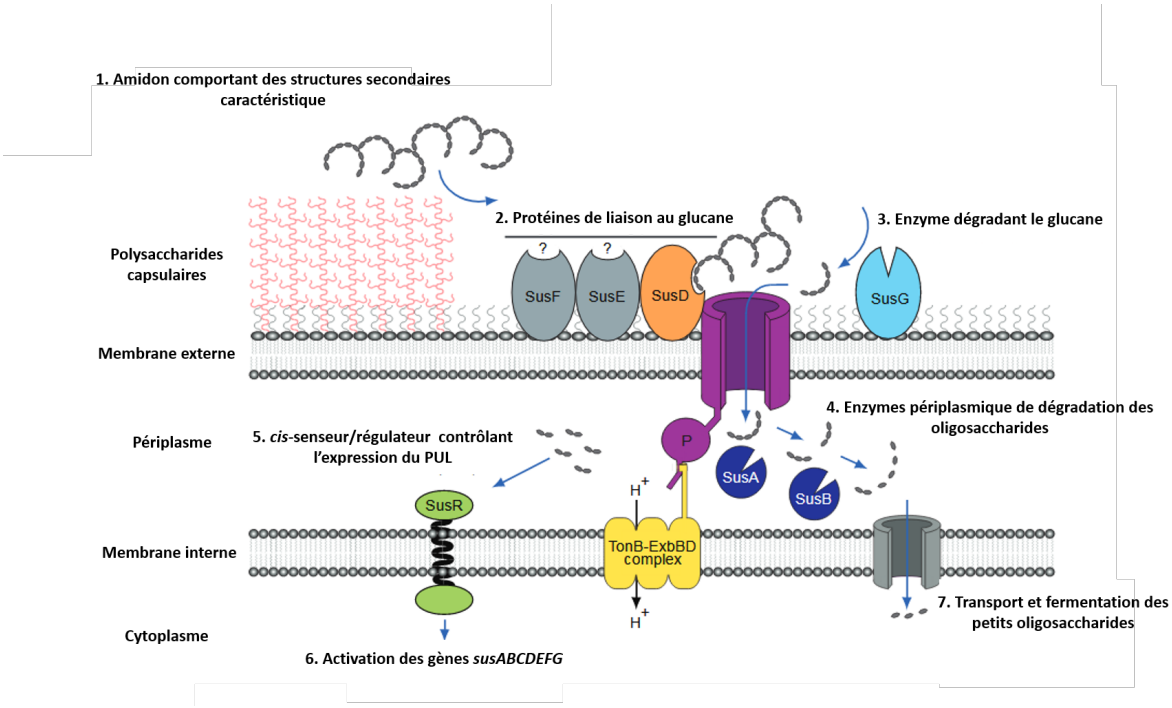


Figure 21 : Modèle de la dégradation de l'amidon par *B. thetaiotaomicron* basé sur les huit gènes du *Starch Utilisation System* (Sus) d'après Martens *et al.*, 2009.

A l'heure actuelle, trois PULs de dégradation de l'agarose ont été prédits chez *Zobellia galactanivorans* (Barbeyron et al, 2016). Les PUL 40 et 42 contiennent tous les deux des glycoside hydrolases de la famille 16 respectivement : *Zg*AgaB (Jam et al., 2005) et *Zg*PorE (β-porphyranase non caractérisée). Le PUL 49 contient une GH117 qui a été identifiée comme l'une des dernières enzymes de la voie du catabolisme de l'agar (Rebuffet et al., 2011) (**Figure 22**). Il est apparu récemment que plusieurs enzymes du PUL 49 étaient des homologues proches d'enzymes du catabolisme des agars (Lee et al., 2014), notamment dédiées au catabolisme du 3,6 anhydro-L-galactose (Yun et al., 2015).

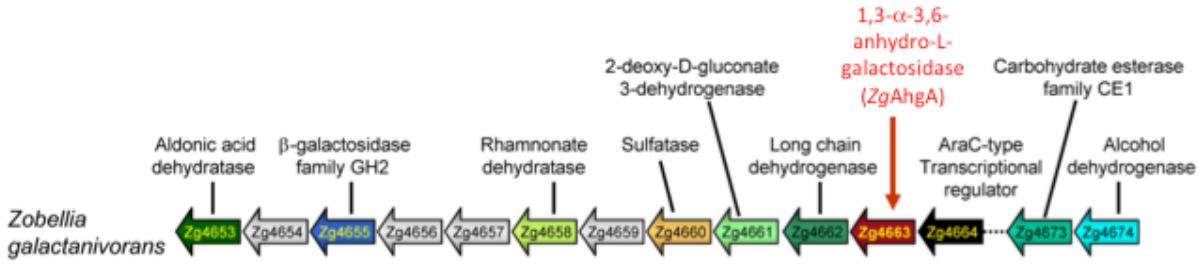


Figure 22 : Représentation schématique du PUL 49 de *Zobellia galactanivorans* (tiré de Rebuffet et al., 2011)

Dans ce PUL putatif, ce trouve *Zg*AghA (*Zg*4663) ainsi que les gènes *Zg*4655, *Zg*4658 et *Zg*4659 tous homologues de gènes impliqués dans la dégradation des agars. Il contient également des gènes qui sont proposés comme codant des transporteurs cytoplasmiques d'ose (*Zg*4654 et *Zg*4656).

I. 3. A. d. Les glycosides hydrolases impliquées dans la dégradation des agars.

A ce jour, il existe cinq familles de GHs contenant des agarases. Les familles 16, 50, 86 et 116 contiennent les β-agarases et la famille 96 regroupe les α-agarases. La famille qui comporte le plus de GHs connues et caractérisées est la famille 16 avec 6564 séquences et 207 enzymes caractérisées toutes spécificités confondues (Donnée tirée de Cazy en Juin 2018). Cinq enzymes de la famille 16 ont été isolées et caractérisées à partir de *Z. galactanivorans* : *Zg*AgaA, *Zg*AgaB, *Zg*AgaD, *Zg*PorA et *Zg*PorB (Jam et al., 2005) (Hehemann et al., 2012a; Hehemann et al., 2012c). Parmi ces enzymes, il n'y a que *Zg*AgaB qui appartient à un PUL.

I. 3. A. d. 1 Famille GH16

Les glycoside hydrolases de la famille 16 sont actives sur de nombreux glucanes et galactanes dont elles vont cliver les liaisons β-1,4 et β-1,3. Dans cette famille, de nombreuses GHs sont d'origines marines telles que les κ-carraghénases, les β-porphyranases et les β-agarases (Michel et al., 2006). Les protéines de cette famille présentent un repliement de type β-jelly-roll. Ce dernier est caractérisé par deux feuillets β antiparallèles incurvés de façon à former une gorge où se fixe le substrat. Tout comme le repliement, les résidus catalytiques sont conservés au sein du motif ExDx(x)E. Le premier glutamate correspond au nucléophile tandis que le second joue le rôle de résidu acide/base. L'aspartate, quant à lui, tiendrait deux rôles. Il a tout d'abord été avancé que ce résidu permettrait de maintenir le nucléophile sous forme activée (Kleywegt et al., 1997). L'autre rôle proposé serait d'accélérer la déglycosylation de l'enzyme (Michel et al., 2001). Ces résidus permettent ainsi la catalyse selon le mécanisme décrit par Koshland (Koshland Jr, 1953) et font que toutes les GH16 agissent par rétention de configuration. Le dernier élément du motif consensus, x entre parenthèses, est absent chez certains membres des GH16. Ce résidu va provoquer un « bourrelet » ou β-bulge au sein du feuillet β (Michel et al., 2001). La raison de sa conservation au sein de certaines GH16 est inconnue (**Figure 23**).

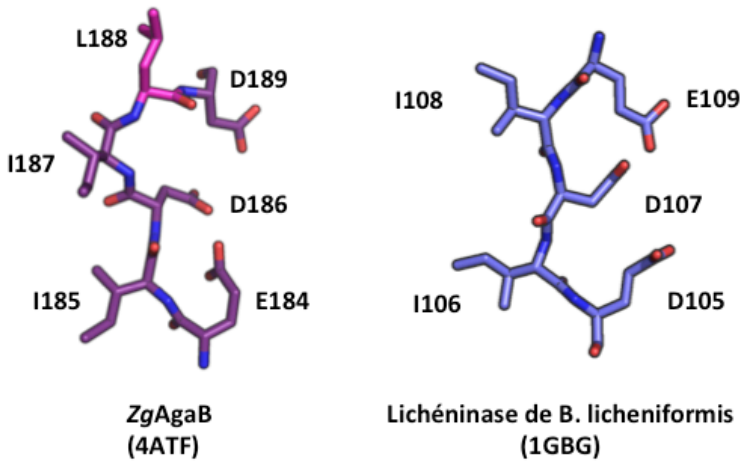


Figure 23 : Visualisation du β-bulge chez *Zg*AgaB (GH16).

Chez les β-agarases le β-bulge est conservé mais chez les lichénases (GH16) ce dernier n'est pas conservé. La présence du β-bulge ne gêne pas le positionnement des résidus catalytiques.

Parmi les enzymes de la famille GH16 caractérisés (Hehemann et al., 2012a,b), peu d'entre elles sont capables de dégrader des agars fortement modifiés. Ainsi le seul agar modifié dégradé par des enzymes est le porphyrane. Les β-porphyranases permettent de cliver les liaisons β-(1,4) lorsque le L-galactose, situé dans les sous-sites +1 et -2, est sulfaté en sur le carbone en position 6 (Correc et al., 2011). Certaines β-agarases sont tolérantes aux L6S si ces derniers sont situés du coté réducteur ou bien en position +4 (*Zg*AgaA, *Zg*AgaB) (Hehemann et al., 2012c) (**Figure 24**). Aucune autre enzyme capable d'accommoder et de cliver un agar comportant d'autres modifications n'a été caractérisée à ce jour.

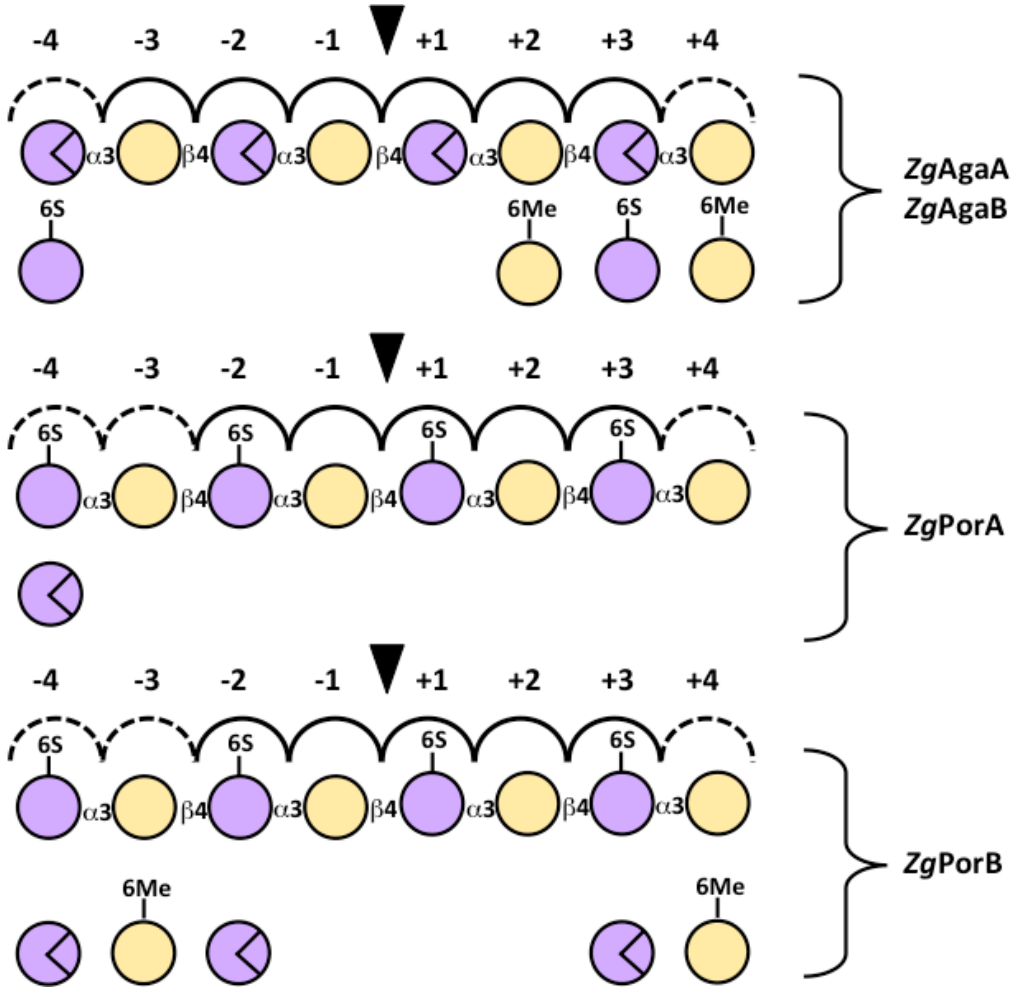


Figure 24 : Représentation des spécificités des sous-sites des β-agarases et β-porphyranases caractérisées chez *Z. galactanivorans*.

Le triangle noir symbolise le site de clivage par l'enzyme. Il se situe au niveau d'une liaison β-(1,4). Les sous-sites sont ainsi numérotés à partir de cette position. La spécificité de chaque enzyme est précisée dans chaque accolade. Seule *Zg*AgaD n'est pas représentée car elle ne reconnaît que l'agarose pur et est totalement intolérante aux modifications. La convention pour la représentation des oses est respectée : jaune pour le D-galactose. Néanmoins comme rien n'est prévu pour les L-galactoses, nous avons choisi de le représenter en violet et quand il s'agit d'un 3,6 anhydro-L-galactose un triangle est ajouté. Si un sous-site peut accommoder différents galactoses modifiés, les autres espèces pouvant être fixées sont ajoutées sous le sous-site en question.

Néanmoins, le travail de Hehemann *et al*., en 2012, laissait présager une β-agarase très différente de celle précédemment décrite. En effet, en réalisant la phylogénie des GH16 présente chez *Zobellia galactanivorans,* il était clairement apparu que *Zg*AgaC a une séquence très divergente par rapport à celles des deux clades regroupant les β-agarases et les β-porphyranases (respectivement 11,13% avec *Zg*AgaA, 10,78% et 18,42% pour *Zg*PorA et *Zg*PorB) (Alignement réalisé avec ClustalW à partir des séquences Uniprot des enzymes). L'enzyme sauvage *Zg*AgaC a été étudiée très succinctement dans un travail essentiellement dédié aux β-agarases *Zg*AgaA et *Zg*AgaB

(Jam et al., 2005). Ce qui a fait de cette enzyme une cible de choix pour améliorer la compréhension des systèmes agarolytiques.

I. 3. A. d. 2 Vues d'ensemble des enzymes et des voies de catabolisme des agars

Le catabolisme des agars se scinde en deux étapes, tout d'abord les polysaccharides sont dégradés par des enzymes extracellulaires et dans un second temps les oligosaccharides ainsi produits sont importés dans le périplasme bactérien pour y être dégradés.

Réactions extracellulaires.

Les glycoside hydrolases de la famille 16 de *Z. galactanivorans* agissent extracellulairement.

La famille GH96 est la seule famille comportant une activité α-agarase décrite à ce jour (Flament et al., 2007). Ces GHs permettent de cliver la liaison α-(1,3) de l'agarose (Potin et al., 1993) mais aussi du porphyrane (*Ta*Aga33) (Ohta et al., 2005). A l'heure actuelle, aucune donnée sur leur mécanisme ou leur structure n'est disponible.

Les familles de β-agarases non traitées de façon exhaustive (GH50, 86 et 118) clivent toutes (sauf la famille 118 pour laquelle l'information n'est pas disponible), à notre connaissance, la liaison β-(1,4) comme les GH16 (Hehemann et al., 2012b) (Ferrara et al., 2014).

Pour la famille GH50, aucune preuve expérimentale n'a été apportée, la stéréochimie du produit et le mécanisme de double déplacement sont proposés car les GH50 font partie du même clan que les GH16 (ClanA). Les clans sont des regroupements de familles de GHs qui sont similaires en terme de structure, de résidus catalytiques et de mécanisme (Henrissat and Bairoch, 1996). A l'exception d'une GH86, toutes les enzymes caractérisées des familles 86 ont un mode d'action de type *endo* (Hehemann et al., 2012b). Comme pour les GH16, la famille GH86 contient aussi une enzyme active sur l'agar modifié, le porphyrane (Hehemann et al., 2012b).

La spécificité de la famille GH50 est de contenir trois β-agarases avec une activité *exo* qui permettent de produire du neoagarobiose (Pluvinage et al., 2013). Cette activité est proposée comme permettant, chez *Saccharophagus degradans*, de clore la phase de

dégradation extracellulaire de l'agarose. En effet, après la première dégradation de l'agar par des enzymes *endo*, comme celles retrouvées chez les GH16 et 86, il est nécessaire de réduire l'oligosaccharide produit à la plus petite unité possible pour permettre un import dans le périplasme. De façon intéressante aucune homologue de ces enzymes n'est retrouvée chez *Z. galactanivorans* ce qui laisse supposer une stratégie différente pour l'acquisition d'oligosaccharides.

Réactions intracellulaires.

Juste après l'internalisation de l'oligoagar intervient une 1,3-α-3,6-anhydro-L-galactosidase (*Zg*AhgA), appartenant à la famille GH117. Cette enzyme agit de façon *exo* et permet de cliver la liaison α-(1,3) du coté non réducteur d'oligo-agars et ainsi de libérer le 3,6 anhydro-L-galactose et des oligoagars impairs (ou du D-galactose dans le cas du neoagarobiose) (**Figure 22**) (Rebuffet et al., 2011).

Une autre enzyme est impliquée dans la dégradation des oligoagars impairs. Chez *Vibrio* sp., une exo-β-galactosidase (*Vej*ABG GH2) permet de libérer le galactose du coté non réducteur en clivant la liaison β-(1,4) (Lee et al., 2014). *Z. galactanivorans* contient deux enzymes homologues *Zg*4655 (48% d'identité de séquence) et *Zg*3633 (33 % d'identité de séquence). *Zg*3633 a été identifiée comme étant une β-galactosidase impliquée dans le catabolisme des oligo-carraghénanes (Ficko-Blean et al., 2017). *Zg*4655 est particulièrement intéressante dans notre étude du catabolisme des agars car son gène est situé dans le PUL49 où se situe *ZgahgA* (Rebuffet et al., 2011). De plus, *Zg*4655 est une β-galactosidase active sur substrat artificiel de type *para*-nitrophénol-Galactose mais également sur le carraghénane. Il a en effet été démontré qu'elle pouvait cliver une liaison β-(1,4) au sein d'un oligo-carraghénane impair libéré par une GH127 ou GH129 (Ficko-Blean et al., 2017). Mais son gène n'est pas induit par les carraghénanes à la différence de *Zg3633*. Par conséquent il est probable que *Zg*4655 ait la même fonction que *Vej*ABG (Lee et al., 2014).

Une première avancée sur les agars modifiés a récemment été réalisée avec la caractérisation de la première méthylase spécifique des agars. Il a en effet été mis en évidence que le 6-O-méthyl-D-galactose pouvait être clivé par une oxydation catalysé par un cytochrome P450 monooxygénase présent chez les bactéries *Formosa agariphila* et *Z. galactanivorans* (Reisky et al., 2018). Cette enzyme contient un hème et nécessite

donc un donneur d'électron tel que la ferredoxine et par conséquent la ferredoxine réductase est requise pour régénérer la ferrodoxine. Cette réaction serait l'une des dernières de la voie du catabolisme des agars avec les réactions de recyclage du 3,6-anhydro-L-galactose (AHG).

La voie de dégradation du AHG a été caractérisée chez *Vibrio* sp., une bactérie marine capable de ne croitre qu'avec du AHG comme source de carbone (Yun et al., 2015). La première étape consiste en une oxydation de la fonction aldéhyde du AHG, disponible lors de l'ouverture du cycle, n'impactant pas le pont 3,6-anhydro. Cette première réaction permet la formation du 3,6-anhydro-L-galactonate (AHGA). Cette réaction est catalysée par une AHG deshydrogénase : VejAHGD, homologue à Zg4659. Les hydroxyles de l'AHGA sont ensuite isomérisés et le pont 3,6-anhydro clivé par une AHGA cycloisomérase (*Vej*ACl, homologue à Zg4658). Le produit ainsi formé est du 2-keto-3-deoxy-galactonate (**Figure 22**). Ce dernier peut être facilement inclus par la suite dans le métabolisme primaire de la bactérie *via* la voie de DeLey-Doudoroff (voie de dégradation oxydative du galactose) (Szumiło, 1981) (Wong and Yao, 1994).

Voie catabolique chez *Zobellia galactanivorans*.

Le paragraphe ci-dessus est un résumé de toutes les enzymes et stratégies connues à disposition des bactéries agarotrophes, mais qu'en est-il chez notre bactérie modèle ?

Z. galactanivorans semble être une bactérie possédant des β-agarases pour commencer la dégradation des agars. De façon plus spécifique, il est certain qu'elle a recours à l'utilisation de GH16 (Hehemann et al., 2012 b) pour réaliser une première dégradation de l'agar. Ensuite se pose la question de l'import intracellulaire des oligosaccharides. En effet, les oligosaccharides produit par *Zg*AgaB et *Zg*AgaA (Jam et al., 2005) sont de grandes tailles, et c'est seulement arrivé à complétion que ces enzymes produisent des petits oligosaccharides de type DP4. En l'absence d'une GH50, comment *Z. galactanivorans* peut-elle intégrer ces premiers produits de dégradation ? La première hypothèse est la présence d'une GH avec une nouvelle spécificité permettant de dégrader les longs oligosaccharides en sucres de plus petit DP pouvant ensuite être facilement internalisés. La seconde est une stratégie récemment décrite, nommée égoïste ou *selfish* en anglais (Cuskin et al., 2015). Cette stratégie repose sur le fait que la bactérie après avoir généré de longs oligosaccharides (>600 Da), les internalise avec un

système de type *sus* et les dégrade une fois dans le périplasme pour éviter la dispersion de petits oligosaccharides facilement assimilables par d'autres bactéries (Reintjes et al., 2017).

Concernant la dégradation dans le périplasme, cela demanderait une confirmation expérimentale, mais *Z. galactinovorans* semble partager avec *Vibrio* sp. la voie de dégradation du 3,6 anhydro-L-galactose (Yun et al., 2015). De même, cette bactérie semble disposer du même système de déméthylation que *Formosa agariphila* (Reisky et al., 2018). Les informations connues sur le catabolisme des agars chez Z. galactanivorans (ainsi que les pourcentages d'identité des protéines impliqués dans les voies caractérisées chez *Vibrio* sp. et *Formosa agariphila)* sont récapitulées dans la **Figure 25**. Les sulfatases, absentes dans ce schéma sont présentées dans le paragraphe suivant.

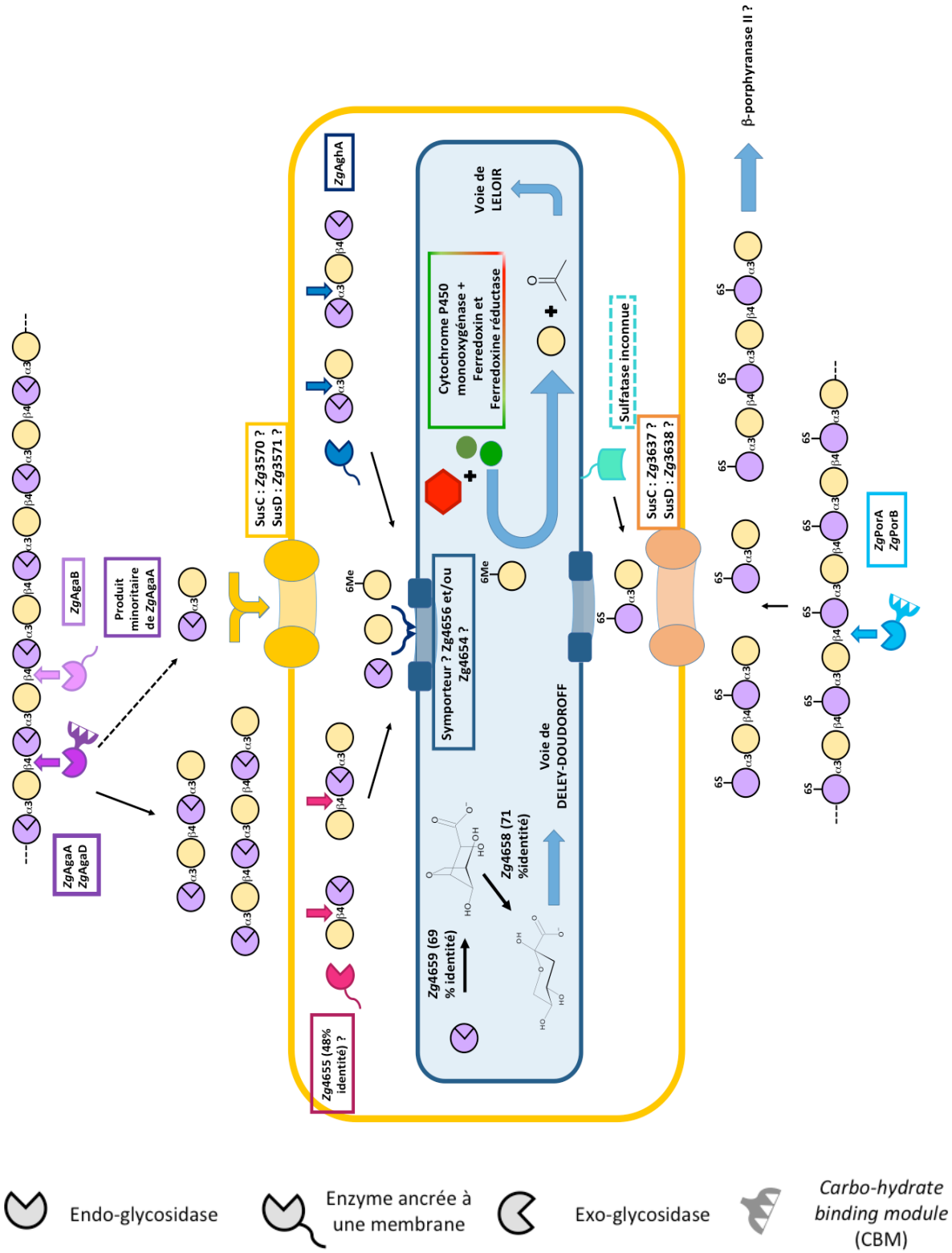


Figure 25 : Modèle de dégradation des agars chez *Zobellia galactanivorans*

Bilan des voies (connues et putatives) cataboliques de l'agar chez *Z. galactanivorans*. Les oligosaccharides générés par *Zg*AgaA, *Zg*AgaB et *Zg*AgaD sont représentés sous forme de DP4 et de DP6, mais nous ne savons pas quelles sont les tailles d'oligosaccharides générés et internalisés par la bactérie physiologiquement. Les protéines pour lesquelles nous avons mis des points d'interrogation et/ou des identités de séquences correspondent aux enzymes qui ont été caractérisés chez d'autres bactéries et dont nous trouvons des homologues chez *Z. galactanivorans*. Concernant la voie de dégradation du porphyrane (en bas de la figure), nous ne savons s'il existe des enzymes intermédiaires à *Zg*PorA et *Zg*PorB qui dégraderaient en petit oligosaccharide les oligosaccharides générés par *Zg*PorA et *Zg*PorB d'où la question de la présence de β-porphyranases II ?

I. 3. B. Les sulfatases actives sur les polymères marins

Afin de dégrader les polysaccharides sulfatés dont les agars, *Zobellia galactanivorans* dispose de 22 PUL contenant au moins une sulfatase (Barbeyron et al., 2016b). Au total, 71 sulfatases ont été dénombrées dans son génome faisant de cette bactérie un modèle pour l'étude de la dégradation de polysaccharides sulfatés d'algues.

Ces enzymes permettent le clivage des liaisons ester-sulfate (CO-SO3-) et sulfamates (CN-SO3-) retrouvés sur un large panel de composés biologiques. Actuellement, il existe quatre classes de sulfatases (Barbeyron et al., 2016a)(**Figure 26**). Ces dernières sont définies en fonction de leurs similitudes de séquence. Mais à ce jour, seule la famille S1 est active sur les carbohydrates. De plus, il s'agit de la seule famille de sulfatases retrouvée chez *Z. galactanivorans*, ainsi ce chapitre n'abordera que cette famille. Une nouvelle base de données (SulfAtlas) ayant pour but de classifier toutes les séquences de sulfatases disponibles est désormais disponible. La famille S1 a été subdivisée en 71 sous familles sur la base d'analyses phylogénétiques (Barbeyron et al., 2016a).

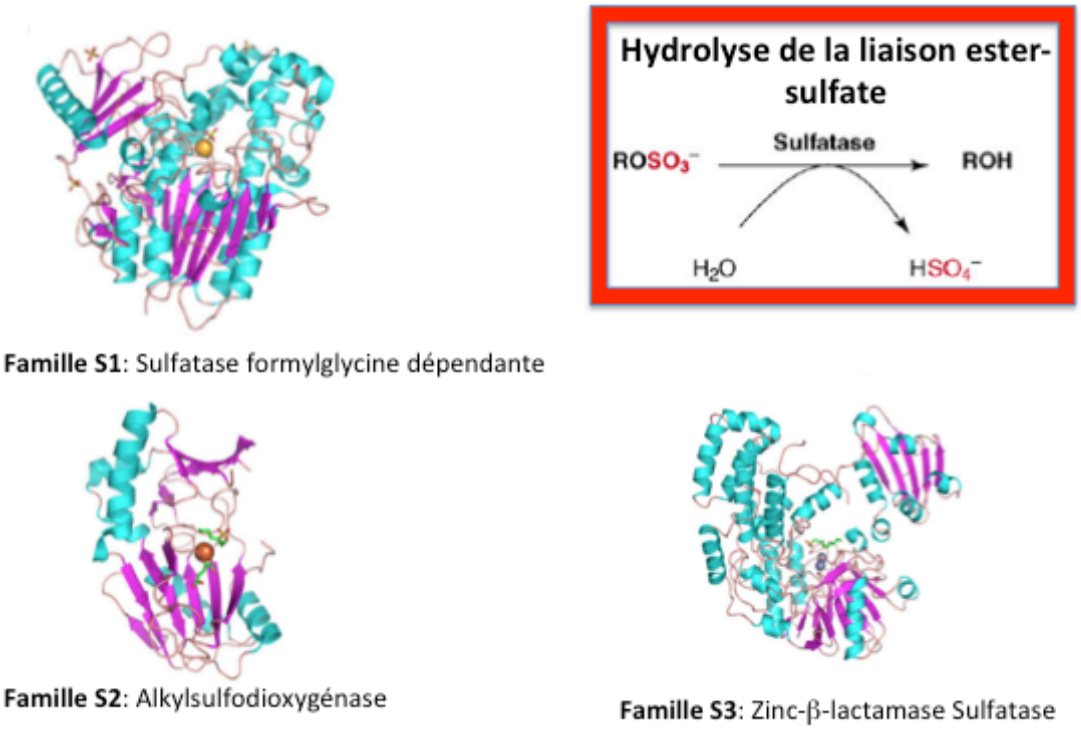


Figure 26 : Les familles de sulfatase

Les structures des sulfatases des différentes familles sont très divergentes. La réaction générale en haut à gauche, ne concerne pas la famille S2 qui ne réalise pas la désulfatation par hydrolyse. Actuellement, il existe 4 familles de sulfatases mais la famille S4 n'est pas caractérisée structuralement.

I. 3. B. a. Famille S1 ou sulfatase formylglycine dépendante.

Les sulfatases de cette famille sont des estérases qui vont procéder par hydrolyse. Pour ce faire, il est nécessaire que l'enzyme comporte une modification co-traductionelle : la formation d'un résidu formylglycine. Il s'agit d'une cystéine ou d'une sérine qui est oxydée de façon à obtenir un aldéhyde au niveau du C_β. Ce résidu unique différencie les sulfatases des autres hydrolases, même si une phosphatase utilise aussi ce résidu (Jonas et al., 2008). Ce résidu a été mis en évidence lors de recherches sur l'étiologie de maladies impliquant des sulfatases comme la maladie d'Austin (maladie de surcharge lysosomale). En effet, des tests de traduction à partir d'ADNc codant des sulfatases dans des cellules de patients atteint de DSM (Déficit Multiple en Sulfatases) menaient à la production de protéines comprenant une activité sulfatase très faible. Ce résultat a orienté des recherches sur l'existence d'une modification co- ou post-traductionnelle indispensable à l'activité. C'est en 1995 que l'équipe de Kurt von Figura identifia, sur des sulfatases digérées par la trypsine, une formylglycine à la place de la cystéine attendue d'après la séquence de l'ADNc (Schmidt et al., 1995).

I. 3. B. b Activation des sulfatases par la *Formylglycine Generating Enzyme* (FGE)

Le processus de modification des sulfatases possède encore de nombreuses zones d'ombres. Les connaissances actuelles sont le résultat des recherches menées sur la FGE humaine (Roeser et al., 2006). Etant donné que les sulfatases S1 eucaryotes et procaryotes montrent un haut degré de similarité notamment au niveau du motif catalytique consensus, il est raisonnable de se baser sur les connaissances acquises sur le modèle eucaryote pour extrapoler sur les sulfatases procaryotes. La modification des sulfatases se déroule dans le lumen du réticulum endoplasmique (RE) ou le peptide néo-synthétisé de la sulfatase est modifié grâce à de l'oxygène moléculaire mais sans aucun cofacteur et ion métallique (Bojarova and Williams, 2008). Une fois le motif consensus des sulfatases S1 eucaryotes CX(P/A)XR identifié, la FGE réaliserait l'oxydation de la cystéine grâce à l'oxygène moléculaire menant à la formation de la formylglycine (**Figure 27**).

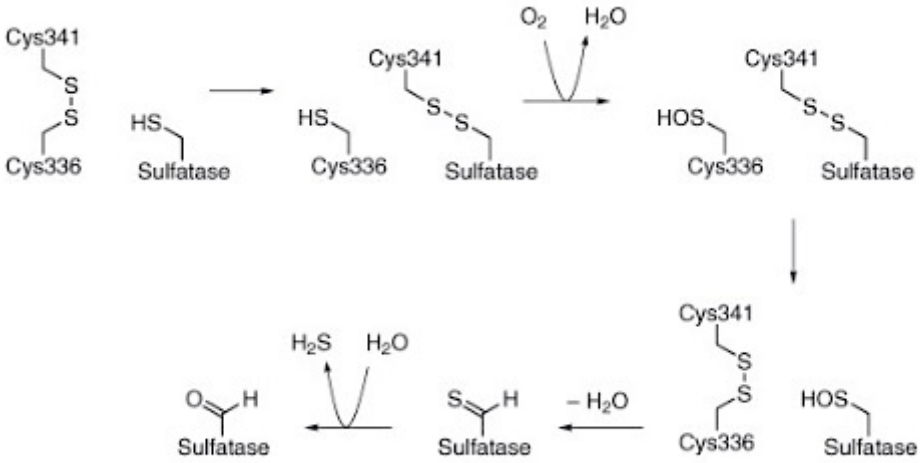


Figure 27 : Mécanisme putatif de la FGE (extrait de Bojarova and Williams, 2008)

Le mécanisme proposé comprend un échange de pont disulfure entre la protéine et le peptide à modifier, Suivi de l'oxydation de la seconde cystéine de la FGE en acide sulfénique (ou en peracide sulfénique SOOH). La quatrième étape consiste en un transfert de l'oxygène sur la sulfatase pour former un acide sulfénique et la reformation du pont disulfure intramoléculaire de la FGE. Une étape d'élimination de l'eau permet de former un thioaldéhyde qu'une hydrolyse transforme en l'aldéhyde.

Une structure cristallographique de l'enzyme humaine a pu être obtenue en complexe avec les peptides CTPSR et LCTPSRA. Cette structure révèle un site actif sous forme de gorge dans laquelle se fixent les acides aminés du motif consensus plus deux acides aminés (Roeser et al., 2006) (**Figure 28**). La façon dont le peptide est reconnu ainsi que la structure globale de la FGE plaide pour que la modification soit réalisée de façon co-traductionelle. En effet, une fois repliés, les résidus du motif consensus de la sulfatase se trouvent dans une conformation incompatible avec la gorge de la FGE et la formylglycine se trouve au fond d'une poche, donc inaccessible. Cette topologie rend impossible la modification une fois que la sulfatase est repliée (Dierks et al., 1998).

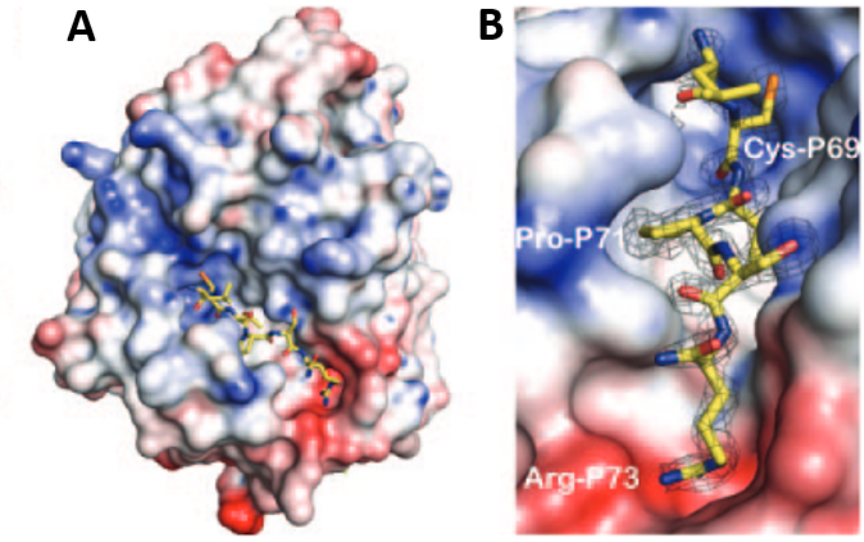


Figure 28 : Structure de la FGE humaine en complexe avec le peptide CTPSR (extrait de Roeser et al., 2006)

La gorge de fixation du peptide est largement ouverte (panel A) et le peptide consensus est particulièrement bien stabilisé au sein de la gorge (panel B).

Le système de modification procaryote peut produire une formylglycine à partir d'une sérine ou d'une cystéine. Ainsi le motif consensus devient pour les procaryotes (C/S)X(P/A)XR. Concernant *Z. galactanivorans*, une seule FGE a été trouvée dans le génome. Elle a été annotée comme étant localisée dans le périplasme. De plus, sa séquence comporte un peptide d'ancrage à la membrane, il donc très probable que la FGE soit ancrée à la membrane cytoplasmique. 61 sulfatases de *Z. galactanivorans* sont prédites comme étant périplasmiques ou ancrées à la membrane. Ainsi, au cours de la traduction la sulfatase serait transloquée sous forme non repliée dans le périplasme, ce qui permettrait à la FGE de la modifier (Barbeyron et al., 2016b).

I. 3. B. c Les sulfatases : mécanismes enzymatiques et structures

I. 3. B. c. 1 Structure tertiaire des sulfatases de la famille S1

Les sulfatases de la famille S1 partagent un repliement commun, même avec une similarité de séquence basse pouvant atteindre les 20% seulement. Certaines sulfatases comportent des domaines transmembranaires permettant leur ancrage dans la membrane, comme l'estrone sulfatase humaine qui possèdent deux hélices α-transmembranaire (Hernandez-Guzman et al., 2003) (PDB : 1P49). A l'exception de ce

domaine, les sulfatases sont des protéines globulaires composées de deux domaines. Le premier domaine est constitué de la partie N-terminale de la sulfatase ainsi que de l'extrémité C-terminale. Ce domaine se présente sous la forme d'un feuillet de brins β mixte isolé et entouré par des hélices α. Certaines sulfatases ont leur extrémité C-terminale incluse sous forme de brin β. Le second domaine contient moins de structures secondaires et est composé d'un feuillet antiparallèle de brin β (4 en moyenne contre 10-11 pour le domaine I) adossé aux hélices α du premier domaine. Ce feuillet est lui aussi entouré par un nombre variable d'hélices α. Le site actif se trouve quant à lui au niveau du domaine N-terminal (**Figure 29**).

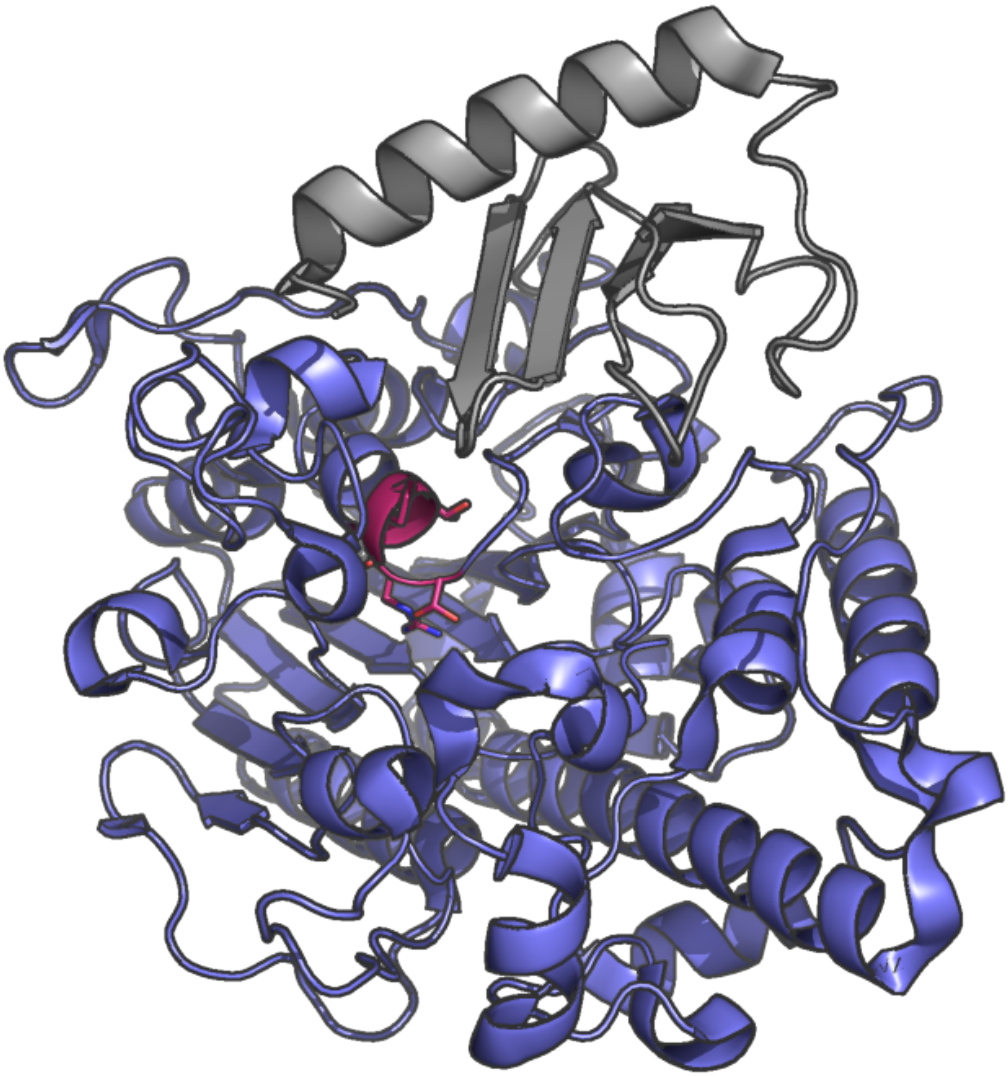


Figure 29 : Structure de l'arylsulfatase A de *P. aeruginosa* (PDB : 1HDH)

Il s'agit de la première sulfatase obtenue à haute résolution (1.3 Å) sur laquelle la formylglycine serait sous forme hydratée dans le site actif. Le motif consensus a été coloré en rose et les résidus sont affichés en bâton. Le domaine N-terminal est en bleu et le C-terminal est en gris. (Boltes et al., 2001)

Une étude menée sur des sulfatases humaines (ARSA, ARSB, STS et PARS) a révélé dix résidus conservés au sein du site actif (Ghosh, 2005). Neuf d'entre eux sont strictement conservés : Asp 35, Asp36, FGly75, Arg79, Lys134, His136, His290, Asp342 et Lys368 (numérotation de la STS). Le dernier résidu est la Gln343 pour la STS tandis que pour les quatre autres il s'agit d'une asparagine. Ce résidu, mis à part, la position des résidus est parfaitement conservé entre ces quatre structures avec un *rmsd* de 0,4 Å. La superposition est encore plus significative pour les chaines latérales des résidus Asp35, Asp36, Asp342 et Gln343 et de leurs équivalents au sein des trois autres structures. En effet, le *rmsd* pour la superposition des chaines latérales de ces résidus est de 0,2 Å. Les oxygènes des chaines latérales de ces résidus interagissent avec un cation divalent. Selon les structures considérées, ce cation est soit un magnésium (Mg^{2+}) (Lukatela et al., 1998) (Boltes et al., 2001), soit un calcium (Ca^{2+}) (Hernandez-Guzman et al., 2003) (Bond et al., 1997). La parfaite conservation de ces résidus accrédite la thèse d'un mécanisme conservé entre toutes les sulfatases.

Le résidu formylglycine (FGly) est essentiel pour réaliser la catalyse et permet un schéma réactionnel bien particulier (Dierks et al., 1998). Le mécanisme a longtemps été sujet à débat et reste encore sujet à interrogations. Néanmoins, il est communément admis que le mécanisme le plus probable est un mécanisme de transestérification-élimination (**Figure 30**). Le cycle catalytique commencerait avec l'hydratation de la FGly pour former un gem-diol (étape A), dont l'un des hydroxyles va réaliser une attaque nucléophile, après avoir été activés par l'aspartate conservé (étape B), sur le soufre de la liaison ester-sulfate ; la forme gem-diol étant stabilisée par des liaisons hydrogène réalisés avec une des histidines conservées (HisA) et un aspartate. La seconde histidine (HisB) et l'ion calcium/magnésium permettraient le positionnement du sulfate. L'attaque nucléophile a pour conséquence la formation d'un adduit sulfate au niveau de la FGly et la libération du composé désulfaté (étape B). Le second hydroxyle du gem-diol va permettre d'éliminer l'adduit et de régénérer la formylglycine. Pour cela, l'histidine A va jouer le rôle de base générale sur l'hydroxyle γ2 permettant un réarrangement électronique et ainsi la rupture de la liaison ester-sulfate (étape C). Il est fort probable que le cation joue un rôle important dans la déstabilisation de la liaison C_{β}-Oγ1 grâce à des interactions électrostatiques (Hanson et al., 2004).

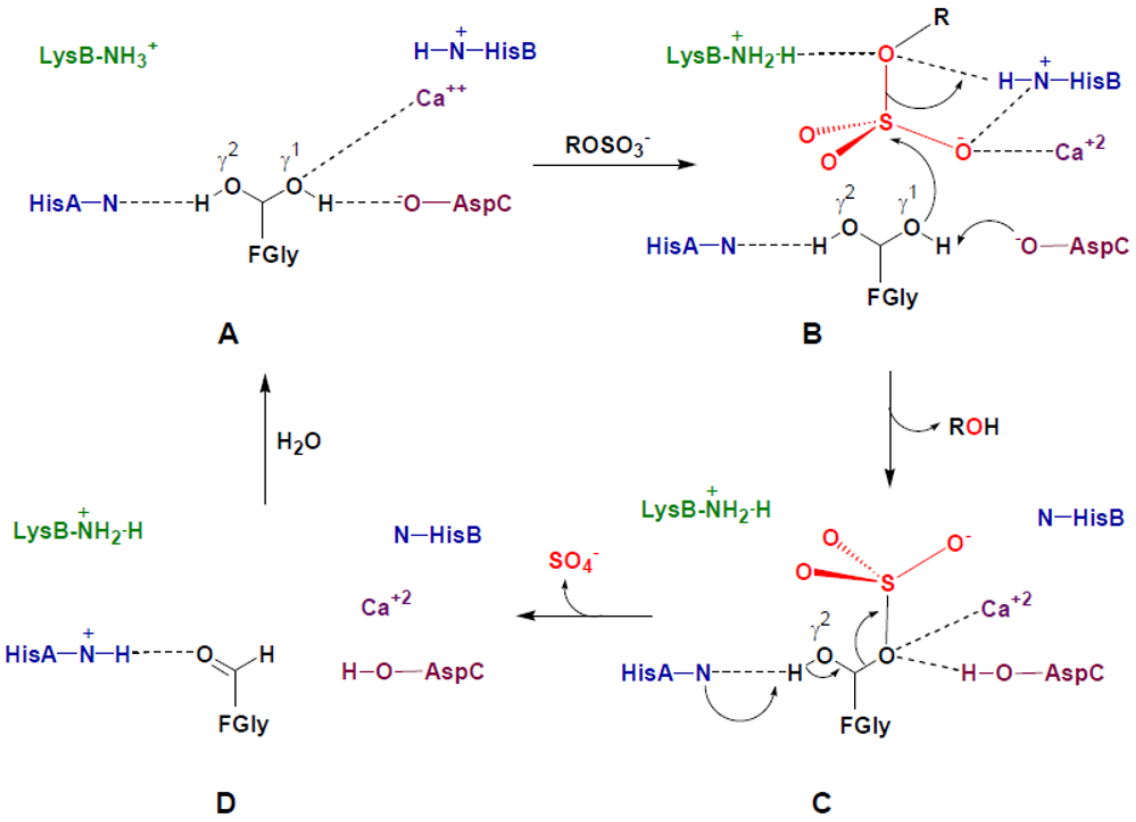


Figure 30 : Mécanisme des sulfatase de la famille 1 proposé par Hanson *et al.*, 2004

I. 3. B. c. 2 Les sulfatases actives sur les polysaccharides d'origine marine.

Etonnamment, parmi toutes les sulfatases permettant de dégrader les galactanes sulfatés (Helbert, 2017), aucune ne provient des algues. De plus, dans les génomes de *Chondrus crispus* et *Porphyra umbilicalis*, aucune sulfatase formylglycine dépendante n'a été prédite (Collen et al., 2013) (Brawley et al., 2017). Ainsi, toutes les sulfatases présentées ici sont d'origines bactériennes. Aucune sulfatase active sur les agars n'a été isolée ou caractérisée à ce jour, mais plusieurs sulfatases agissant sur les carraghénanes sont dénombrés (**Figure 31**). Récemment la voie de dégradation des carraghénanes chez *Z. galactanivorans* ainsi que les sulfatases impliquées ont été caractérisées (Ficko-Blean et al., 2017).

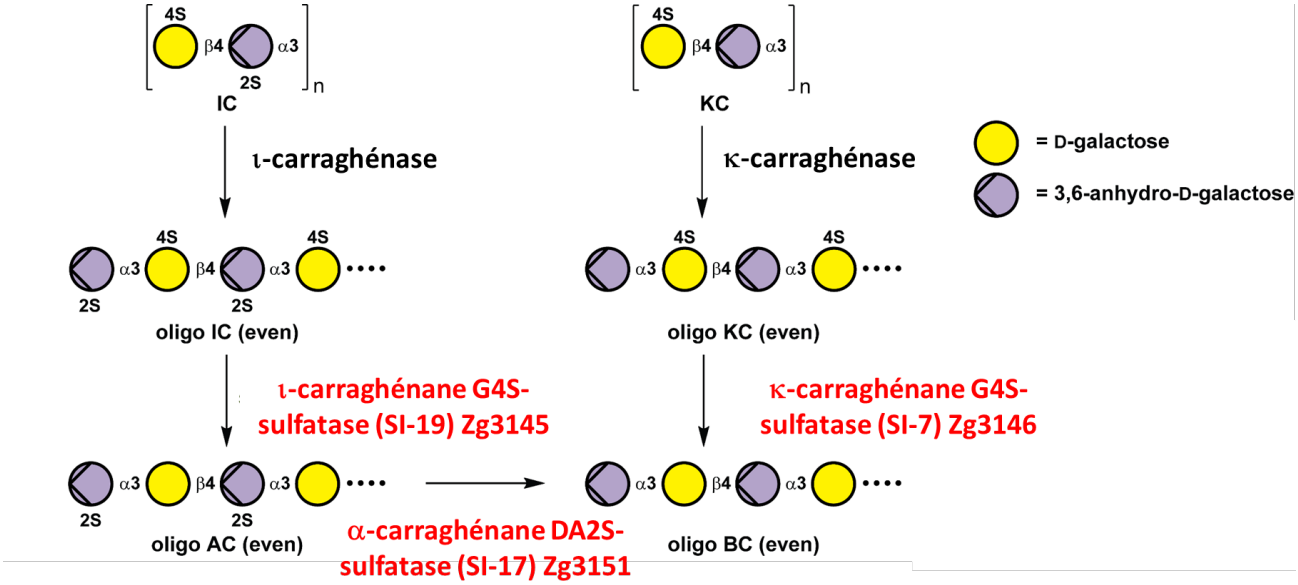


Figure 31 : Voie de dégradation partielle des carraghénanes chez Z. galactanivorans (d'après de Ficko-Blean et al., 2017)

Dans cette voie sont indiquées toutes les sulfatases actives sur les carraghénanes connus (entouré en rouge). Ces dernières sont des homologues de sulfatases déjà caractérisées dans d'autres organismes. La kappa-carraghénane sulfatase, qui clive le sulfate en position 4 du D-galactose mais uniquement du coté non-réducteur, est la première à avoir été décrite en 1966 (Weigl and Yaphe, 1966) puis caractérisée en 1979 (McLean and Williamson, 1979) chez *P. carrageenovora* (enzyme sauvage). La première sulfatase spécifique des carraghénanes à avoir été caractérisée au niveau moléculaire est la iota-carraghénane G4S sulfatase de *Pseudoalteromonas atlantica* permettant ainsi la conversion du iota-carraghénane en alpha-carraghénane (Préchoux et al., 2013). Une deuxième sulfatase a été caractérisée chez *P. atlantica* qui permet de cliver le sulfate en position 4 du D-galactose sur toute la chaine polysaccharidique ce qui en fait une endo-sulfatase (Préchoux et al., 2016). Trois nouvelles sulfatases ont été caractérisées chez *Z. galactanivorans* qui expliquent la désulfatation des kappa- et iota-carraghénanes (Ficko-Blean et al., 2017)

Actuellement, il n'existe que trois structures de sulfatase en complexe avec un oligosaccharide. Les premières parues concernent deux sulfatases appartenant à un PUL dégradant l'héparane sulfate (Cartmell et al., 2017). Ces enzymes ont un mode d'action de type exo. Ces structures ont montré que le positionnement du substrat vis-à-vis des histidines conservées était idéal pour permettre la protonation du groupe partant. Cartmell et coll. ont également défini les premières bases de reconnaissance moléculaire du substrat par une sulfatase. Mais le travail le plus complet sur la reconnaissance sulfatase-substrat a été mené sur une iota-carraghénane G4S-sulfatase (Hettle et al., 2018).

Ainsi, il semblerait que les dix résidus conservés du site actif permettent bien de positionner et cliver le sulfate tandis que ce sont les résidus autour du site actif qui sont déterminants pour la spécificité. En effet, ces résidus annexes, en stabilisant et reconnaissant la partie saccharidique du substrat, détermineraient la spécificité.

Cette étude a par ailleurs proposé de définir des sous-sites en fonction du site de clivage. Les sous-site 0 et S (pour sous-site de liaison au sulfate) étant ceux où sont positionnés respectivement l'ose et le sulfate à cliver. Puis comme pour les glycoside hydrolases, les sous-sites positifs sont ceux du coté réducteur tandis que les sous-sites négatifs sont du coté non réducteur.

I. 4. Objectifs de la thèse

Afin de pouvoir, un jour, produire en quantité industrielle des oligosaccharides avec des propriétés thérapeutiques, il est essentiel de comprendre leur biogénèse et leur catabolisme. L'agar est aujourd'hui un des polymères le plus connu et l'un des plus utilisés notamment au sein de nos laboratoires.

Pourtant nous savons encore peu de choses sur la façon dont il peut être dégradé enzymatiquement. Pour pallier à ce manque, je me suis attaché à caractériser des enzymes agissant sur les agars complexes.

Tout d'abord, nous avons sélectionné une agarase de la famille GH16 qui est très divergente des β-agarases et des β-porphyranases : *Zg*AgaC. En effet, cette enzyme paraissait particulièrement intéressante, tout d'abord de part sa séquence, du fait du peu d'homologie avec les autres β-agarases et β-porphyranases. Ensuite, son originalité s'est confirmée puisque lors de la purification de cette enzyme native depuis le milieu extracellulaire de *Zobellia galactanivorans*, la protéine ne se fixe pas sur l'agarose utilisé dans les colonnes Sephadex 200 mais présente quand même une activité agarolytique (Jam et al., 2005). Une première phylogénie l'avait classée parmi les β-agarases, tout en étant très divergente (Hehemann et al., 2010). Il était légitime de penser qu'elle pouvait avoir une activité uniquement sur des agars complexes. L'objectif de ce premier projet était de combiner différentes méthodes : enzymologie, cristallographie aux rayons X et spectrométrie de masse, afin de caractériser finement l'action de cette enzyme. L'une des étapes cruciales a été de mettre au point une méthode permettant de mettre en évidence rapidement sur quels agars étaient actives *Zg*AgaC par rapport aux GH16 déjà caractérisées. Cette étude a par ailleurs permis de mettre au point une technique de criblage des activités sulfatases sur agar.

Une des grandes inconnues dans le catabolisme des agars concerne l'action des sulfatases. L'autre versant de cette thèse a donc été la caractérisation de nouvelles activités sulfatase. Deux cibles ont pour cela été sélectionnées. La première provient du PUL 42, celui où se situe *Zg*AhgA (Rebuffet et al., 2011) ainsi que les homologues de la voie de dégradation du 3,6-anhydro-L-galactose de *Vibrio* sp. (Yun et al., 2015). En raison de ce contexte génétique, il semblait probable que *Zg*4660 soit active sur l'agar. Mais cette dernière, lorsqu'elle est produite de façon recombinante, est très peu active

sur des substrats artificiels de type 4-méthyllumbelliféryl-sulfate (MUS)(résultats préliminaires antérieurs à ma thèse).

Pour anticiper ces problèmes de manque d'activité de la forme recombinante de *Zg*4660 produite en système hétérologue bactérien, nous avons sélectionné comme seconde cible la sulfatase recombinante la plus active sur substrat artificiel : *Zg*3515. Cette protéine appartient à un PUL qui comprend une GH30 : *Zg*3516 ayant été annotée comme étant une xylanase. Le substrat de ces enzymes semblait relativement aisé à déterminer, puisque du xylane sulfaté est connu chez l'algue rouge *Palmaria palmata* (Deniaud Boüet, 2002).

Les travaux sont présentés sous forme de 3 chapitres indépendants. Le premier présentera l'étude sur *Zg*AgaC sous forme d'un article (Chapitre III). Les deux autres chapitres seront consacrés à l'étude de la sulfatase *Zg*4660 (Chapitre IV) et du couple sulfatase *Zg*3515 / GH30 *Zg*3516 (Chapitre V).

Chapitre II : Matériels et méthodes

II.1 Clonage des différents gènes cibles

II.1. A Vecteurs d'expression

Au cours de notre travail, nous avons utilisé différents plasmides dont les caractéristiques sont résumées dans le **Tableau 3**. Le choix du plasmide était fonction de la cible à cloner et de l'utilisation de la construction résultante. Les stratégies de clonage et les caractéristiques des protéines recombinantes sont également résumées.

Le Plasmide pFO4 est un don du laboratoire du Pr. M. Cygler (Canada). Il possède une cassette de restriction qui permet d'utilisé un grand nombre d'enzymes de restrictions dont Xho1 et Nsi1. En outre, il permet d'exprimer les protéines en fusion avec une étiquette de 6 histidines, placé en N-terminale- afin de faciliter la purification de la protéine recombinante par chromatographie d'affinité. Le vecteur pET-DUET (Société Novagen) est utilisé pour réaliser une co-expression de deux gènes cibles, en effet son origine de replication est compatible avec celle du plasmide pFO4. Le plasmide vecteur pET-DUET- FGE a été réalisé par le Dr. E. Ficko-Blean et D. Mico-Latorre en utilisant le Kit In-Fusion® HD Cloning Plus (Clontech).

Le plasmide pNEA provient de la collection personnelle du Dr. C. Romier (Diebold et al., 2011). Ce plasmide est utilisé pour l'expression périplasmique de protéines recombinantes qui possèdent une étiquette poly-histidine à l'extrémité C-terminale. Le pNEA-3MC-HEXIM a été réalisé par les Dr. A-C. Dock-Bregeon et D. Martinez-Zapien (non publié). Ce dernier contient une version modifié de la MBP en N-terminale. Cette construction présente l'avantage de pouvoir purifié par affinité au Nickel ainsi qu'au maltose. Trois résidus de surface de la MBP ont été modifié et ce dans le but d'améliorer la cristabilité de cette dernière, cette MBP modifié est abrégé 3MC.

Le plasmide pBAD (Thermofisher) est un plasmide dédié à l'expression des protéines toxiques, en effet du fait de son promoteur pBAD les fuites d'expression sont limitées. L'avantage de ce promoteur est que l'expression du gène placé sous son contrôle peut être régulée avec la concentration d'inducteur. De plus, l'arabinose n'est pas un inducteur gratuit ce qui permet de rapidement faire cesser l'expression de la cible et ainsi évité une trop forte surexpression. Le pBAD-PSU a comme insert le gène codant la GFP. Ce plasmide nous a été donné par le Dr F. Leroux.

Plasmide	Expression	Gènes clonés	Clonage	Etiquette protéine recombinante	Utilisations	Remarques
pFO4	Cytoplasmique	*zg3515* *zg3516* *zg4660* *zgAgaC*	Groisillier *et al.*, 2010	N terminal	Caractérisation biochimique structurale	Clonage réalisé par Larocque
pNEA	Périplasmique		Via technique InFusion	C-terminal	Caractérisation biochimique	Plasmide utilisé pour le système de *Zobelli galactanivorans* chez
pBAD	Périplasmique	*zgFGE*	Via technique InFusion	Pas d'étiquette	Activation de la sulfatase *Zg*4600	Plasmide utilisé pour reproduire le système *Z. galactanivorans* ch *coli*
pHEXIM_ construction 1		*zg4660*	Via technique InFusion	N terminal + MBP modifiée+ linker 3 résidus alanine	Cristallisation	
pHEXIM_ construction 2			Via technique InFusion	N terminal + MBP modifiée+ linker 6 ala	Cristallisation	

Tableau 3 : Description des différents plasmides utilisés pour le clonage de toutes les cibles étudiées dans ce manuscrit.

II. 1. B Analyse bio-informatique et définition des séquences à cloner

Pour pouvoir recréer le système d'expression et d'activation des sulfatases chez *E.coli*, il était impératif de conserver l'intégralité des gènes codant pour les protéines. En effet, pour l'expression périplasmique des enzymes ; ou l'inclusion dans la membrane cytoplasmique dans le cas de *Zg*FGE ; il est nécessaire que les protéines recombinantes possèdent les peptides signaux *ad hoc* à l'extrémité N-terminale. Chez *Z. galactanivorans*, les gènes codant pour ces protéines codent également pour ces signaux. Comme il s'agit d'un transfert de gènes entre deux bactéries et que ces signaux sont extrêmement conservés, nous avons décidé d'utiliser les séquences de *Z. galactanivorans*. Pour obtenir la séquence codant pour les protéines et les peptides d'adressage périplasmique, nous avons utilisé eu recours à la banque du génoscope : Mage Microscope. Dans cette banque, les gènes sont annotés et disponibles sous leur forme entière. C'est donc ainsi que nous avons défini les séquences à cloner.

II.1.C ADN Matrice

L'amplification des gènes cibles a été réalisée à partir de l'ADN génomique de *Z. galactanivorans*. Ce dernier a été extrait et purifié par gradient de chlorure de césium par F. Thomas (Thomas, 2011).

II. 1. D Conception des amorces

Le clonage des différents gènes ayant été réalisé en utilisant le kit In-Fusion HD Cloning, la conception des amorces était un peu particulière. En effet, en plus de la séquence spécifique du gène à cloner, les amorces sont dessinées en ajoutant 15 paires de bases correspondantes au vecteur dans lequel la séquence doit être insérée.

Afin de pouvoir réaliser différents clonages en parallèle, nous avons défini une température d'hybridation normalisée fixée à 60°C, comme préconisé par le kit. Les amorces ont toutes été conçues avec l'outil InFusion intégré au programme SnapGene.
Le Tm est calculé par cet outil selon des paramètres spécifiques (non explicités) qui font que le Tm donné ne correspond pas à ce que nous pourrions obtenir avec la formule

2(nombre de A et T) + 4(nombre de G et C). Les différentes amorces utilisées lors des clonages sont répertoriées dans le **Tableau 4.**

A

Gène cloné- vecteur	Amorces	N°amorce	Site de restriction
zg4660-pFO4	5'-TTTTTT**CTCGAG**CAAGACAAACCGAACATTGTTTTTATA-3'	1	XhoI
zg4660-pFO4	5'-TTTTTT**ATGCAT**TTAGTGTTTTCTATATGAATCCATTCG-3'	2	NsiI
zg4660-pNEAVHC	5'-ACTTCCAGGGC**CATATG**ATGGGCAGCAGCCATCATCA-3'	3	NdeI
zg4660-pNEAVHC	5'-TAGACTATTA**GGATCC**GTGTTTTCTATATGAATCCATTCGT-3'	4	BamHI
zgFGE-pBAD	5' GGATGGAGTGAAACGATGTACAATATGATTAAGAGTTATATTCCCCTTGCC 3'	5	
zgFGE-pBAD	5' CTGTCTAGA**CTCGAG**TTATTTTTTTGAATTGCGGATCATATCAGGGG 3'	6	
zgFGE-pBAD	5' CTCGAGTCTAGACAGCGCTTTTCC 3'	7	
zgFGE-pBAD	5' CGTTTCACTCCATCCAAAAAAACGGG 3'	8	
zg4660-3MC-3A	5'AGACTAATGCA**CATATG**GCAGCAGCGCAAGACAAACCGAACATTGTTTTTATATTGTCC 3'	9	NdeI
zg4660-3MC-3A	5'TAGACTATTA**GGATCC**TTAGTGTTTTCTATATGAATCCATTCGTTTGGCC 3'	10	BamHI
zg4660-3MC-6A	5'AGACTAATGCA**CATATG**GCAGCAGCGGCAGCAGCGCAAGACAAACCGAACATTGTTTTTATATTGTCC 3'	11	NdeI
zg4660-3MC-6A	5'TAGACTATTA**GGATCC**TTAGTGTTTTCTATATGAATCCATTCGTTTGGCC 3'	12	BamHI

B

Gène cloné-mutation	Amorces	N°amorce
*zg*3515 C86S	5‘ GCGCCGGTGCACTTAACATGTTGGAATAGGCATTG 3'	13
*zg*3515 C86S	5‘ CAATGCCTATTCCAACATGTTAAGTGCACCGGCGC 3'	14
*zg*3515 C86A	5‘ CGCGCCGGTGCAGCTAACATGTTGGAATAGGCATTGG 3'	15
*zg*3515 C86A	5‘ CCAATGCCTATTCCAACATGTTAGCTGCACCGGCGCG 3’	16
*zg*3515 C86G	5’ CAATGCCTATTCCAACATGTTAGGTGCACCGGCGC 3’	17
*zg*3515 C86G	5’ GCGCCDGTGCACCTAACATGTTGGAATAGGCATTG 3’	18
*zg*3515 H186Q	5‘ CCATGCGCCCGAACCTGGTCGAGATAACC 3'	19
*zg*3515 H186Q	5‘ GGTTATCTCGACCAGGTTCGGGCGCATGG3'	20
*zg*3515 H186N	5‘ GGTTATCTCGACCAGGTTCGGGCGCATGG 3'	21
*zg*3515 H186N	5‘ CTATGGTTATCTCGACAATGTTCGGGCGCATGG 3'	22

Tableau 4 : Amorces utilisées pour les différentes stratégies de clonage de *Zg*4660 et *Zg*FGE (A) ou pour réaliser la mutagenèse dirigée des résidus Cys86 et His 186 de Zg3515 (B)

II.1. E Clonage

Pour toutes les cibles (pNEA-HEXIM-*zg*4660, pBAD-*zg*FGE, pNEA-*zg*4660$_{full}$), hormis celles clonées par R. Larocque (voir **Tableau 4**), le clonage a été réalisé à l'aide du kit In-Fusion® HD Cloning Plus (Clontech) dont le principe est résumé dans la **Figure 32**. L'utilisation de ce kit présente de multiples avantages. Le premier est qu'il est extrêmement rapide et permet de réaliser un clonage en moins d'une journée. De plus, l'utilisation de la polymérase CloneAmp HiFi incluse dans le kit permet d'avoir des séquences avec un taux d'erreur de 12 bases pour 542,580, y compris lors de l'amplification d'un vecteur entier qui peut se faire en une seule étape. Enfin, en une seule réaction, il est possible d'inclure plusieurs fragments dans un vecteur.

Amplification des gènes cibles par PCR

Toutes les amplifications ont été réalisées avec l'enzyme CloneAmp HIFI de chez Takara. Cette enzyme est extrêmement processive et rapide ce qui permet de réduire significativement les temps d'amplification Les temps et températures utilisés étaient les suivants :

Etapes	Temps	Température
Dénaturation	30 sec	98°C
Elongation (30 cycles)	10 sec	98°C
	15 sec	55°C
	Selon la taille du fragment à amplifier (5 sec/kpb à amplifier)	72°C
Fin	5 min	72°C

Nous avons linéarisé les vecteurs grâce à des enzymes de restriction. Cependant, dans le cas du plasmide *Zg*FGE-pBAD, cette dernière stratégie s'est révélée inefficiente. Le plasmide a donc été amplifié par PCR sous forme linéaire avec les amorces 7 et 8, puis l'insert a été intégré selon le protocole InFusion classique. Ainsi, pour le clonage de

*Zg*FGE dans le plasmide pBAD aucune étape de restriction n'a été nécessaire. Toutes les amplifications ont produit des inserts purs sur gel et c'est la raison pour laquelle la purification de ces amplicons sur colonne n'a été nécessaire.

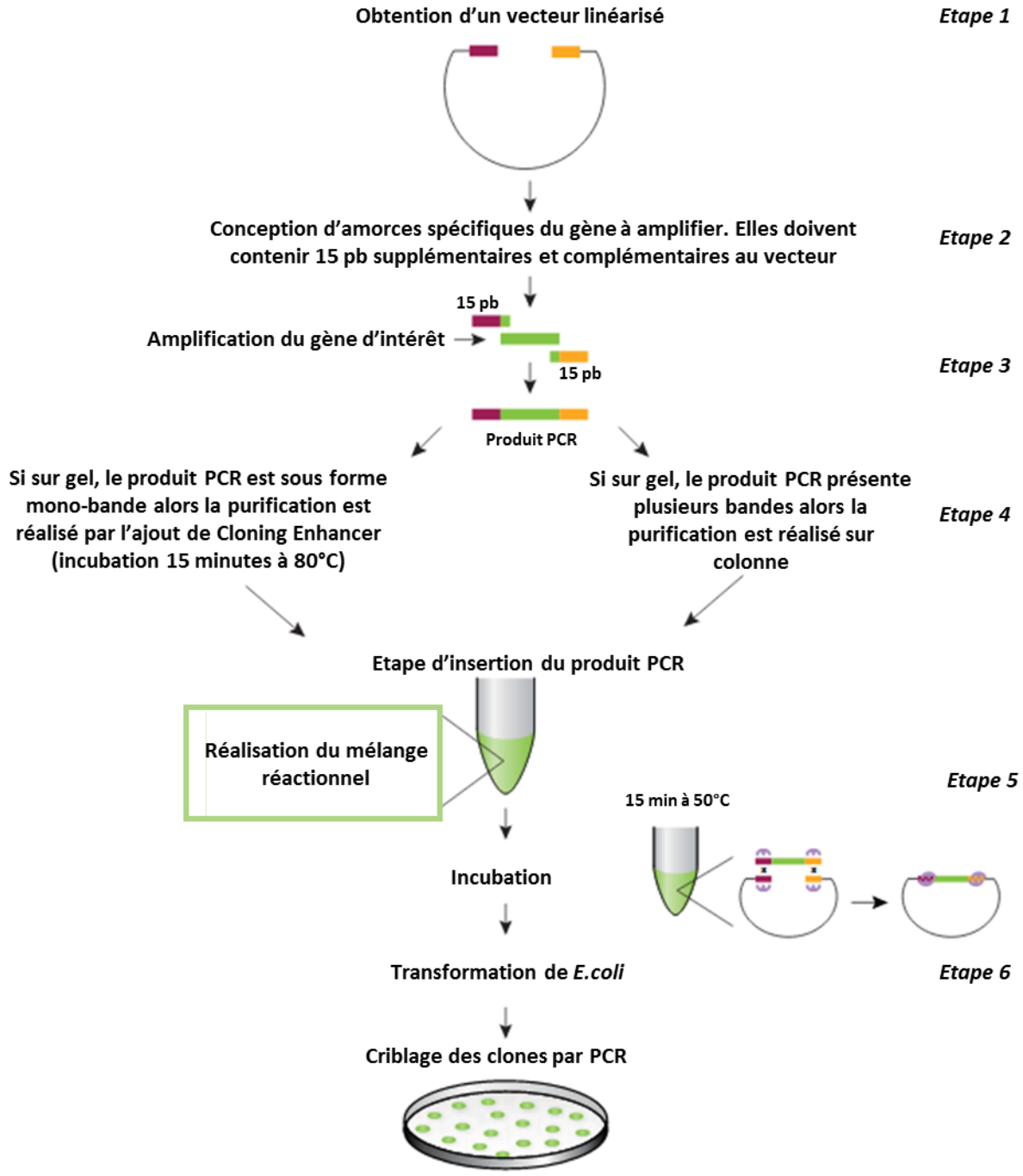


Figure 32 : Protocole de clonage selon le kit InFusion HD Cloning.

II.1.F Transformation des souches d'*E. coli*

Un microlitre des plasmides obtenus (10 ng/µl) ont été ajoutés à 50 µL de bactéries chimiquement compétentes BL21(DE3) (Takara). Le mélange a été incubé sur glace pendant 30 minutes avant de réaliser un choc thermique à 42°C pendant 45 secondes. Après une courte incubation sur glace, les bactéries sont régénérées dans 500 µl de LB pendant une heure.

II.2 Mutagénèse dirigée

Afin d'étudier l'importance de la cystéine 86 et de l'histidine 186 de *Zg*3515, nous avons réalisé des mutations ponctuelles à l'aide du kit Quick Change II XL (Stratagene), des doubles mutant C86G et H186N et Q ont également été réalisés. La cystéine 86 a été remplacée par une sérine, une alanine ou une glycine et l'histidine 186 a été mutée en glutamine et asparagine. Les amorces utilisées pour réaliser ces mutations sont récapitulées dans le **Tableau 3B**. Des réactions de séquençages ont été réalisées afin de vérifier que les mutations avaient été correctement effectuées. Les plasmides ont ensuite été utilisés pour transformer la souche de stockage *E. coli* DH5α et la souche d'expression *E. coli* BL21 (DE3), comme décrit précédemment.

II.3 Expression des différentes protéines cibles

Les souches BL21 (DE3) contenant les plasmides pFO4-z*g*4660**,** pNEA-HEXIM-3A-*zg*4660, pNEA-HEXIM-6A-*zg*4660, pFO4-*zg*3516 ont été utilisées pour inoculer 30 ml de milieu Luria-Bertani additionné de 100 µg/mL d'ampicilline. L'incubation a été réalisée sur la nuit à 37°C Cette préculture a été utilisée pour ensemencer du milieu autoinductible ZYP 5052 ((Studier, 2005), contenant 100 µg/mL d'ampicilline. Le volume de préculture utilisé était tel que la densité optique à 600 nm (DO_{600}) était de 0,1. La culture est ensuite incubée durant deux heures à 37°C sous agitation (180 rpm) avant d'être placée à 20°C pendant 14 à 16 heures (180 rpm).

Dans le cas de la sulfatase *Zg*3515, la préculture est utilisée pour ensemencer du milieu LB-ampicilline (100 µg/mL) dans des conditions similaires de DO_{600}. La culture ainsi obtenue est incubée à 37°C. Lorsque la DO_{600} une valeur entre 0,6 et 0,8,

l'expression de la protéine est induite par l'ajout de 0,1 mM d'IPTG et la culture est ensuite incubée à 20°C durant 14 à 16 heures.

*Zg*3515 sous forme séléniée a été produite à l'aide du kit : SelenoMethionine Medium Complete (Molecular Dimensions). Le milieu SelenoMet Medium Base a été additionné de 40 mg de sélénométhionine et ensemencé avec à l'aide d'une préculture de façon à avoir une DO_{600} de départ de 0,09 de bactéries comprenant le plasmide pFO4-*zg*3515. Le milieu a ensuite été placé à 37°C et lorsque la culture a atteint 0,8 DO_{600} l'expression de la protéine a été induite par l'ajout de 1mM d'IPTG, l'incubation a ensuite été poursuivie sur le nuit à 20°C.

En fin de culture, une centrifugation à 8000 *g* pendant 20 minutes à 4°C est réalisée afin de séparer les cellules du surnageant. Les culots de cultures sont conservés à – 20°C jusqu'au moment de leur utilisation.

II.4 Co-expression cytoplasmique de la *Zg*FGE et de *Zg*4660 chez *E.coli*

La transformation a été réalisée comme précédemment décrit dans le paragraphe XX. A la différence que les deux plasmides : pET-DUET-*Zg*FGE (10ng/µl) et pFO4-*Zg*FGE (10 ng/µl) ont été ajoutés ensemble aux bactéries chimiquement compétentes. Les bactéries ayant intégré les deux plasmides sont sélectionnés grâce à un étalement sur une boite LB-agar chloramphénicol (36 µg/mL) et ampicilline (100 µg/mL).

La préculture est réalisée avec trois clones différents dans du LB contenant les deux antibiotiques de sélection.

Pour la culture, le milieu LB-ampicilline (100 µg/mL) (sans chloramphénicol) est ensemencé de façon à obtenir 0,1 unité de DO_{600} finale. La culture est ensuite placée à 37°C jusqu'à ce que la DO_{600} atteigne 0,3 -0,4. A ce moment, de l''éthanol absolu à une concentration finale de 3% (v/v) est ajouté à la culture et celle-ci est placée à 4°C pendant 2 heures. A l'issue de cette incubation, la production de *Zg*FGE et de *Zg*4660 est induite par l'ajout de 0,1 mM d'IPTG et la culture est incubée sous agitation pendant seize heures à 16°C.

II.5 Co-expression périplasmique de *Zg*FGE et de *Zg*4660 chez *E.coli*

Un microlitre de pBAD-ZgFGE (10ng/µl) et pNEA-*Zg*4660$_{full}$ (10 ng/µl) sont ajoutés à 50 µL de bactéries chimiquement compétentes BL21(DE3) et la transformation est réalisée comme décrit dans le paragraphe II.1.6. Les bactéries ayant intégré les deux plasmides sont sélectionnés grâce à un étalement sur une boite LB-agar chloramphénicol (36 µg/mL) et ampicilline (100 µg/mL). Une fois que la culture a atteint 0,3 ou 0,6 de DO_{600}, l'arabinose est ajouté à une concentration finale de 0,1 ou 0,2 % (w/v). La culture est poursuivie à X°C. Lorsqu'elle atteint 0,6 ou 1,2 de DO_{600}, l'IPTG est ajouté à 0,1mM et l'incubation est poursuivie durant 16 heures à 20°C.

II.6 Purification des différentes protéines

II.6. A Purification de la sulfatase *Zg*4660.

Les culots bactériens sont lysés chimiquement. Le culot est resuspendu dans 30 mL du tampon de resuspension (50 mM Tris pH 8, 25% de sucrose (w/v)) et mis à incubé sous agitation à température ambiante avec ~10 mg de lysozyme. Les bactéries sont ensuite lysées par l'ajout de 60 mL de tampon de lyse (1% de deoxycholate (w/v), 1% de triton X-100 (w/v), 20 mM de Tris pH 8, 100mM de NaCl). Ce mélange est additionné de $MgCl_2$ de façon à avoir 5 mM finale. Pour éliminer l'ADN du milieu, 10 unités d'ADNase sont ajoutées et le mélange est agité doucement à température ambiante pendant 10 minutes. Une centrifugation à 40000 g durant 20 minutes permet d'éliminer les débris cellulaires.

Tous les protocoles de purification ont en commun la première étape : une chromatographie d'affinité sur une colonne d'affinité de type HisTrap (GE Healthcare, 5 ml). Dans le cas de *Zg*4660, il s'agit de deux colonnes HisTrap FF de 5 mL qui ont été mise en série. La colonne est équilibrée à un débit de 2 ml/ min dans le tampon A (Tris 50 mM pH8, NaCl 400 mM, Imidazole 20mM) avant injection (1mL/min) du surnageant de lyse. Après injections, les contaminants non fixées à la colonne sont éliminés par un lavage extensif avec le tampon A jusqu'à ce que la DO à 280nm soit négligeable. Les protéines sont éluées à 2 mL/min *via* un gradient d'imidazole réalisé avec le tampon B

Tris 50mM, NaCl 400mM et Imidazole 500mM) durant 5 volumes de colonne. Des fractions de 1.5 ml sont récoltées et conservées à 4°C.

Pour les protéines exprimées par les constructions pnEA-HEXIM, deux colonnes (HisTrap crude 5mL et HisTrap FF 5mL) sont montées en série et équilibrées avec le tampon C (Tris 50 mM pH8, NaCl 400 mM, Imidazole 5 mM). L'injection se fait à 0,5 mL/min et l'élution des contaminants dans les mêmes conditions. La majorité de la protéine est éluée lors du lavage. L'élution de la protéine fixée est réalisée par palier : le premier avec 25% de tampon B (125 mM imidazole) afin d'éliminer les contaminants. Le second à 60% de tampon B (300 mM imidazole) permet d'éluer la protéine. Des fractions de 1,5 ml sont récupérées aux différentes étapes.

Pour éliminer les contaminants de la protéine élué lors du lavage de l'étape d'affinité au nickel, une colonne échangeuse d'anion est réalisé. Il s'agit d'une QTrap 5mL de chez GE préquilibré avec le tampon suivant : Tris 20 mM pH8, NaCl 100 mM. Avant injection la protéine est préalablement diluée dans le tampon d'équilibration de la colonne pour diminuer les concentrations en sels. Les protéines non fixées sur la colonne sont éliminées par le passage de cinq volumes colonnes de tampon d'équilibration. L'élution est réalisée par un gradient de cins volumes colonnes de Tris 20 mM pH8, NaCl 1M, les fractions collectées sont de 1.5 mL.

L'analyse des différentes fractions d'élution par électrophorèse sur gel de polyacrylamide permet de repérer les fractions contenant nos protéines d'intérêt. Les protéines destinées aux tests enzymatiques sont par la suite dessalées grâce à une colonne de dessalage (HiPrep Desalting 26/10(GE Healthcare) préalablement équilibrée dans le tampon D (Tris 50mM pH8, NaCl 400mM).

Toutes les constructions destinées à la cristallogenèse (pFO4-*Zg*4660, pnEA-HEXIM-3A-*Zg*4660 et pnEA-HEXIM-6A-*Zg*4660) subissent une étape supplémentaire de purification. Une chromatographie d'exclusion de taille de type Sephadex G200 (GE Healthcare) (180cm et 3 cm) est réalisée avec le tampon D pour pFO4-*Zg*4660, ou avec le tampon E (Tris 50mM pH8, NaCl 150 mM) pour pNEA-HEXIM-3A-*Zg*4660 et pnEA-HEXIM-6A-*Zg*4660.

II.6. B Purification de *Zg*3515 et *Zg*3516

Le lysat bactérien est chargé sur une colonne d'affinité HiTrap FF (5mL GE Healthcare) préalablement équilibrée avec le tampon Tris 50mM pH8, NaCl 200 mM, $CaCl_2$ 1mM et Imidazole 20mM pour la purification de la sulfatase *Zg*3515. Dans le cas de de *Zg*3516, le tampon utilisé est : Bicine 50mM pH8NaCl 400mM et imidazole 20mM. L'élution est réalisée avec un gradient d'imidazole de 20 à 500 mM en 5 volumes de colonne

La purification de ces protéines est ensuite complétée par une chromatographie d'exclusion de taille sur une colonne de type HiLoad Superdex S200 16/60 (GE Healthcare). Dans le cas de *Zg*3515, le tampon utilisé alors est Tris 50mM pH8, NaCl 200 mM, $CaCL_2$ 1mM, tandis que pour *Zg*3516 le tampon est Bicine 50mM pH8, NaCl 300mM. Des fractions de 1,5 ml sont récoltées et analysées par électrophorèse sur gel de polyacrylamide (SDS- PAGE)

II.6. C Etude de la thermostabilité des protéines.

La thermostabilité des protéines est étudiée par la technique de diffusion dynamique de la lumière (DLS, *Dynamic Light Scattering*) à l'aide de Zetasizer Nano instrument (Malvern). Pour chaque mesure 75 µL de protéine à une concentration de 1mg/mL sont préalablement filtrés sur une membrane de 0,2 µm. L'échantillon ainsi préparé est déposé dans une cuvette spéciale et chauffé dans une gamme de températures de 20°C à 70°C par paliers de 2 ou de 5°C selon la protéine. La température de dénaturation correspond à la température pour laquelle le rayon hydrodynamique (Rg) augmente de façon significative.

II.7 Cinétiques des sulfatases sur substrats artificiels

La méthylumbelliferyl-sulfate (MUS, Sigma), un substrat synthétique, a été utilisée pour mesurer l'activité des sufatases. En effet, sous l'action de la sulfatase, la liaison ester-sulfate est clivée libérant ainsi la 4-méthylumbelliferone (MUF) qui est un composé fluorescent.

Pour pouvoir corréler la fluorescence à une quantité de liaison clivée, nous utilisons de la méthylumbeliférone. La gamme étalon en µM est de 0 ; 0,001 ; 0,01 ; 0,05 ; 0,1 ;

0,5 et 1. Les solutions stock de MUF (20 µM) et de MUS (85 µM) sont préparées dans du formamide.

Détermination des constantes cinétiques

La gamme de substrat (MUS) varie en fonction de l'enzyme. Concernant Zg3515, les tests ont été réalisés avec 4nM d'enzyme finale pour une gamme de MUS entre 0 et 3,4 mM de MUS, pour Zg4660 la concentration d'enzyme varie entre 0,8 nM et 100 µM avec une gamme de substrat comprise entre 0 et 272 µM. Les tests sont réalisés dans des plaques Greiner Half Black. Les puits sont pré-remplis avec les différentes solutions de MUS (60 mL). Au dernier moment 40 ml d'enzyme sont ajoutés et le suivi de la réaction est réalisé à 24°C. La lecture s'effectue à l'aide d'un lecteur de microplaques Spark Tecan 3000 avec les paramètres suivants : la longueur d'onde d'excitation de la MUF est de 360 nm et sa longueur d'onde d'émission est 449 nm, le gain de l'appareil est fixé à 60, la Z-position à 19902 µM. Tous les tests sont réalisés en triplicat.

Détermination du pH optimum

Pour déterminer le pH optimum de *Zg*3515, nous avons utilisé un tampon tripartite (1M d'acide acétique, 0,5 M de Bis Tris et 0,5 M de Triéthanolamine) dilué au dixième et ajusté à différents pH compris entre : 3,4 et 9,9. La réaction a été réalisée avec 1nM d'enzyme purifié et 409 µM d'enzyme, le substrat et l'enzyme sont préalablement dilués dans le tampon au pH désiré. Un aliquot de 100 µL a été prélevés toutes les 2 minutes pendant 10 minutes et dilué dans 900µL de tampon glycine 150 mM pH 10,3 et 150 mM NaCl. Le mélange ainsi obtenu est incubé durant 2 minutes à 60°C. 100 µL sont ensuite transférés dans des plaques Half Black Flat bottom (Greiner) ainsi que la gamme étalon, la même que celle utilisée précédemment, mais avec la MUF a dilué dans le tampon glycine décrit plus haut.

Influence des métaux sur l'activité enzymatique

Pour les tests de chélation des métaux, *Zg*3515 a été dilué (de façon à avoir 10,5 µM d'enzyme finale) dans le tampon Tris 50 mM pH8, NaCl 200mM. Le tampon pouvait contenir soit de l'EGTA soit de l'EDTA à 10 ou 100mM. L'enzyme a été incubé sur la nuit avec ces chélatants. Après quoi, l'activité de l'enzyme a été testée sur 15 µg/mL de MUS.

II.8 Gels de polyacrylamide FACE ou « *Fluorophore Assited Carbohydrate electrophoresis* »

Les oligosaccharides sont couplés à des fluorophores, chargés ou neutre, et mis à migrer sur des gels en polyacrylamide très concentré (31% d'acrylamide pour le gel de séparation) pour permettre la séparation en fonction de l'encombrement stérique de ces petits molécules. De plus, seules les molécules chargés négativement vont migrer pour ces gels si le fluorophore de couplage est neutre. Ainsi, dans le cas de molécules portant plusieurs charges, la séparation se fera en fonction de l'encombrement et de la charge.

Le gel de séparation est composé de 31% d'acrylamide (w /v), 0,33 M Tris pH 8,8 et 41 mM de glycine et le gel de concentration est à 6,8% d'acrylamide (w /v), 130mM Tris pH6,8 et 49 mM de glycine.

Selon la nature des oligosaccharides étudiés, deux fluorophores ont été utilisés : le disodium 8-aminonaphtalene-1, 3, 6-trisulfonate ou ANTS, qui est un fluorophore chargénégativement, utilisé pour marquer des oligosaccharides neutres. Le second fluorophore est le 2-aminoacridone, ou AMAC, ce dernier est neutre et utilisé pour marquer les oligosaccharides chargés.

Toutes les étapes de préparation, de réaction et d'électrophorèse sont réalisées à l'abri de la lumière.

L'ANTS est préparé à une concentration de 150 mM dans une solution acide acétique : eau (3 /17) (v /v). L'AMAC est à une concentration de 100 mM dans une solution acide acétique /DMSO (3 /17) (v /v) Les oligosaccharides déshydratés sont mis en présence de 2µL de fluorophore et de 2µL de sodium cyanoborohydride (1M dans du DMSO). Le mélange ainsi obtenu est incubé durant une nuit à 37°C avant d'être additionné de 20 µL de glycérol 25% (v /v).

5µL du mélange réactionnel sont déposés sur gel de polyacrylamide. La migration est réalisée durant environ 2 heures à 175 V (voltage constant) et à 4°C dans du tampon Tris 25mM pH 8,5 et glycine 0,192M préalablement refroidi. La visualisation des

oligosaccharides est réalisée sous UV à l'aide d'un UV Transiluminator (Thermofisher Scientific Bioblock).

II.9 Criblage des substrats naturels des sulfatases (*Zg*4660 et *Zg*3515)

Pour étudier la spécificité des sulfatases *Zg*3515 et *Zg*4660, nous avons utilisé les oligosaccharides produits lors de l'étude de la spécificité de *Zg*AgaC. Ces derniers ont été générés à partir de différentes algues agarophytes digérées par des GH16 (*Zg*AgaB, *Zg*PorB et *Zg*AgaC). Ces mélanges d'oligosaccharides sont issus de la digestion réalisée durant 24 heures à 35°C de 0,1 g/mL d'algues broyées avec 1µM d'enzyme (voir le **chapitre III**).

Les sulfatases *Zg*3515 et *Zg*4660 (0,5 µM ou 5 µM d'enzyme purifié respectivement) sont incubées durant 24-heures à 37°C avec 100 µL du mélange d'oligosaccharides décrit plus haut. Les échantillons (150 µ L) sont ensuite séchés puis marqués avec l'AMAC comme décrit dans le § II.8, avant d'être analysés par la technique de FACE (§ II.8).

II. 10 Criblage de l'activité enzymatique de la GH30 (*Zg*3516) sur substrats artificiels et naturels

Afin d'élucider la spécificité de l'enzyme, plusieurs substrats artificiels ont été testés et notamment le *p*-nitrophényl-β-D-xylobioside (1mg/mL), le *p*-nitrophényl-β-L-fucopyranoside (3,6mg/mL), *p*-nitrophényl-β-D-glucopyranoside (2mg/mL) et 4-méthylumbelliferyl-β-D-glucopyranoside (2mg/mL). L'hydrolyse des substrats p-nitrophényl libère du *p*-nitrophénol qui est jaune et dont l'absorbance à 420 nm est proportionnelle à la concentration. Pour le 4-méthylumbelliferyl-β-D-glucopyranoside, c'est la mesure de la fluorescence qui permet de déterminer s'il y a une activité enzymatique. Ces substrats ont été préparés dans du tampon Tris-HCl 50mM pH 8 contenant du NaCl 400mM. La réaction enzymatique est réalisée en incubant 50 µL de substrat avec 10 µL d'enzyme (1,66 µM) pendant une nuit à 30°C. La mesure d'activité a été réalisée avec un lecteur de microplaques Spark Tecan 3000.

Pour étudier l'activité de Zg3516, nous avons utilisé plusieurs polysaccharides préparés et ou mis à disposition par M. Jam: porphyrane de *Porphyra sp.*, des fucoidanes

de diverses origine (fs28 ; *Ascophyllum nodosum* ; *Laminaria hyperborea*), différents carraghénanes (κ, ι et λ) et du xylan de *Palmaria palmata*. Ces polysaccharides ont été dissouts dans du tampon Tris- HCl 50mM pH8, NaCl 400mM à une concentration finale de 0,5% (w/v). Afin d'éliminer les oligosaccharides naturellement présents une ultrafiltration avec un filtre Amicon Ultra-50 10kDa (Millipore) a été réalisée. 100 µL des polysaccharides ainsi préparés ont été incubés durant une nuit à 30°C avec 5 µL d'enzyme à 1,66 µM.

Un test a également été réalisé sur du *Palmaria palamata* frais qui a été préalablement broyé puis resuspendu dans le même tampon que les polysaccharides de façon à avoir un mélange à 6 % (w/v). Un premier volume d'enzyme (10 µL à 1,66 mM) a été ajouté au début de l'incubation, réalisée à 30°C sous agitation, puis après 5 heures, 5 µL d'enzyme ont été rajoutés et l'incubation a été poursuivie sur la nuit. Le milieu a ensuite été fractionné par centrifugation, puis filtré à l'aide d'un Amicon Ultra-50 10kDa (Millipore) Les différentes fractions (fraction totale, surnageant de centrifugation et ultrafiltrat) ont alors été analysées par FACE.

Des algues entières cryo-broyées ont également été utilisée pour tester l'activité de l'enzyme. Dans ce contexte, nous avons notamment utilisés *Osmundea pinnatifida*, *Dumontia contorta*, *Polysiphonia simulans*, *P. elongata*, *P. brodiei*, *Rhodomella sp.*, *Chondria dasyphylla*, *Cryptopleura ramosa*, *Gracillaria sp.*, *Vertebrata lanosa*, *Ceramium rubrum*, *Chylocladia verticillata*, *Dumontia contorta, Ulva sp.*, *Fucus spiralis*, *F. vesiculosus*, *F. serratus*, *Delseisseria sanguinea*, *Saccarina latissima* et *Condrus crispus*. Dans tous les cas, les cryo-broyats ont été repris dans le tampon Bicine 50mM pH8, NaCl 400mM. La réaction enzymatique est réalisée en incubant 900 µl de cryo-broyat avec 100 µL d'enzyme à 10 µM, ce qui a permis d'avoir dans le volume réactionnel des algues cryo-broyées à une concentration de 1 % (w /v) et une concentration enzymatique de 1 µM. Les mélanges sont incubés à 35°C pendant une nuit et sous agitation.

Des tests enzymatiques ont également été réalisés sur des oligosaccharides purifiés par F. Lesourd. Dans ce cas, les réactions enzymatiques étaient réalisées de la manière suivante : 120 µg d'oligosaccharides (20 µg/µL dans de l'eau MilliQ) sont incubés à 30°C avec 1 µM d'enzyme pendant 60 heures.

Toutes ces fractions résultant de l'hydrolyse enzymatique des poly- ou oligo-saccharides ont été marquées à l'ANTS et analysées par la technique du FACE décrite précédemment (§ II.8).

II. 11 Tests de *Zg*3515 et *Zg*3516 sur la fraction oligoporphyrane F50

La fraction F50 est une fraction d'oligoporphyranes générés à partir du porphyrane par *Zg*PorB. Cette fraction a été obtenue à la suite d'une purification sur chromatographie d'exclusion de taille de l'hydrolysat totale du porphyrane par *Zg*PorB. Cette fraction a été identifiée comme présentant des oligosaccharides substrat de la sulfatase Zg3515.

Une partie de la fraction F50 a été utilisée comme substrat pour étudier l'ordre d'action des enzymes Zg3515 et Zg3516 et voir si l'action d'une enzyme sur ce substrat permettait de générer le substrat de la seconde enzyme. Concrètement nous avons procédé de la manière suivante :

Dans le premier test, les oligosaccharides sont incubés pendant 24 heures avec 5 µM de sulfatase (*Zg*3515). L'enzyme est alors inactivée durant 5 minutes à 95°C puis le mélange réactionnel est incubé durant 24 heures à 30°C avec 5 µM de glycoside hydrolase (*Zg*3516). Le mélange réactionnel est séché et marqué à l'ANTS comme décrit précédemment (§ II.8).

Le second test est réalisée dans des conditions expérimentales (T°, temps) similaires mais la première incubation est réalisée avec la glycoside hydrolase (*Zg*3516) et la seconde avec la sulfatase (*Zg*3515)

II. 12 Chromatographie sur couche mince (CCM)

La technique de chromatographie sur couche mince (CCM) a été utilisée pour étudier la spécificité de Zg3516 sur des algues cryobroyées (les mêmes que dans le § II.10). Concrètement, les cryo-broyats ont été repris dans le tampon Bicine 50mM pH8, NaCl 400mM. La réaction enzymatique est réalisée en incubant 900 µl de cryo-broyat avec 100 µL d'enzyme à 10 µM, ce qui a permis d'avoir dans le volume réactionnel des algues cryo-broyées à une concentration de 1 % (w /v) et une concentration enzymatique de 1 µM. Les mélanges sont incubés à 35°C pendant une nuit et sous agitation.

80 µL du mélange réactionnel sont ensuite prélevés et séchés pour être repris dans 5 µL d'eau MilliQ. 4 µL sont déposés sur des plaques d'aluminium recouvertes d'une matrice de silice (TLC Silica gel 60 Millipore). La migration est effectuée à température ambiante avec le mélange éthanol : butanol : eau (5 :5 :3) dans une enceinte close. Après

séchage, la révélation est réalisée par vaporisation de naphtorésorcinol 0,1%, dilué dans un mélange 10% d'acide sulfurique (v /v) dans de l'éthanol. Les plaques sont ensuite séchées à 95°C sur une plaque chauffante. Une solution de xylose 2% (w/v) sert de référence de témoin de migration pour les oses simples.

II.13 Extraction et digestion des polysaccharides

II. 13. A Extraction

Afin d'étudier la spécificité de nos enzymes sur substrats naturels, nous avons récolté différentes algues rouges et notamment *Porphyra sp*. *Chondria dasyphylla, Polysiphonia elongata* et *Osmundea pinnatifida.*

D'une manière générale, les algues sont prétraitées pour obtenir un extrait insoluble dans l'alcool ou AIR (*Alcool insoluble residues*) selon le protocole décrit par Torode et al., 2015. Les algues sont ensuite resuspendues dans de l'eau MilliQ pour obtenir une solution à 1% (w/v). Selon l'algue, le traitement des AIR est différent.

- Dans le cas de *Porphyra sp*., l'extrait AIR est autoclavé pendant 1 heure à 121°C sous une pression de 1 bar afin d'extraire les polysaccharides.
- Les polysaccharides de *Chondria dasyphylla* sont extraits en chauffant les AIR resuspendus une heure à 95°C et à pression atmosphérique. Cette étape est réalisée deux fois sur le même AIR.
- Deux techniques sont utilisées pour l'extraction des polysaccharides de *Polysiphonia elongata*.
  - Soit le mélange est chauffé à 70°C pendant une heure, puis le surnageant et l'AIR sont séparés par filtration. Cette étape est répétée trois fois sur l'AIR.
  - Soit le mélange est porté à 120°C, pendant 5 minutes et sous une pression de 1 bar. Une filtration permet alors de récupérer le surnageant, dans ce cas l'opération n'est pas répétée.

Les surnageant récupérés à l'issue de ces extractions sont concentrés par évaporation. Les polysaccharides sont précipités par ajout de 5 volumes d'éthanol à

96% (v/v) pour un volume de surnageant. Une filtration sur filtre nylon de 0,2 µm permet de récupérer les polysaccharides qui sont ensuite lyophilisés. Le polysaccharide est ensuite repris dans de l'eau MilliQ pour obtenir une solution à 1% (w/v).

II. 13. B Digestion enzymatique

Pour le polysaccharide de *Polysiphonia elongata* et *Chondria dasyphylla*, l'enzyme *Zg*AgaB (purifiée selon le protocole de Jam *et al.,* 2005) est ensuite ajoutée de façon à avoir une concentration finale d'enzyme à 1 µM pour 1% (w/v) de polysaccharides. Le mélange est ensuite incubé à 37 °C pendant 60 heures.

Le porphyrane est digéré par l'ajout de 1µM (concentration finale) de *Zg*PorB pour 1% (w/v) de polysaccharides. La digestion est réalisée durant 60 heures à 37°C.

Toutes les digestions sont ensuite ultrafiltré avec des Amicon Ultra-50 (Millipore) dont le seuil de coupure est de 3kDa. Seule la fraction contenant des oligosaccharides de taille inférieure à 3 kDa est conservée.

II.14 Tests des différentes chromatographies échangeuses d'anions.

Pour la mise au point de la purification nous avons testé plusieurs colonnes échangeuse d'anion de 1 mL fournies par GE Healthcare : DEAE Sepharose Fast Flow, HiTrap Capto DEAE, ANX Sepharose 4 Fast Flow, Q Sepharose XL et HiTrap Capto adhere.

Les colonnes étaient préalablement équilibrées avec de l'eau MilliQ fraiche. Les colonnes ont été chargées avec 1mL d'oligosaccharide de *P. elongata* digérés par *Zg*AgaB à une concentration d'environ 1% (w/v). Les espèces non retenues sont éliminées par 10 volumes colonnes d'eau MilliQ. Puis les oligosaccharides ont été élué avec un gradient de NaCl allant jusqu'à 1M en dix volumes colonnes, et des fractions de 1 mL ont été récoltées. Les différentes fractions ont ensuite été analysées par FACE.

II.15 Purification des oligosaccharides de *Polysiphonia elongata* et *Chondria dasyphylla* par chromatographie de type échangeuse d'anion

Pour réaliser les purifications des oligosaccharides, 5 colonnes de HiTrap CaptoAdhere (5mL, GE Healthcare) sont mises en série. Toute la purification doit impérativement se dérouler à 4°C, en effet à 20°C le carbonate d'ammonium se dégrade en CO2 et en ammoniaque. Ce qui entrainerait d'une part la diminution de la concentration de l'éluant et d'autre part cela produirait des bulles préjudiciables aux colonnes. Toute la purification est effectuée à un débit de 1 mL/min comme recommandé par le fournisseur.

Les colonnes sont équilibrées avec de l'eau MilliQ fraiche Les oligosaccharides (20 mL à une concentration d'environ 1% w/v) sont chargés sur les colonnes grâce à un système AKTA start (GE). Les oligosaccharides non chargés, et donc non fixé, sont élués de façon isocratique par 5 volumes de colonne (VC) d'eau milliQ.

L'élution des oligosaccharides chargés est réalisée grâce à une solution de carbonate d'ammonium 1M.

Concernant *P. elongata*, la première phase d'élution est un gradient qui permet d'atteindre 6% de carbonate d'ammonium en 3,5 VC avec des fractions récoltées de 15 mL. La seconde partie du gradient permet d'atteindre 15% de carbonate d'ammonium en 11 VC et les fractions récoltées sont de 5 mL. La dernière partie de l'élution consiste en 4 VC de carbonate d'ammonium à 100% pour décrocher les derniers oligosaccharides fixés.

L'élution des oligosaccharides de *Chondria dasyphylla* est réalisée avec un gradient de carbonate d'ammonium (1M) : 0 à 30% (v/v) en 5 VC, puis 30 à 100 % (v/v) en 3 VC. La dernière partie de l'élution consiste en un plateau où du carbonate d'ammonium 1M permet d'éluer les derniers oligosaccharides. Les fractions récoltées lors de l'élution ont un volume de 15mL.

Les fractions sont regroupées, lyophilisées et marqué à l'AMAC pour permettre l'identification des fractions d'intérêt.

II.16 Cristallisation

Différents cribles commerciaux ont été utilisés pour déterminer les premières conditions de cristallisation : PACT, JCSG+ (Quiagen) et Morpheus I et II (Molecular Dimensions). Les criblages ont été réalisés à l'aide d'un robot Tecan et dans des plaques Greiner 96 puits. Les différentes concentrations de protéines testées étaient de 28,5 mg /mL pour *Zg*3515, 21 mg/mL et 30mg/mL pour *Zg*3516 et 20 mg/mL pour Zg4660. Les réservoirs contenaient 100 µL de précipitant et les gouttes étaient déposées avec un rapport protéine/précipitant de 2/1. Tous les criblages, ainsi que les reproductions en boites Linbro ont été réalisé à 20 °C.

La production des cristaux a été effectuée selon la technique de la goutte suspendue en boite Linbro 24 puits, au sein des réservoirs contenant 500 µL de précipitant et des gouttes constituées de 2 µL de protéines et 1 µL de précipitant. Une exception a été faite dans le cas de Zg4660 dont la reproduction des cristaux a été réalisée en plaque Greiner 96 puits.

II.17 Collecte des données et traitement

Les cristaux ont été préalablement congelés, sous flux d'azote à 100K, pour limiter les dégâts dûs aux radiations et à la diffusion des radicaux libres induits par les rayons X. La présence d'un cryoprotectant (glycérol, PEG400, MPD…) lors de l'étape de congélation est essentielle pour éviter la formation de glace, le choix du cryoprotectant est empirique. Dans le cadre de ce travail, le cryoprotectant utilisé est du glycérol à une concentration de 14 % mais cela varie en fonction des cristaux. Dans le cas de *Zg*3516 la solution de cristallisation contenait déjà du glycérol, il n'a donc pas été nécessaire de rajouté un cryoprotectant. Dans le cas de *Zg*AgaC les données étaient généralement de meilleures qualités avec 10 % de glycérol. Les données ont été collectées à l'ESRF (ID29, ID30) ou à Soleil (Proxima 1, Proxima 2). Lors de la collecte, les cristaux sont placés sur un goniomètre et sous flux d'azote à 100K. Les cristaux sont ensuite irradiés (selon un temps d'exposition variant de 0,1 s à 0,5s) et avec un pas d'oscillation déterminé en fonction des caractéristiques du cristal.

Les spectres de diffraction ont été indexés et intégrés avec XDS, puis mis à l'échelle avec XSCALE (dans le cas des collectes effectuées à Soleil) (Kabsch, 2010;Otwinowski

and Minor, 1997). La qualité des jeux de données sont évaluées en fonction de différents facteurs de qualité : complétude, I/σI, résolution, redondance et le facteur R_{merge}.

II.18 Résolution du problème de phase et affinement

Cas de *Zg*3515 : phasage par diffusion anomale simple (SAD)

Les données ont été collectées sur Proxima 1 (Soleil), un jeu a été collecté à la longueur d'absorbance du sélénium (0,97 Å). Le phasage a été réalisé via l'interface graphique HKL2MAP qui utilise la suite SHELX (Sheldrick, 2010). Une fois les méthionines positionnées, le logiciel ARP/wARP a été utilisé pour construire un premier modèle.

Cas des autres protéines : phasage par remplacement moléculaire

Plusieurs programmes de la suite CCP4 ont été utilisés selon le cas Molrep (pour Zg3516) ou MoRDa (pour ZgAgaC)(Vagin and Lebedev, 2015; Vagin and Teplyakov, 1997). Une fois qu'un modèle a été placé par ces logiciels, nous avons utilisé soit ARP/wARP (pour améliorer le modèle dans le cas de *Zg*AgaC), soit Coot (Emsley and Cowtan, 2004; Perrakis et al., 2001) pour modifier et améliorer le modèle manuellement (cas de *Zg*3516.

Les modèles sont tous complétés et modifiés manuellement avec Coot avant d'être affinés plusieurs fois avec REFMAC (Winn et al., 2003).

II.19 Tests d'activation/inhibition de *Zg*3515 par différents oses.

Tous les tests ont été réalisés avec l'enzyme sauvage et le mutant C89S, pour l'étude de stéréospécificité le mutant H186N a été testé en plus.

Les différents mutants ont été testés à une concentration de 10 nM finale sur de la MUS à 200 µM en présence de différents ose simples à 50 mM finale (D-Galactose, D-Fructose, D-Glucose, D-Fucose, Mannose, Sucrose et Mélibiose).

Le même test a été réalisé mais avec 100 mM d'ose final (L-Galactose, D-Galactose, D-Fructose et L-Fucose).

Les plaques et les paramètres du lecteur de microplaques sont les mêmes que pour les études cinétiques.

II.20 Essai de *trans*-sulfatation avec *Zg*3515

Des mutants de la sulfatase ont été utilisés pour vérifier si il était possible qu'ils catalysent le transfert d'un sulfate depuis un donneur de sulfate à du L-galactose.

Le test a été réalisé avec le mutant C89G/H186Q. L'enzyme était à une concentration de 1 µM soit dans le tampon succinate 100mM pH6, soit dans le tampon Tris 100 mM pH8. Différentes concentrations d'acide sulfamique ont été testées (40mM, 80mM et 200mM) contre une seule concentration de L-galactose (40mM). Le L-galactose, l'acide sulfamique et l'enzyme ont été mis à incuber ensemble à 20°C pendant 20 heures.

L'activité trans-sulfatation est détectée par chromatographie sur couche mince. 4 mL de mélange réactionnel sont déposés sur des plaques TLC Silica gel 60 Millipore. La migration est réalisée à température ambiante avec un éluant composé d'un mélange butanol : éthanol : eau (5 :5 :3) et la révélation avec le mélange résorcinol (cf. II.13).

Chapitre III : Caractérisation d'une nouvelle sous-famille de GH16

The agar-specific hydrolase *Zg*AgaC from *Zobellia galactanivorans* defines a new GH16 subfamily

Anaïs Naretto[1], Mathieu Fanuel[2], David Ropartz[2], Hélène Rogniaux[2], Mirjam Czjzek[1], Charles Teillier[3*] and Gurvan Michel[1*]

Dans cet article j'ai réalisé les expériences suivantes :

- Le clonage de l'enzyme dans le plasmide pGex, ainsi que la mutagénèse dirigée.
- La surexpression et la mise au point des conditions de purification de l'enzyme sauvage et du mutant sans étiquette histidine.
- La caractérisation enzymatique de *Zg*AgaC, ce qui inclut la comparaison des profils de dégradation des différentes GH16, la détermination des conditions de pH et de températures optimales ainsi que la détermination des paramètres cinétiques de l'enzyme.
- La cristallogenèse et la résolution de la structure de *Zg*AgaC.
- La mise au point de la purification (par chromatographie d'exclusion de taille et par échange d'anion) des oligosaccharides, ainsi que leur accumulation.

Résumé :

Zobellia galactanivorans est une bactérie marine du phylum des Bacteroidetes capable de dégrader de nombreux polysaccharides d'origine marine, dont les agars (Barbeyron et al., 2016) qui sont extraits à partir de la paroi cellulaire des algues rouges. Ce polysaccharide est composé d'un motif répété comme suit : 3-O-β-D-galactopyranose-1,4-α-L-galactopyranose. Dans son milieu naturel, ce motif présente souvent de nombreuses modifications chimiques (sulfatation, méthylation…) masquant le motif de répétition.

Le génome de *Z. galactanivorans* contient seize glycosides hydrolases de la famille 16 (GH16) et neuf d'entre elles sont annotées comme étant soit des β-agarases, soit des β-porphyranases. Ainsi, nous retrouvons *Zg*AgaA, *Zg*AgaB, *Zg*AgaC, *Zg*AgaD dans le clade des β-agarases et *Zg*PorA, *Zg*PorB, *Zg*PorC, *Zg*PorD, *Zg*PorE dans le clade des β-

porphyranases (Hehemann et al, 2010). Bien que *Zg*AgaC ait été classifiée dans le clade des β-agarases, sa séquence protéique est très divergente de celles des autres β-agarases. Au moment de sa classification, le manque de séquences disponibles et l'absence de caractérisation biochimique de *Zg*AgaC ne permettaient pas une classification plus détaillée. De plus, des études préliminaires avait démontré que *Zg*AgaC avait un comportement non classique pour une β-agarase (Jam et al., 2005). *Zg*AgaC se trouve dans le milieu extracellulaire de *Z. galactanivorans* et est capable de dégrader un gel d'agarose mais ne se fixe pas sur un gel de sépharose (bille d'agarose modifié) *a contrario* de ce qui a été observé pour les autres enzymes (Jam et al., 2005). Ces éléments suggèrent que *Zg*AgaC n'est pas une β-agarase classique. La phylogénie désormais actualisée des GH16 démontre que *Zg*AgaC, avec d'autres protéines disponibles aujourd'hui, forme bien un nouveau clade, distinct de celui des β-agarases.

En vertu de ces éléments, nous supposions que *Zg*AgaC avait une nouvelle spécificité de substrat qui lui conférait la capacité d'hydrolyser une forme d'agar modifiée.

La caractérisation biochimique de *Zg*AgaC présentait un défi dans le sens où cela impliquait de développer un protocole permettant de discriminer son action de celles des autres GH16 agissant sur les agars. Pour cela une nouvelle stratégie a été mise au point. Pour conserver les modifications naturelles des agars durant l'extraction nous avons choisi de faire agir les enzymes directement sur les algues broyées. Plusieurs algues, produisant des agars différents, ont été selectionnées. Elles ont été incubées avec les 3 enzymes suivantes *Zg*AgaC, une β-agarase (*Zg*AgaB) et une β-porphyranase (*Zg*PorB). Les produits d'hydrolyse ont été marqués et analysés par électrophorèse sur gel de polyacrylamide selon la technique du FACE (Fluorophore Assited Carbohydrate Enzymes). La comparaison des profils de dégradation de ces trois enzymes ont permis d'identifier plusieurs algues pour lesquelles *Zg*AgaC produisait des oligosaccharides très différents de ceux générés par les deux autres enzymes. C'est le cas par exemple de l'agar extrait de l'algue *Osmundea pinnatifida*. C'est dans ce contexte que j'ai entrepris la purification des oligosaccharides produits par *Zg*AgaC à partir de l'agar extrait de cette algue.

Les différentes fractions purifiées ont été analysées en spectrométrie de masse (Plateforme BIBS, INRA NANTES). Le Dr. D. Ropartz et M. Fanuel ont résolu la structure d'un des oligosaccharides. Ce dernier comprend de nombreuses modifications, parmi lesquelles le 3,6-anhydro-L-galactose-2-sulfate (LA2S), le 3,6-anhydro-L-galactose-2-

methyle et un D-galactose avec un branchement xylose. C'est la première fois qu'une GH16 est documentée comme permettant l'hydrolyse d'un agar comportant ces modifications.

La structure tridimensionnelle de l'enzyme a également révélé un site actif très divergent de celui des β-porphyranases et des β-agarases connues à ce jour. En effet, la comparaison entre *Zg*AgaC, *Zg*PorA et *Zg*AgaB ne révèle que cinq résidus rigoureusement conservés, dont deux sont les résidus catalytiques. Quatre de ces résidus conservés sont impliqués dans la reconnaissance et la stabilisation des galactoses aux sous-sites +1 et -1. Le cinquième résidu, un tryptophane, permet la reconnaissance et la stabilisation du substrat en position -3. En sus, trois résidus basiques sont présents dans le site actif de *Zg*AgaC. Ces trois résidus sont conservés dans la séquence de tous les membres de cette sous-famille. Ils permettent probablement la reconnaissance et la stabilisation de groupement sulfate, ceux en position 2 du résidu LA par exemple.

Concernant le comportement cinétique de *Zg*AgaC, l'une des découvertes les plus intéressantes est la mise en évidence de l'influence, sur le comportement de l'enzyme, du microenvironnement de pH local des polysaccharides. En effet, le pH optimum de *Zg*AgaC sur de l'agarose (qui est neutre) est de 6,6, tandis que sur un agar acide (tel que celui extrait d'*O. pinnatifida*) le pH optimum est de 9,1. Par la suite, les essais cinétiques ont été réalisés à chaque fois au pH optimum pour le polysaccharide testé. Avant de présenter les résultats de ces cinétiques une remise en contexte est nécessaire.

Les enzymes qui agissent sur des polysaccharides ont pour substrat des polymères, qui ne sont pas homogènes en taille et en poids, et présentent de nombreuses modifications. Pour réaliser des études enzymologiques sur ces composés, nous devons disposer de leur masse moléculaire, or au mieux nous ne disposons que des poids moléculaires moyens. Dès 1988, le travail de Rochas et Lahaye (REF) mettait en lumière ce problème. En fonction du fournisseur et du lot d'agarose testé, le poids moléculaire moyen du polysaccharide variait de 17 à 134 kDa. Ainsi, tous calculs enzymatiques menés sur de l'agarose présentent un biais important. Lors de l'étude de polysaccharides présentant de nombreuses modifications, une seconde difficulté s'ajoute : le nombre de motifs de reconnaissance de l'enzyme et leurs répartition au sein du polysaccharide est inconnu.

Lors de nos expérimentations, nous avons déterminé que les valeurs de K_M de *Zg*AgaC étaient de : 1.26 mM ± 1.73 e-1 lorsque l'agarose (molecular grade) était utilisé comme substrat, et de 9.87 mM ± 9.25 e-1 lorsque le substrat était de l'agar extrait à partir de *O. pinnatifida*. Le calcul du K_M s'effectue selon une méthode conventionnelle dans le milieu de la glycobiologie (Jam et al., 2005 ; Hehemann et al., 2010) qui consiste à convertir les concentrations des polysaccharides (exprimé en pourcentage) en mM en utilisant le poids moléculaire du néoagarobiose (3,6- anhydro-L-galactose—1,3-D-galactose) comme référence.

Ainsi, si on prend le cas de *Zg*AgaC avec l'agarose comme substrat, on obtient, si on considère les poids moléculaires de Rochas et Lahaye, des valeurs de K_M qui se trouvent entre 24,03 µM et 3,05 µM, soit 50 à 400 fois inférieures à celles estimées en utilisant la méthode « conventionnelle ». Si nous voulons pouvoir comparer, de façon pertinente, les différentes enzymes agissant sur l'agar, il nous faudrait idéalement avoir la masse moléculaire moyenne du polymère testé. De plus, pour pouvoir comparer les valeurs obtenus avec l'agarose et celles obtenus avec l'agar issu d'*O. pinnatifida*, il faudrait déterminer le nombre de sites de fixation de l'enzyme disponible sur ce dernier.

Les données et les moyens à notre disposition actuellement ne nous permettent donc que de déterminer approximativement les constantes cinétiques de nos enzymes. Toutes comparaison entre enzymes ou bien entre polysaccharides substrat devra par conséquent être très prudente.

Les éléments présentés dans l'article (la structure, la phylogénie, la caractérisation cinétiques) nous ont permis de définir les éléments spécifiques aux AgaC-*like* et de caractériser la première enzyme de cette sous-famille.

The agar-specific hydrolase *Zg*AgaC from *Zobellia galactanivorans* Dsij^T defines a new GH16 subfamily

Anaïs Naretto[1], Mathieu Fanuel[2], David Ropartz[2], Hélène Rogniaux[2], Robert Larocque[1], Mirjam Czjzek[1], Charles Teillier[3*] and Gurvan Michel[1*]

[1]Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France

[2]INRA, Unité de Recherche Biopolymères Interactions Assemblages (BIA), Nantes, France

[3]UFIP, UMR 6286 CNRS, Université de Nantes, 2, rue de la Houssinière, 44322, Nantes, France

Running title: Characterization and structure of *Zg*AgaC

*Co-corresponding authors to whom correspondence should be addressed: Gurvan MICHEL; gurvan@sb-roscoff.fr; Phone number: + 33 2 98 29 23 75; Charles Teillier; Charles.Tellier@univ-nantes.fr; Phone number: + 33 2.

Keywords: marine bacteria, red algae, agar, glycoside hydrolase, GH16, sulfated polysaccharide, pH microenvironment, mass spectrometry, crystal structure, evolution

ABSTRACT

Agars are sulfated galactans from red macroalgae composed of D-galactose (G unit) and L-galactose (L unit) alternatively linked by α-1,3 and β-1,4 glycosidic bonds. In their natural context, these polysaccharides display a high complexity, with numerous modifications of their backbone (e.g., presence of 3,6-anhydro-bridge (LA unit), sulfations, methylation). Currently bacterial polysaccharidases active on agars (β-agarases and β-porphyranases) have been essentially characterized on simple agarose and more rarely on porphyran, a polymer constituted by agarobiose (G-LA) and porphyranobiose (G-L6S) motifs. How bacteria can cope with complex agars remains thus an open question. Here, we studied an enzyme from *Zobellia galactanivorans* (*Zg*AgaC) which is distantly related to the GH16 family β-agarases and β-porphyranases. Using a large collection of red algal species, we demonstrated that *Zg*AgaC hydrolyzes not only agarose but also complex agars from *Ceramiales* species. Notably, we elucidated, by tandem mass spectrometry, the structure of a purified hexasaccharide product released by the action of *Zg*AgaC on the agar extracted from *Osmundea pinnatifida*: L6S-G-LA2Me-G(2Pentose)-LA2S-G. We also resolved the crystal structure of ZgAgaC at high resolution (1.3 Å). By comparison with the structures of ZgAgaB and ZgPorA, in complex with their respective substrates, we determined that *Zg*AgaC recognizes agarose in a different way from classical β-agarases and we identified conserved residues involved in the binding of complex oligoagars.

Furthermore, we demonstrate a probable influence of the acidic polysaccharide's pH microenvironment on enzyme activity. Finally, a phylogenetic analysis supports that ZgAgaC homologues define a new GH16 subfamily distinct from β-porphyranases and classical β-agarases.

The main cell wall polysaccharides of marine red macroalgae are unique sulfated galactans, carrageenans or agars (1). These polysaccharides consist of a backbone of galactopyranose units linked by alternating α-1,3 and β-1,4 linkages. While all the 3-linked residues of these galactans are in the D configuration (G unit), the 4-linked galactose units are in the D configuration in carrageenans (D unit) and in the L configuration in agars (L unit). A further layer of complexity is introduced by the systematic occurrence of either a 3,6-anhydro bridge or a sulfate group at C6 in the 4-linked galactose residues. Galactose-6-sulfate (referred to as D6S in carrageenans and L6S in agars) is considered as the biogenic precursor of 3,6-anhydro-galactose (referred to as DA in carrageenans and LA in agars). Indeed the conversion of galactose-6-sulfate into 3,6-anhydro-galactose is catalyzed by galactose-6-sulfurylases (2,3), enzymes which have been identified only in genomes of red macroalgae (4,5). The regular structure of the backbone of red algal galactans is often masked by additional chemical modifications, such as ester sulfate groups, methyl groups, or pyruvic acid acetal groups (6-8). This complexity has been taken into account in carrageenan nomenclature and traditionally carrageenans are identified by a Greek prefix, indicating the major component of the sample (7). This Greek prefix nomenclature is widely used in literature, but also in industry and even in legislation. However this system is not sufficient to describe more complex carrageenans and a letter code based nomenclature, inspired by the IUPAC nomenclature, was proposed by Knutsen and coworkers for systematically described these complex galactans (9). A Greek prefix nomenclature has never been introduced for agars, probably because academic and private sectors have essentially focused on agarose, a high-gelling agar essentially devoid of modifications which is used worldwide in food industries, in molecular biology and for chromatography matrices. However, natural agars are largely as complex as carrageenans (6,8,10-12) and Knusten's nomenclature is also used to describe these polymers (as we do here).

To investigate the structure of natural agars, the combination of biophysical methods and specific enzymes has become a powerful strategy. For instance the use of bacterial β-porphyranases with ^{1}H-NMR (13,14) demonstrated that the agar from *Porphyra umbilicalis*, commonly referred to as porphyran, is composed of one third of agarobiose motifs (G-LA) and of two third of porphyranobiose motifs (G-L6S). Tetrasacharides with a C6-methylated D-galactose (L6S-G6Me-L6S-G) were also identified. Recent reinvestigation of these olio-porphyrans using tandem mass spectrometry (MS/MS) have revealed an even greater complexity with different degrees of methylation, substitution by uronic acids and branching with a pentose unit (15,16). *Zobellia galactanivorans* Dsij^{T} is a model alga-associated bacterium (17) which has already provided several enzymatic tools to study agar structure. Its agarolytic system includes four potential β-agarases and five potential β-porphyranases (all belonging to the family 16 of glycoside hydrolases; GH16; http://www.cazy.org/; (18)) and at least two 1,3-α-3,6-anhydro-L-galactosidases (GH117 family) (19,20). Three β-agarases (*Zg*AgaA, *Zg*AgaB and *Zg*AgaD) and two β-porphyranases (*Zg*PorA and *Zg*PorB) have already been characterized at the biochemical and structural level (13,21-24). These studies notably revealed a gradient of tolerance for modifications in the agar chain, from *Zg*AgaD which is strictly specific for long stretches of unsubstituted agarobiose motifs to *Zg*AgaB and *Zg*PorB which can tolerate substituents at some subsites (24).

The last putative β-agarase, which has not been characterized in *Z. galactinovorans*, is *Zg*AgaC. The first mention in literature of this enzyme was the purification of the wild-type enzyme from the culture medium of *Z. galactanivorans* (21). The authors demonstrated that the extracellular agarolytic activity of this marine flavobacterium is encompassed by two enzymes: *Zg*AgaA, which was retained on a Sepharose CL6B affinity

chromatography column, and *Zg*AgaC, which was found in the flowthrough of the agarose-containing column. By Edman degradation, an oligopeptide of *Zg*AgaC was sequenced (ATYDFTGNTP) but the *agaC* gene was not successfully cloned (21). Later, the sequencing of the *Z. galactanivorans* genome revealed that the ATYDFTGNTP peptide was only found in the open reading frame ZGAL_4267, which was thus named *agaC* (17). Here, we present the characterization of *Zg*AgaC on agarose and on a natural complex agar, revealing a previously undescribed substrate specificity. The crystal structure of this enzyme gives insight into the molecular bases of its recognition of agars. Finally, an updated phylogeny of the GH16 galactanases indicate that *Zg*AgaC and its close homologues constitute a new subfamily within the GH16 family.

RESULTS

The agaC gene and its genomic context

The *agaC* gene encodes a protein of 328 amino acids with a theoretical molecular mass of 37.6 kDa. The protein sequence contains a lipoprotein signal peptide (1-17), with a lysine at the +2 position (Lys19) suggesting that *Zg*AgaC is anchored to the outer membrane (25), a propeptide rich in glutamate and lysine (Cys18-Asn67) and a catalytic module of the GH16 family (Ala68-Glu328). Interestingly, this protein does not feature a C-terminal type IX secretion system (T9SS) domain, which is unique to *Bacteroidetes* (26), whereas *Zg*AgaC was shown to be an extracellular enzyme (21). This suggests that *Zg*AgaC is indeed anchored to the outer membrane and likely oriented towards the external medium. Its secretion might be the result of a fortuitous proteolytic cleavage of the low complexity propeptide.

The position of *agaC* in the genome is noteworthy. Although this gene is not included in a polysaccharide utilization locus (PUL), it is nonetheless found in a carbohydrate-related genomic context. Indeed, *agaC* (locus ID: ZGAL_4267) is located next to the mannitol utilization operon (ZGAL_4259 - ZGAL_4264) (27), the iota carrageenase gene *cgiA1* (locus ID: ZGAL_4265) (28,29) and a gene predicted to encode a β-helix fold protein (locus ID: ZGAL_4268). In the recent characterization of the carrageenolytic regulon in *Z. galactanivorans* (30), the genes *cgiA1* and ZGAL_4268 were found to be strongly induced in the presence of carrageenans (κ- and ι-carrageenans for *cgiA1*; only κ-carrageenan for ZGAL_4268). But this was not the case of *agaC* despite its neighboring location (30). In another transcriptomic study on *Z. galactanivorans* (31), *agaC* was found to be induced by agar but not by porphyran. More surprisingly this gene was also induced in the presence of laminarin, the storage polysaccharide of brown algae.

Phylogenetic analysis of GH16 family galactanases

Currently the GH16 family includes three different types of enzymes specific for sulfated galactans from red algae: κ-carrageenases (32), β-agarases (21) and β-porphyranases (13). Based on pairwise sequence comparisons with characterized galactanases from *Z. galactanivorans*, the GH16 module of *Zg*AgaC appears highly divergent from these different enzymes. The closest homologue is *Zg*PorA (31% identity), distantly followed by *Zg*AgaA (23%), ZgAgaB (21%), *Zg*PorB (21%) and *Zg*CgkA (19%). To have a better idea on the relationships between these enzyme groups, we searched Genbank for homologues of these GH16 galactanases from *Z. galactanivorans*. We thus identified 154 sequences, essentially originating from diverse phyla of marine heterotrophic bacteria. Using the GH16 family laminarinases *Zg*LamA (33) and *Zg*LamB (34) as outgroups, we calculated a phylogenetic tree of these sequences (Fig. 1; supplementary Fig. S1). This tree is divided into 5 very solid clades, with bootstrap values ranging from 86% to 99%. As expected κ-carrageenases and β-agarases form two different clades. This analysis also indicates that *Zg*AgaC and its closest homologues (25 sequences) constitute a monophyletic group distinct from "classical" β-agarases. The most surprising result is that β-porphyranases form two distinct monophyletic groups, the *Zg*PorA-like group and the *Zg*PorB-like group. Some relative positions of these clades are also solid. The κ-carrageenase clade is the closest to the root and is a sister group of all the agar-specific enzymes (bootstrap value: 99%). The *Zg*PorA-like clade is the most ancestral group of agar-specific enzymes, solidly rooting a cluster

composed of the three other clades (bootstrap value: 97%). Within this cluster the relative positions between the "classical" β-agarases, the *Zg*AgaC-like clade and *Zg*PorB-like clade are unclear, since all the nodes connecting these clades have bootstrap values below 50%.

ZgAgaC is an agar-specific enzyme degrading highly modified agars

The recombinant *Zg*AgaC corresponds to the catalytic GH16 module encoded by *agaC* (Thr69-Glu328) without the lipoprotein signal peptide and the N-terminal propeptide. This protein (31.3 kDa, isoelectric point: 5.81) was produced in soluble form in *E. coli* BL21 (DE3) with a high yield of ~185 mg per liter of culture. A two-step purification, immobilized metal ion affinity and size exclusion chromatography (SEC), was necessary to purify *Zg*AgaC to electrophoretic homogeneity. The SEC analysis suggested that *Zg*AgaC is a monomer in solution, and this result was confirmed by dynamic light scattering (DLS).

To understand the specificity of *Zg*AgaC, we compared the degradation pattern of this GH16 enzyme with the already characterized β-agarase *Zg*AgaB and β-pophyranase *Zg*PorB (13,21,24) on diverse natural agars. Thirteen red agarophyte algae were tested with *Zg*AgaC, *Zg*PorB, *Zg*AgaB and the combination of *Zg*PorB and *Zg*AgaB. The oligosaccharide degradation patterns of these GH16 enzymes were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE). For most algae (e.g. *Vertebrata lanosa*, *Rhodomella sp*. *Polysiphonia simulans* and *Polysiphonia brodei)*, the degradation pattern of *Zg*AgaC was very close to that of *Zg*AgaB and was different from that of *Zg*PorB. This suggests that *Zg*AgaC essentially behaves as a classical β-agarase on the agars from these algal species. Nonetheless, the degradation pattern of *Zg*AgaC was unique for *Polysiphonia elongata* and *Osmundea pinnatifida* (formerly named *Laurencia pinnatifida*) (Fig. 2). On *P. elongata*, most of the bands of the *Zg*AgaC profile are common with the *Zg*AgaB profile; however, several bands of the *Zg*AgaB profile are missing in the ZgAgaC profile. Conversely, a band corresponding to a low molecular mass oligosaccharide is only released by *Zg*AgaC (Fig. 2A). The major differences between the degradation patterns were observed for *O. pinnatifida* (Fig. 2B). *Zg*PorB degraded the agar of this species, but did not produce a large amount of oligosaccharides. The degradation of this polysaccharide by *Zg*AgaB appears even less efficient, with only very few bands. The combination of *Zg*PorB and *Zg*AgaB resulted in a larger release of oligosaccharides, suggesting a synergistic effect on this substrate. In contrast, *Zg*AgaC was able alone to produce numerous oligosaccharides, with a profile which partially resembles that of *Zg*PorB but also features bands unique to ZgAgaC (Fig. 2B).

ZgAgaC enzymatic characterization

The enzymatic activity of *Zg*AgaC was further tested on three purified polysaccharides: agarose (Sigma), "pure porphyran" (a native porphyran pre-treated with *Zg*AgaB to remove the agarobiose motifs) and the native agar extracted from *O. pinnatifida*. Preliminary tests demonstrated that *Zg*AgaC is active on agarose and on the *O. pinnatifida* agar, but has no activity on "pure porphyran". To determine the mode of action of *Zg*AgaC, the enzymatic hydrolysis of agarose was monitored by FACE for 1 hour at 40°C (Fig. 3). After 1 min a large range of oligosaccharides was released by *Zg*AgaC, from oligosaccharides with high degrees of polymerization to smaller oligosaccharides. All along the kinetic experiment, the population of larger oligosaccharides decreased with a concomitant increase of the apparent quantity of small oligosaccharides. This evolution of the degradation patterns indicates that *Zg*AgaC proceeds with an endolytic mode of action.

The pH dependence of *Zg*AgaC was studied on both agarose and the *O. pinnatifida* agar. Due to precipitation problem of *Zg*AgaC with the universal buffer Teorell-Stenhagen (35), several buffers were used separately to measure the pH dependence. Although, this strategy generally results in gaps between the different curves this was not the case here, likely due to the use of 150 mM of NaCl in each buffer to maintain a constant ionic strength. Interestingly, the pH optimum of *Zg*AgaC is strongly dependent of the substrate used, pH 6.5 with agarose and pH 9 with the *O. pinnatifida* agar (Fig. 4). The temperature dependence of *Zg*AgaC was evaluated only with the *O. pinnatifida* agar as agarose gels

below 40°C while the *O. pinnatifida* agar remained soluble at all the tested temperatures. With the *O. pinnatifida* agar the activity of *Zg*AgaC was optimal at 50°C. The kinetic parameters of *Zg*AgaC have been thus evaluated at 50°C but at the pH optimum determined for each substrate. The kinetic curves obtained are typical of a Michaelian behavior. The kinetic parameters of *Zg*AgaC are k_{cat}= 2.39 e+2 s^{-1} ± 1.21e1, K_M= 1.26 mM ± 1.73 e-1 for agarose and k_{cat}= 1.34 e+2 s^{-1} ± 1.21e1, K_M= 9.87 mM ± 9.25 e-1 for the *O. pinnatifida* agar.

Purification and structure of the O. pinnatifida oligosaccharides

To determine the molecular basis of the recognition of complex agars by *Zg*AgaC, the characterization of agar oligosaccharides from *O. pinnatifida* produced by this enzyme was undertaken. The purification of agar oligosaccharides was realized by SEC as previously described (21). 48 fractions were collected and were analyzed by FACE gels. The three fractions (named OP30, OP36, and OP44) that contained the smallest oligosaccharides identified by FACE gels were characterized by electrospray - mass spectrometry (ESI-MS, Fig. 5). The major species in the OP44 fraction was measured at m/z 723.17 as $[M-H]^-$ (Fig. 5A). This oligosaccharide is attributed to a neoagarotetraose that contain two LA units, two G units, one sulfate group and one methyl group (exact mass of the $[M-H]^-$ calculated at 723.1655916). This tetrasaccharide seems to be the terminal form of the products released from *O. pinnatifida* by the action of *Zg*AgaC. The OP30 fraction contained a larger species. This oligosaccharide was measured as a $[M-2H]^-$ at m/z 629.14 (Fig. 5C) and attributed to a species containing two LA units, four G (or L) units, two sulfate group and one methyl group (exact mass of the $[M-2H]^{2-}$ calculated at 629.131679). In order to decipher the complete structure of this species, this ion was further studied using extreme ultraviolet photon activation tandem mass spectrometry (XUV-DPI) at the synchrotron SOLEIL facility on the DISCO beamline (36). In contrast to classical tandem MS approach, XUV-DPI allow to obtain a definitive structural characterization of oligosaccharides and, especially, a complete description of the methylation and sulfate patterns (15). The XUV-DPI spectrum (Fig. 6A) allow to attribute the specie at m/z 629.14 to a L6S-G-(2-O-Me)-LA-G(2-Pentose)-LA2S-G as described in details in Fig6. B. This structure highlights the diversity of the subunits and linkages which can be found in the *O. pinnatifida* agar and the originality of the structures which can be released by *Zg*AgaC.

Structural comparison of ZgAgaC with β-agarases and β-pophyranases

The crystal structure of *Zg*AgaC was solved at high resolution (1.3 Å) using the automatic molecular replacement pipeline MoRDa (37). The crystal was orthorhombic ($P2_12_12$) and its unit cell dimension was as follows: a = 60.6 Å, b = 101.5 Å, c = 46.7 Å. Its asymmetric unit contains one protein chain, one calcium ion and 331 water molecules. The electron density map was of high quality, allowing the modelling of the complete recombinant *Zg*AgaC (Thr69-Glu328), and even of the residues corresponding to the BamHI cloning site (Gly-Ser) and of the N-terminal His_6-tag. Interestingly the His_6-tag of the *Zg*AgaC chain in the asymmetric unit interacts with residues of the active site of a symmetric *Zg*AgaC chain. This unusual interaction likely favored the crystallization of *Zg*AgaC and strengthened the crystal packing.

*Zg*AgaC adopts a β-jelly roll fold typical of the GH16 family (Fig. 7) and, despite extreme sequence divergences (21-31% sequence identity), superimposed well to β-agarases and β-porphyranases with a root mean square deviation (rmsd) of 1.323 Å with *Zg*AgaB (PDB 4ATF) and 0.921 Å with *Zg*PorA (PDB 3ILF). The sequence alignment of *Zg*AgaC with structurally-characterized β-agarases and β-porphyranases confirms that the secondary structures of these GH16 galactanases are essentially conserved (Fig. 8). The calcium ion is bound to the carbonyl group of Glu95 and Gly128 and to the side chain of Asp318 (OD1). These residues constitute a calcium-binding site which is conserved in most GH16 enzymes (22,32,33) and was shown in lichenases to enhance the protein thermostability (38). The catalytic machinery of the GH16 family, ExDx(x)E, is well conserved in *Zg*AgaC with Glu188, Asp190 and Glu193 predicted as the nucleophile, the catalytic helper and the acid/base catalyst, respectively (Fig. 8 and 9). A fourth residue, a

histidine (His215, *Zg*AgaB numbering, Fig. 9B) which forms a hydrogen bond with the conserved aspartate, is usually considered to cooperate with this catalytic helper in proton trafficking during the deglycosylation step (32,39). This histidine residue is conserved in most GH16 enzymes with a β-bulged active site (e.g. laminarinases (33), κ-carrageenases (32), β-agarases (22) and β-porphyranases (13)) and even in the GH7 family β-glucanases (39) which constitute the clan GH-B with the GH16 family. Interestingly, this histidine is substituted by a phenylalanine in *Zg*AgaC (Phe215, Fig. 8). This hydrophobic residue, which is strictly conserved in the *Zg*AgaC-like homologues (Supplementary Fig. S2), cannot play a similar function. Nonetheless Gln213, which is also located in strand β13 (Fig. 9), is hydrogen bonded to Asp190 (Gln213 NE2 – Asp190 OD2: 2.97 Å). This glutamine is invariant in the *Zg*AgaC-like subgroup (Supplementary Fig. S2) and may functionally replace the usually conserved histidine.

The representation of the molecular surface of *Zg*AgaC indicates that its active site is an open groove (Fig. 10). Such an active site topology is consistent with the endolytic mode of action of this enzyme (Fig. 3). With the exception of the acidic catalytic residues, the active groove displays a strong basic character (Fig. 10). The two major positively charged patches are due to Arg186 (negative subsites) and to Arg224 and Lys226 (positive subsites). These basic amino acids are invariant among the *Zg*AgaC-like homologues (Supplementary Fig. S2) and are strong candidates for interacting with the negatively charged substituents of complex agars (e.g. sulfate groups, uronic acids).

To have a more precise idea about the molecular bases of agar recognition by *Zg*AgaC, we compared this structure to the structure of *Zg*AgaB (24) and of *Zg*PorA (13) in complex with their respective substrates (Fig. 9). Among the residues involved in substrate recognition, only two residues are strictly conserved with either *Zg*AgaB or *Zg*PorA: Trp110 and Trp177 (*Zg*AgaC numbering). The tryptophans equivalent to Trp117 stack against the pyranose ring of the D-galactose moiety at subsite -1 in *Zg*PorA (Fig. 9A) and *Zg*AgaB (Fig. 9B). An aromatic residue is found at this position in most GH16 enzymes (13,22,33,40-44). Trp110 is more significant for the specificity of *Zg*AgaC. Indeed, the equivalent tryptophans in *Zg*AgaB and *Zg*PorA do not superimpose with Trp110, despite their invariance in term of sequence (Fig. 8). The distance between the Cα of Trp110 and of its equivalent tryptophan in *Zg*AgaB (Trp109) is 2.65 Å (2.31 Å with Trp56 in *Zg*PorA) and the differences are even greater for the side chains. This is due to the dissimilarities in length and composition of the loops between the strands β2 and β3 (8, 10 and 11 residues in *Zg*AgaC, *Zg*PorA and *Zg*AgaB, respectively; Fig. 8) which result in a different positioning of the conserved tryptophan in each type of enzyme. In the *Zg*PorA complex, Trp56 is nearly perpendicular to the ring of the L-galactose-6-sulfate at the -2 subsite while its NE1 atom establishes a hydrogen bond with the O6 of the D-galactose at the –3 subsite (Fig. 9A). Trp109 forms a parallel hydrophobic platform for the D-galactose moiety at -3 subsite in the *Zg*AgaB complex (Fig. 9B). Trp110 (*Zg*AgaC) overlaps neither Trp109 (*Zg*AgaB) nor Trp56 (*Zg*PorA), suggesting that this residue likely interacts with a sugar moiety in the subsite -2 or -3 of ZgAgaC, but in a fashion differing from both β-agarases and β-porphyranases. Some residues of *Zg*AgaC occupy similar three-dimensional positions than non-equivalent residues from *Zg*AgAB or *Zg*PorA. Thus, the guanidine group of Arg186 (located in the strand β11 of *Zg*AgaC) and of Arg133 (located in the strand β10 of *Zg*PorA) nearly overlap (Fig. 9A). Arg133 is a key residue of *Zg*PorA which recognizes the sulfate group of L6S at the -2 subsite (13), but it is not conserved in the *Zg*AgaC sequence. The functional group of Arg219 (*Zg*AgaB) is also apparently close to that of Arg186 (*Zg*AgaC, Fig. 9B). However, a global look at the superimposition of these enzymes indicates that the side chain of Arg219 (*Zg*AgaB) would clash sterically with the loop Asp181-Arg186 of *Zg*AgaC and thus cannot play a similar role. The side chain of Trp297 (*Zg*AgaC) and Trp312 (*Zg*AgaB) are also roughly at the same position (Fig. 9B), although these tryptophans originate from loops differing in length and sequence (Fig. 8). Trp312 forms a hydrophobic platform for the D-galactose bound in the +2 subsite of *Zg*AgaB (24).

There are also some striking substitutions between *Zg*AgaC and the other galactanases. Notably, Glu308 (*Zg*AgaB

numbering) is strictly conserved not only in β-agarases but also in the *Zg*PorA-like and *Zg*PorB-like β-porphyranases (Fig. 8). This glutamate is hydrogen-bonded to the axial hydroxyl group OH4 of the D-galactose at the cleavage subsite -1 and is considered as an essential difference between β-galactanases and β-glucanases in the GH16 family (23,24). In *Zg*AgaC, this key glutamate is replaced by a tryptophan (Trp291, Fig. 8 and Fig. 9) which is strictly conserved in the *Zg*AgaC homologues (Supplementary Fig. S2). Similarly, Gly111 (*Zg*AgaB numbering) is strictly conserved in β-agarases and β-porphyranases (Fig. 8), likely to make room for the neighboring residues Trp312 and Glu308. This glycine is replaced by a tyrosine in *Zg*AgaC (Tyr112) which stacks against Trp291 and points towards the -1 subsite (Fig. 9). Tyr112 is also invariant in the *Zg*AgaC subgroup (Supplementary Fig. S2). Such a drastic substitution was rendered possible by the above mentioned shift of Trp110. Considering their position, it is most likely that Tyr112 and Trp291 are involved in the recognition of the D-galactose moiety at the -1 subsite, possibly through their functional groups OH and NE1, respectively.

Finally, due to the particular crystal packing involving the His6-tag, it was not possible to attempt a co-crystallization with agar oligosaccharides using the initial *Zg*AgaC construct. To solve this problem, we succeeded in overproducing another construct of *Zg*AgaC with an N-terminal GST-tag and a C-terminal His-tag. This recombinant enzyme was inactivated by the mutation of the acid/base catalyst Glu193 into a threonine. We developed a protocol for the cleaving of the GST-tag and the purification of the cleaved enzyme *Zg*AgaC$_{E193T}$. Unfortunately, we were not able to crystallize this recombinant protein.

DISCUSSION

Zobellia galactinovorans Dsij^T was isolated from *Delesseria sanguinea* (45), a red alga which produces a highly substituted agar (46,47). In its natural environment, this agarolytic bacterium is expected to feed on such complex agars, notably originating from red algae of the orders Bangiales, Corallinales, Gracilariales, and Ceramiales (which includes *D. sanguinea* and *O. pinnatifida*) (8). To date three β-agarases and two β-porphyranases from *Z. galactinovorans* have been characterized (13,21-24). Notably they were shown to be more or less tolerant to some substitutions, such as the sulfation at C6 of the L unit and the methylation of the C6 of the G unit (24). However, it is yet unclear whether *Z. galactanivorans* can cope with other modifications found in complex agars. Our present work on *Zg*AgaC gives some insight into this question.

Although *Zg*AgaC was annotated as a putative β-agarase in *Z. galactinovorans* genome (17), this enzyme is highly divergent from characterized GH16 β-agarases (23% and 21% with *Zg*AgaA and ZgAgaB, respectively). Despite this distant relationship, we have demonstrated here that a recombinant *Zg*AgaC can degrade agarose (Fig. 3) and is inactive on pure porphyran, indicating at first sight that this enzyme is a classical β-agarase. However, an activity screening on a collection of 13 agarophytes has revealed that the degradation patterns of *Zg*AgaB and *Zg*AgaC differed significantly for several species (e.g. *P. elongata* and *O. pinnatifida*), suggesting that the substrate specificity of *Zg*AgaC is broader than expected. Before discussing in more details the exact specificity of *Zg*AgaC, it is noteworthy that the pH optimum of *Zg*AgaC strongly varies depending on the nature of the substrate (pH 6.5 and pH 9.1 with agarose and the *O. pinnatifida* agar, respectively; Fig. 4). The presence of a microenvironment around the polymer may explain this phenomenon. There is little documentation on the influence of a microenvironment on carbohydrate-active enzymes; however, a previous study by Li an coworkers (48), demonstrates a significant positive influence of charged (non-substrate) polysaccharides on enzyme stability over a wide pH range. For highly sulfated agars, protons are likely required as counter ions, creating an acidic environment around the surface of the polymer. In contrast, agarose does not display charged modifications and thus there is probably no such pH microenvironment. If this is to be placed in a biological context, the pH microenvironment around the acidic polysaccharide will be different than that which is observed further from the algal surface. Thus, an apparent pH optimum of 9.1 is likely overestimated on the molecular level near the active site and doesn't reflect the reality of the local environment.

This also raises the interesting question of whether the pH of the local environment plays a molecular role in the regulation of the activities of the different GH16 enzymes from *Z. galactanivorans.*

By a combination of UHPLC and XUVPDI-MS/MS, we have succeeded in characterizing purified oligosaccharides released by the action of *Zg*AgaC on the *O. pinnatifida* agar. The smallest characterized product is mono-sulfated and mono-methylated neoagarotetraose (LA-G-LA-G), although we were unable to determine the exact position of the substituents. Nonetheless, the structure of a hexasaccharide was completely elucidated: L6S-G-LA2Me-G(2Pentose)-LA2S-G. This result is consistent with the chemical composition of the *O. pinnatifida* agar which was shown to contain D-xylose branching (10,11). Thus, the pentose on the hexasaccharide is likely a D-xylose moiety. The structure of the hexasaccharide is reminiscent of that proposed for the tetratrasccharide product. This suggests that the sulfate and methyl groups of this tetrasaccharide could be also on the C2 of the LA units at the reducing-end and non-reducing end, respectively. In any case, this MS analysis demonstrates that *Zg*AgaC can act on a highly complex agar motif. Taking into account these oligosaccharide structures, the activity of *Zg*AgaC on agarose and its inactivity on pure porphyran, we can deduce the following characteristics of the active site of *Zg*AgaC: (i) the subsites -2 and -1 are specific for a neoagarobiose unit, either unmodified (agarose) or sulfated at C2 on the LA moiety; (ii) the subsite +1 and 2 can bind either neoagarobiose unit (as in agarose) or neoporphyranobiose (L6S-G; as found on the non-reducing of the hexasaccharide). The fact that *Zg*AgaC cannot hydrolyze pure porphyran whereas it is able to bind neoporphyranobiose at the subsites +1 and +2 strongly indicates that the subsites -2 and -1 cannot recognize a neoporphyranobiose unit; (iii) the subsites -3 and -4 can bind either neoagarobiose unit (as in agarose) or LA2Me-G(2Pentose).

The crystal structure of *Zg*AgaC is not particularly different from β-agarases and β-porphyranases at the fold level, almost all secondary structures being conserved (Fig. 7 and 8). In contrast, the *Zg*AgaC active site is strongly modified in comparison to that of the other galactanases. Based on the structural comparison with the complex structures of *Zg*AgaB (24) and *Zg*PorA (13) (Fig. 9) and on the sequence alignment of the *Zg*AgaC homologues (Supplementary Fig. S2), we can predict that the following conserved residues of *Zg*AgaC are involved in substrate recognition: Trp110 (subsite -2 or -3), Arg186 (subsite -2), Tyr112, Trp177, Trp291 (subsite -1), Lys226, Arg244 (subsite +1) and Trp297 (subsite +2). Among these residues, only Trp110 and Trp177 are conserved with β-agarases and β-porphyranases, but the position of Trp110 is shifted in comparison to the equivalent tryptophan in the other enzymes. The most spectacular change is the substitution of a tryptophan (Trp291) for the conserved glutamate (Glu308, *Zg*AgaB numbering) that is important in recognition of the D-galactose unit at the -1 cleavage subsite in both β-agarases and β-porphyranases (Fig. 9). We propose that both Trp291 and Tyr112 functionally replace this crucial glutamate. Considering their spatial location, Arg186 and Lys226/Arg224 most likely interact with the sulfate groups of the complex hexasaccharide (LA2S (-2 subsite) and L6S (+1 subsite), respectively). Altogether, *Zg*AgaC recognizes agarose in a completely different way than classical β-agarases and is also adapted to bind highly substituted agar hybrids (i.e. sulfated, methylated and/or branched oligoagars).

The updated phylogenetic tree of the GH16 galactanases (Fig. 1 and Supplementary Fig. S1) unambiguously supports that the *Zg*AgaC homologues constitute a monophyletic clade distinct from the classical β-agarases and the β-porphyranases. Considering their structural and activity differences, we thus propose that the *Zg*AgaC-like enzymes form a new subfamily within the GH16 family. Another important result of this phylogenetic analysis is that *Zg*PorA-like and *Zg*PorB-like enzymes now constitute two solid, distinct clades. This is not a complete surprise for two reasons: (i) *Zg*PorA and *Zg*PorB display only 24% of sequence identity; (ii) in the initial phylogenetic tree of the GH16 galactanases in 2010 (13) the node connecting the *Zg*PorA-like and the *Zg*PorB-like enzymes was supported by a low bootstrap value (55%). Without entering into details, the superimposition of *Zg*PorA and *Zg*PorB (Fig. 9C) confirms that these GH16 enzymes

recognize porphyran in a partially different way (e.g. the L6S at -2 subsite is not recognized by the same arginine). More generally, this updated phylogenetic analysis supports that the *Zg*PorA homologues are the most early diverging type of agar-specific enzymes, solidly rooting a group comprising the *Zg*PorB-like clade, the *Zg*AgaC-like clade and the classical β-agarases. Moreover, the κ-carrageenases constitute a sister clade of all the agar-specific enzymes. Therefore, the common ancestor of GH16 galactanases was most likely an enzyme acting on complex, sulfated galactans rather than on a neutral galactan. The *Zg*AgaC and *Zg*PorB homologues also act on sulfated galactans suggesting that the classical β-agarases, which are more specific for neutral agarose, have diverged more recently.

To summarize, this study highlights the diversity of the GH16 agar-specific enzymes which appeared as a bacterial response to red macroalgae cell walls in order to cope with the complexity of natural agars. However, this creates a practical difficulty in terms of enzyme nomenclature. How can we simply distinguish *Zg*AgaC-like enzymes from classical β-agarases or the *Zg*PorA-like from the *Zg*PorB-like β-porphyranases? The Greek letter nomenclature of carrageenans is particularly useful for naming the different types of carrageenases (e.g.. κ-, ι-, λ-, β-carrageenases). A similar nomenclature for agars and agarases could be a solution. But until a satisfying nomenclature is created, we recommend mentioning the type of β-agarases (classical or ZgAgaA-like; ZgAgaC-like) and β-porphyranases (*Zg*PorA-like; *Zg*PorB-like).

EXPERIMENTAL PROCEDURES

Except when mentioned, all chemicals were purchased from Sigma (France)

Phylogenetic analysis

Homologues of *Zg*AgaC (17), of the β-agarase *Zg*AgaA (21), of the β-porphyranases *Zg*PorA and *Zg*PorB (13) and of the κ-carrageenase *Zg*CgkA (40) were identified in the Genbank database using BlastP (49). After removal of duplicated sequences, all these GH16 galactanases were aligned, with the laminarinases *Zg*LamA (33) and *Zg*LamB (34) as outgroup sequences, using MAFFT with the iterative refinement method and the scoring matrix Blosum62 (50). These alignments were manually edited with BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) on the basis of the superposition of the crystal structures of *Zg*AgaC, *Zg*AgaA (22), *Zg*PorA, *Zg*PorB (13), *Zg*CgkA (40), *Zg*LamA (33) and *Zg*LamB (34). This multiple alignment allowed calculation of model tests and maximum likelihood trees with MEGA version 6.0.6 (51). Tree reliability was tested by bootstrap using 100 resamplings of the dataset. The trees were displayed with MEGA 6.0.6.

Cloning and site-directed mutagenesis of the agaC gene

The *agaC* gene from *Z. galactanivorans* encoding the *Zg*AgaC protein (locus identifier: ZGAL_4267, GenBank™ accession number FQ073843.1) was cloned as by *Groisillier* and coworkers (52). Briefly primers were designed to amplify the coding region corresponding to the catalytic module of *Zg*AgaC (forward primer AAAAAAGGATCCACCTATGATTTTACCGGAAACACCC, reverse primer TTTTTTCTGCAGTTATTCCTCTACCAATTGATAGGTATG) by PCR from *Z. galactanivorans* genomic DNA. After digestion with the restriction enzymes BamHI and PstI, the purified PCR product was ligated using the T4 DNA ligase into the expression vector pFO4 (52) predigested by BamHI and NsiI, resulting in a recombinant protein with a N-terminal hexa-histidine tag. This plasmid, named pZG234, was transformed into *Escherichia coli* DH5α strain for storage and in *E. coli* BL21(DE3) strain for protein expression. The putative nucleophile Glu193 was replaced by a threonine by site-directed mutagenesis performed using the QuickChange II XL site directed mutagenesis kit (Stratagene), the pZG234 plasmid and the primers CAGGAGGGGTTGGAGTTATTCGTTATTCGTTATAACGTCAATTTCGTTACG and GTCCTCCCCAACCTCAATAAGCAATATTGCAGTTAAAGCAATGC. The resulting plasmid is named pZG234$_{E193T}$. For expressing a variant of ZgAgaC$_{E193T}$ with a C-terminal His-tag instead of an N-terminal His-Tag, the cloning was performed using the In-Fusion HD Cloning Kit (Clontech) and the manufacturer's protocol was followed. Briefly the gene was

amplified from pZG234$_{E193T}$ with the primers: 5' GGGGCCCCTGGGATCCGGATCCACCTA TGATTTTACCGGAAA 3' and 5' AGTCACGATGCGGCCGCTTATTCCTCTA CCAATTGATAGGTATGTATCCAGTCTAT TTTCATGGTGT3', these primers bearing the 15-bp homology necessary for InFusion cloning in pGex-6p3. pGex6p3 was digest by BamHI and NotI. All PCR amplifications were done with the high-fidelity polymerase CloneAmp (Clontech). The resulting plasmid is named pGEX_ZG234$_{E193T.}$ Plasmid amplifications were performed in *E. coli* XL10-Gold Ultracompetent Cells (Stratagene).

Overexpression and purification of ZgAgaC and ZgAgaC$_{E193T}$

E. coli BL21(DE3) cells harboring the plasmid pZG234 were cultivated at 20°C in a 1 L auto-induction ZYP 5052 medium (53) supplemented with 100 µg.mL^{-1} ampicillin. Cultures were stopped when the cell growth reached the stationary phase and were centrifuged for 35 min at 4°C, 3,000 *g*. The cells were resuspended in a 20 mL of buffer A (tris 50 mM pH 8, NaCl 500 mM, imidazole 20 mM) and chemically lysed as previously described (20). Afterwards the lysate was clarified at 12,000 *g* for 30 min at 4°C and the supernatant was filtered on 0.22 µm. The supernatant was loaded onto a HyperCell PAL column charged with $NiCl_2$ (0.1 M) and pre-equilibrated with buffer A. The column was washed with buffer A and the protein was eluted with a linear imidazole gradient produced by the mixing of buffer A and buffer B (tris 50 mM pH 8, NaCl 500 mM, imidazole 500 mM) at a flow rate of 1 mL.min^{-1}. The different fractions were concentrated on Amicon Ultra 15 (10 kDa) Merck Millipore to reach a volume of 2 mL. Finally, the protein was injected onto Sephacryl S-200 size exclusion column (GE Healthcare) pre-equilibrated with buffer C (tris 50 mM pH8, NaCl 300mM).

The fusion protein GST-*Zg*AgaC$_{E193T,}$ which has an N-terminal GST tag and a C-terminal His_6 tag, was produced from *E. coli* BL21 cells harboring the plasmid pGEX_ZG234$_{E193T}$ with the same protocol used for *Zg*AgaC. The cells were resuspended in 20 mL of buffer D (tris 50 mM pH 8, NaCl 200 mM, DTT 1 mM) and chemically lysed. After clarification as described above, the supernatant was loaded onto a 5 mL GST trap 4B column (GE Healthcare) equilibrated with buffer D. The column was washed extensively with buffer D and the elution was performed with buffer E (tris 50 mM pH 8, NaCl 200 mM, DTT 1 mM, glutathione 50 mM, agarose oligosaccharides 7 g.L^{-1}). The agarose oligosaccharides were produced as described by Jam and coworkers (21). For removal of the GST tag, GST-*Zg*AgaC$_{E193T}$ was incubated with the GST-tagged human rhinovirus 3C protease (GST-PreScission) (1 µM) for 16 h. The cleaved protein, referred to as *Zg*AgaC$_{E193T}$, was separated from the free GST tag and the GST-PreScission protease by injection onto a 5 mL GST trap 4B column. The column was washed with buffer F (tris 50 mM pH 8, NaCl 400 mM, DTT 1 mM, glutathione 50 mM) and *Zg*AgaC$_{E193T,}$ which has affinity for the sepharose matrix, was eluted with buffer G (tris 50mM pH8, NaCl 200mM, DTT 1mM, agarose oligosaccharides 7 g.L^{-1}). A final size-exclusion chromatography was undertaken with a Sephacryl S-200 column (GE) pre-equilibrated with buffer C.

Comparison of the degradation pattern between ZgAgaC, ZgAgaB and ZgPorB

Thirteen species of agarophyte red algae were harvested in June 2016 in Roscoff (Brittany, France): *Osmundea pinnatifida*, *Dumontia contorta*, *Polysiphonia simulans*, *Polysiphonia elongata*, *Polysiphonia brodiei*, *Rhodomella sp.*, *Chondria dasyphylla*, *Cryptopleura ramosa*, *Gracillaria sp.*, *Vertebrata lanosa*, *Ceramium rubrum*, *Chylocladia verticillata* and *Dumontia contorta*. These algae were cryo-ground with CryoMill (Retsch). For comparison purposes, the β-agarase *Zg*AgaB and the β-porphyranases *Zg*PorB from *Z. galactinovorans* were produced and purified as previously described (13,21). The ground algae were resuspended in the buffer of the tested enzymes. Each reaction mixture contained 1 µM of enzyme and 0.1 g.mL^{-1} of ground algae and was incubated at 35°C under agitation during 24 h. These reaction mixtures were centrifuged at 11,000 *g* over 20 min and the supernatants were conserved at -20°C for subsequent fluorophore-assisted carbohydrate electrophoresis analyses.

Fluorophore-assisted carbohydrate electrophoresis

The different degradation reactions and the oligosaccharide fractions were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE) (54). 100 µL of the enzymatically degraded algae or 1 mL of the purified oligosaccharide fractions were dried in a speed-vacuum centrifuge. Oligosaccharides were labeled with 2-aminoacridone (AMAC) or 8-aminonaphtalene-1,3,6-trisulfonate (ANTS) as previously described (55). Briefly, for fluorophore labeling, the dried oligosaccharide pellet was dissolved with 2 µL of AMAC or ANTS solution and 2 µL of 1 M sodium cyanobohydride in dimethyl sulfoxyde (DMSO) was added. The mixture was incubated at 37°C for 16 h in the dark. Glycerol 20% (20 µL) was added to the samples before loading onto a 27% polyacrylamide gel. The electrophoresis was carried out at 4°C in the dark during 2 h at 175 V.

Extraction and preparation of natural agars

Natural agars from *Porphyra* sp. and *Osmundea pinnatifida* were extracted as follows. Algae were treated to obtain alcohol insoluble residues (AIR) as described by Hervé and coworkers (56). Briefly, the dried algae were successively washed with 70% ethanol, 96% ethanol, methanol:chloroform (1:1, v/v) and acetone 100%. Each step was repeated three times. After this treatment, agars from *Porphyra sp.*, referred to as porphyran, and from *O. pinnatifida* were extracted using an autoclave at 100°C at 1 bar over 1 hour. The polysaccharides were precipitated by addition of four volumes of ethanol and were retrieved by centrifugation at 6,000 *g* over 30min.

Porphyran is usually constituted of one third of agarobiose motifs and two thirds of porphyranobiose motifs (13). It was necessary to undertake enzymatic assays on a substrate containing only porphyranobiose motifs. Thus, the native porphyran was solubilized in water at 1% and digested by *Zg*AgaB (21) at final enzyme concentration of 1,5 µM. The reaction mixture was filtered on Amicon Ultra 15 (3kDa-cuttoff). The retentate was recovered and freeze-dried for use in the enzymatic assays. This polysaccharidic fraction is referred to as pure porphyran (i.e without agarobiose motifs).

Enzymatic assays

The thermostability of *Zg*AgaC was evaluated by Dynamic Light Scattering (DLS) measurements at different temperatures (Malvern). The enzyme was initially tested by FACE to determine potential substrates. The amount of reducing ends produced by enzymatic digestion was followed using a method adapted from Kidby and Davidson (57). Aliquots of reaction medium (20 µL) were mixed with 180 µL of ferrycyanide solution (300 mg of potassium hexocyanoferrate III, 29 g of Na_2CO_3, 1 mL of 5 M NaOH, completed to 1 L with water). The mixture was heated to 95°C over 5 min and cooled down to 4°C. Its absorbance was measured at 420 nm. The pH optimum of *Zg*AgaC was determined by monitoring enzymatic activity at 25°C and in pH range of 3.85 to 10.8 for each polysaccharide. Several buffers were separately used to measure the pH optimum: (i) for agarose, 100 mM phosphate was used between pH 6 to 8, 100 mM sodium acetate between pH 3.85 to 5 and 100 mM sodium borate for pH 9; (ii) for the *O. pinnatifida* agar, 100 mM tris was used between pH 7 and 9 and glycine NaOH between pH 9 and 10.5. For all these buffers 150 mM NaCl was added. The temperature optimum of *Zg*AgaC was evaluated by activity measurement on the *O. pinatifida* agar at temperatures ranging from 25°C to 65°C.

To determine the kinetic parameters of *Zg*AgaC, the substrate concentrations ranged from 0.0075% to 0.2% (w/v) for agarose, 0.1% to 0.8% (w/v) for the *O. pinnatifida* agar. Reducing end equivalents were determined by the ferrycyanide assay with galactose standard curves. The reaction buffers were 100 mM phosphate pH 6.5, 150 mM NaCl for agarose and 100 mM glycine pH 9, 150 mM NaCl for the *O. pinnatifida* agar. All the enzymatic assays were performed at 50°C. The reactions were monitored over 80 sec with a point every 10 sec and all values were determined in triplicate. The molar concentration of agarose was determined using the molecular mass of the repeating unit neoagarobiose. The K_M and the k_{cat} were determined by a nonlinear regression analysis using the program R.

Purification of Osmundea pinnatifida oligoagars

The agar from *Osmundea pinnatifida* was dissolved at 1% (w/v) and incubated at 37°C for 48 h with 1 µM of purified *Zg*AgaC. Degradation products were ultrafiltrated with Amicon Ultra 15 (3 kDa, Merck Millipore) and the oligosaccharides contained in the filtrate were further purified as previously described (21). Briefly, the oligo-agars were purified by preparative size exclusion chromatography with three columns of Superdex 30 (26/60) (GE) in series which were operated on an HPLC system (Gilson). Detection was performed using a refractive index detector (Spectra System RI-50). The purification was monitored by Unipoint Software (Gilson). The elution was performed using 50 mM of ammonium carbonate ($(NH_4)_2CO_3$) at a flow rate of 1 $mL.min^{-1}$. The fractions were collected and freeze-dried before mass spectrometry analyses.

ESI MS measurements

The mass measurements were performed on a Synapt G2Si high-definition mass spectrometer (Waters Corp., Manchester, UK) equipped with an Electrospray ion (ESI) source. The instrument was operated in negative polarity, as well as in 'sensitivity' mode. Sample were diluted at 10 $µg.ml^{-1}$ in a solution of H_2O/MeOH (50:50) and infused with a flow rate of 3 $µl.min^{-1.}$

XUV-DPI tandem MS measurements

The experimental setup of the extreme ultraviolet dissociative photoIonization (XUV-DPI) was developed at the SOLEIL synchrotron radiation facility at the endstation of the DISCO beamline (58). A bending magnet-based synchrotron beamline was coupled to a linear ion trap (LTQ XL, Thermo Fisher Scientific). An automatic shutter was used to synchronize the photon beam (tuned to 18 eV, 68.9 nm) with the trapped precursor ions. Precursor ions were isolated with a window of 2 Da and exposed to XUV photons for 500 ms and were ejected for their measurement after a delay of 50 ms. Sample was diluted to a concentration of 50 $µg.mL^{-1}$ and infused at a flow rate of 5 $µL.min^{-1}$. Measurements were performed in negative ion mode on the doubly charged species observed at 629.1 m/z. The nomenclature used for annotations is according to that defined by Domon and Costello (59). Raw data were processed with mMass 5.3.0 (60).

Crystallization and structure determination of ZgAgaC

*Zg*AgaC and *Zg*$AgaC_{E193T}$ were concentrated at 50 $mg.mL^{-1}$ and 70 $mg.mL^{-1}$, respectively, and stored at 4°C in buffer C (tris 50 mM pH8, NaCl 300mM). Crystallization screening was undertaken with the nanodrop-robot Crystal Gryphon (Art Robbins instruments) with the four-sparse matrix sampling kits (Qiagen and Molecular Dimensions). For *Zg*$AgaC_{E193T,}$ 0.5 mg of agar oligosaccharides (neoagarotetraose and neoporphyranotetraose, gifts from Dr. F. Le Sourd; terminal products released by the action of ZgAgaC on the *O. pinnatifida* agar) were added to the protein solution prior to the crystallization screening. The initial crystallization conditions were manually optimized and single crystals were obtained using the hanging drop vapor diffusion method as follows. For *Zg*AgaC 2 µL of enzyme (50 $mg.mL^{-1}$) were mixed with 1 µL of reservoir solution containing 2.1 M sodium malonate and 1% glycerol. For *Zg*$AgaC_{E193T}$, we undertook a sparse matrix based on the screening results, unfortunately no crystals were obtained.

Diffraction data for a *Zg*AgaC crystal were collected at 1.3 Å resolution on the ID29 beamline (ESRF, Grenoble, France). X-ray diffraction data were integrated using XDS (61) and scaled with Scala (62). The structure was solved by molecular replacement, using the automatic pipeline MoRDa (37). The model provided by MoRDa was further manually modified and corrected using COOT (63) and refined with REFMAC5 (64).

ACKNOWLEDGEMENT

We are indebted to the local contacts for their support during X-ray data collection at the ID29 Beamline, European Synchrotron Radiation Facility. XUV-DPI Experiments were performed on the DISCO beamline at SOLEIL Synchrotron, France (proposal number 20161299) and we are grateful to the SOLEIL staff, and especially Dr. Alexandre Giuliani, for smoothly running the facility. We are grateful to the Marine Core Facility (Station Biologique de Roscoff, SBR) for its help in collecting red algal species and to Dr. Frédéric Le Sourd for providing agar oligosaccharides. We thank Alexandra Jeudy (SBR crystallization platform) for her help and advices for the crystallization screening. We also thank Dr. Elizabeth Ficko-Blean for helpful discussions and critical reading of our manuscript. A.N. thanks Brittany and the "Pays de Loire" region for her PhD fellowship granted in the context of the Glyco-Ouest interregional network (http://www.glyco-ouest.fr/). M.C., C.T. and G.M. are grateful to the French National Research Agency (ANR) for its support with regards to the investment expenditure program IDEALG (http://www.idealg.ueb.eu/, grant agreement No. ANR-10-BTBR-04). G.M. also acknowledge support from ANR with regard to the "Blue Enzymes" project (reference ANR-14-CE19-0020-01). D.R., M.F. and H.R are grateful for the ANR support under project number ANR-08- BLAN-0065.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest for the contents of this article.

REFERENCES

1. Ficko-Blean, E., Hervé, C., and 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
2. Lawson, C. J., and Rees, D. A. (1970) An enzyme for the metabolic control of polysaccharide conformation and function. *Nature* **227**, 392-393
3. Genicot-Joncour, S., Poinas, A., Richard, O., Potin, P., Rudolph, B., Kloareg, B., and Helbert, W. (2009) The cyclization of the 3,6-anhydro-galactose ring of iota-carrageenan is catalyzed by two D-galactose-2,6-sulfurylases in the red alga *Chondrus crispus*. *Plant Physiol.* **151**, 1609-1616
4. Collen, J., Porcel, B., Carre, W., Ball, S. G., Chaparro, C., Tonon, T., Barbeyron, T., Michel, G., Noel, B., Valentin, K., Elias, M., Artiguenave, F., Arun, A., Aury, J. M., Barbosa-Neto, J. F., Bothwell, J. H., Bouget, F. Y., Brillet, L., Cabello-Hurtado, F., Capella-Gutierrez, S., Charrier, B., Cladiere, L., Cock, J. M., Coelho, S. M., Colleoni, C., Czjzek, M., Da Silva, C., Delage, L., Denoeud, F., Deschamps, P., Dittami, S. M., Gabaldon, T., Gachon, C. M., Groisillier, A., Herve, C., Jabbari, K., Katinka, M., Kloareg, B., Kowalczyk, N., Labadie, K., Leblanc, C., Lopez, P. J., McLachlan, D. H., Meslet-Cladiere, L., Moustafa, A., Nehr, Z., Nyvall Collen, P., Panaud, O., Partensky, F., Poulain, J., Rensing, S. A., Rousvoal, S., Samson, G., Symeonidi, A., Weissenbach, J., Zambounis, A., Wincker, P., and Boyen, C. (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**, 5247-5252
5. Brawley, S. H., Blouin, N. A., Ficko-Blean, E., Wheeler, G. L., Lohr, M., Goodson, H. V., Jenkins, J. W., Blaby-Haas, C. E., Helliwell, K. E., Chan, C. X., Marriage, T. N., Bhattacharya, D., Klein, A. S., Badis, Y., Brodie, J., Cao, Y., Collen, J., Dittami, S. M., Gachon, C. M. M., Green, B. R., Karpowicz, S. J., Kim, J. W., Kudahl, U. J., Lin, S., Michel, G., Mittag, M., Olson, B., Pangilinan, J. L., Peng, Y., Qiu, H., Shu, S., Singer, J. T., Smith, A. G., Sprecher, B. N., Wagner, V., Wang, W., Wang, Z. Y., Yan, J., Yarish, C., Zauner-Riek, S., Zhuang, Y., Zou, Y., Lindquist, E. A., Grimwood, J., Barry, K. W., Rokhsar, D. S., Schmutz, J., Stiller, J. W., Grossman, A. R., and Prochnik, S. E. (2017) Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6361-E6370
6. Lahaye, M., and Rochas, C. (1991) Chemical structure and physico-chemical properties of agar. *Hydrobiologia* **221**, 137-148
7. Van de Velde, F., Knutsen, S., Usov, A., Rollema, H., and Cerezo, A. (2002) 1H and 13C high resolution NMR spectroscopy of carrageenans: application in research and industry. *Trends Food Sci. Tech.* **13**, 73-92
8. Usov, A. I. (2011) Polysaccharides of the red algae. *Adv. Carbohydr. Chem. Biochem.* **65**, 115-217
9. Knutsen, S., Myslabodski, D., Larsen, B., and Usov, A. (1994) A modified system of nomenclature for red algal galactans. *Bot. Mar.* **37**, 163-169
10. Bowker, D. M., and Turvey, J. R. (1968) Water-soluble Polysaccharides of the Red Alga *Laurencia pinnatifida*. Part 1. Constituent Units. *J. Chem. Soc.*, 983-988
11. Bowker, D. M., and Turvey, J. R. (1968) Water-soluble Polysaccharides of the Red Alga *Laurencia pinnatifida*.Part lI. Methylation Analysis of the Galactan Sulphate. *J. Chem. Soc.*, 989-992
12. Ferreira, L. G., Noseda, M. D., Goncalves, A. G., Ducatti, D. R., Fujii, M. T., and Duarte, M. E. (2012) Chemical structure of the complex pyruvylated and sulfated agaran from the red seaweed *Palisada flagellifera* (Ceramiales, Rhodophyta). *Carbohydr. Res.* **347**, 83-94
13. Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908-912
14. Correc, G., Hehemann, J. H., Czjzek, M., and Helbert, W. (2011) Structural analysis of the degradation products of porphyran digested by *Zobellia galactanivorans* b-porphyranase A. *Carbohydr. Polym.* **83**, 227-283

15. Ropartz, D., Giuliani, A., Herve, C., Geairon, A., Jam, M., Czjzek, M., and Rogniaux, H. (2015) High-energy photon activation tandem mass spectrometry provides unprecedented insights into the structure of highly sulfated oligosaccharides extracted from macroalgal cell walls. *Anal. Chem.* **87**, 1042-1049
16. Ropartz, D., Giuliani, A., Fanuel, M., Herve, C., Czjzek, M., and 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
17. Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., Goesmann, A., Leblanc, C., Oliver Glockner, 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
18. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-495
19. 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
20. Ficko-Blean, E., Duffieux, D., Rebuffet, E., Larocque, R., Groisillier, A., Michel, G., and Czjzek, M. (2015) Biochemical and structural investigation of two paralogous glycoside hydrolases from *Zobellia galactanivorans*: novel insights into the evolution, dimerization plasticity and catalytic mechanism of the GH117 family. *Acta Crystallogr. D* **71**, 209-223
21. Jam, M., Flament, D., Allouch, J., Potin, P., Thion, L., Kloareg, B., Czjzek, M., Helbert, W., Michel, G., and Barbeyron, T. (2005) The endo-beta-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
22. Allouch, J., Jam, M., Helbert, W., Barbeyron, T., Kloareg, B., Henrissat, B., and Czjzek, M. (2003) The three-dimensional structures of two beta-agarases. *J. Biol. Chem.* **278**, 47171-47180
23. Allouch, J., Helbert, W., Henrissat, B., and Czjzek, M. (2004) Parallel substrate binding sites in a beta-agarase suggest a novel mode of action on double-helical agarose. *Structure* **12**, 623-632
24. Hehemann, J. H., Correc, G., Thomas, F., Bernard, T., Barbeyron, T., Jam, M., Helbert, W., Michel, G., and Czjzek, M. (2012) Biochemical and Structural Characterization of the Complex Agarolytic Enzyme System from the Marine Bacterium *Zobellia galactanivorans*. *J. Biol. Chem.* **287**, 30571-30584
25. 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
26. McBride, M. J., and Zhu, Y. T. (2013) Gliding Motility and Por Secretion System Genes Are Widespread among Members of the Phylum *Bacteroidetes*. *J. Bacteriol.* **195**, 270-278
27. Groisillier, A., Labourel, A., Michel, G., and Tonon, T. (2015) The mannitol utilization system of the marine bacterium *Zobellia galactanivorans*. *Appl. Environ. Microbiol.* **81**, 1799-1812
28. Barbeyron, T., Michel, G., Potin, P., Henrissat, B., and Kloareg, B. (2000) iota-Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of kappa-carrageenases. *J. Biol. Chem.* **275**, 35499-35505
29. Rebuffet, E., Barbeyron, T., Jeudy, A., Jam, M., Czjzek, M., and Michel, G. (2010) Identification of catalytic residues and mechanistic analysis of family GH82 iota-carrageenases. *Biochemistry* **49**, 7590-7599
30. Ficko-Blean, E., Prechoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., Stam, M., Genicot, S., Jam, M., Calteau, A., Viart, B., Ropartz, D., Perez-Pascual, D., Correc, G., Matard-Mann, M., Stubbs, K. A., Rogniaux, H., Jeudy, A., Barbeyron, T., Medigue, C., Czjzek, M., Vallenet, D., McBride, M. J., Duchaud, E., and Michel, G. (2017) Carrageenan

catabolism is encoded by a complex regulon in marine heterotrophic bacteria. *Nat. Commun.* **8**, 1685
31. Thomas, F., Bordron, P., Eveillard, D., and 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
32. Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) The kappa-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
33. Labourel, A., Jam, M., Jeudy, A., Hehemann, J. H., Czjzek, M., and Michel, G. (2014) The b-glucanase *Zg*LamA from *Zobellia galactanivorans* evolved a bent active site Adapted for efficient degradation of algal laminarin. *J. Biol. Chem.* **289**, 2027–2042
34. Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J. H., Ferrieres, V., Czjzek, M., and Michel, G. (2015) Structural and biochemical characterization of the laminarinase *Zg*LamC$_{GH16}$ from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta Crystallogr. D* **71**, 173-184
35. Ostling, S., and Virtama, P. (1946) A modified preparation of the universal buffer described by Teorell and Stenhagen. *Acta Phys. Scandinav.* **11**, 289- 293
36. Ropartz, D., Lemoine, J., Giuliani, A., Bittebiere, Y., Enjalbert, Q., Antoine, R., Dugourd, P., Ralet, M. C., and Rogniaux, H. (2014) Deciphering the structure of isomeric oligosaccharides in a complex mixture by tandem mass spectrometry: photon activation with vacuum ultra-violet brings unique information and enables definitive structure assignment. *Anal. Chim. Acta* **807**, 84-95
37. Vagin, A., and Lebedev, A. (2015) MoRDa , an automatic molecular replacement pipeline. *Acta Crystallogr. A* **71**, s19-s19
38. Keitel, T., Meldgaard, M., and Heinemann, U. (1994) Cation binding to a *Bacillus* (1,3-1,4)-beta-glucanase. Geometry, affinity and effect on protein stability. *Eur. J. Biochem.* **222**, 203-214
39. Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Stahlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T., and Jones, T. A. (1997) The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes. *J. Mol. Biol.* **272**, 383-397
40. Matard-Mann, M., Bernard, T., Leroux, C., Barbeyron, T., Larocque, R., Prechoux, A., Jeudy, A., Jam, M., Nyvall Collen, P., Michel, G., and Czjzek, M. (2017) Structural insights into marine carbohydrate degradation by family GH16 kappa-carrageenases. *J. Biol. Chem.* **292**, 19919-19934
41. Keitel, T., Simon, O., Borriss, R., and Heinemann, U. (1993) Molecular and active-site structure of a *Bacillus* 1,3-1,4-beta-glucanase. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5287-5291
42. McGregor, N., Yin, V., Tung, C. C., Van Petegem, F., and Brumer, H. (2017) Crystallographic insight into the evolutionary origins of xyloglucan endotransglycosylases and endohydrolases. *Plant J.* **89**, 651-670
43. Johansson, P., Brumer, H., 3rd, Baumann, M. J., Kallas, A. M., Henriksson, H., Denman, S. E., Teeri, T. T., and Jones, T. A. (2004) Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. *Plant Cell* **16**, 874-886
44. Mark, P., Baumann, M. J., Eklof, J. M., Gullfot, F., Michel, G., Kallas, A. M., Teeri, T. T., Brumer, H., and Czjzek, M. (2009) Analysis of nasturtium TmNXG1 complexes by crystallography and molecular dynamics provides detailed insight into substrate recognition by family GH16 xyloglucan endo-transglycosylases and endo-hydrolases. *Proteins* **75**, 820-836
45. Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., and Potin, P. (2001) *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of. *Int. J. Syst. Evol. Microbiol.* **51**, 985-997
46. Potin, P., Patier, P., Floc'h, J. Y., Yvin, J. C., Rochas, C., and Kloareg, B. (1992) Chemical characterization of cell-wall polysaccharides from tank-cultivated and wild plants of

Delesseria sanguinea (Hudson) Lamouroux (Ceramiales, Delesseriaceae): Culture patterns and potent anticoagulant activity. *J. Appl. Phycol.* **4**, 119-128
47. Grunewald, N., and Alban, S. (2009) Optimized and standardized isolation and structural characterization of anti-inflammatory sulfated polysaccharides from the red alga *Delesseria sanguinea* (Hudson) Lamouroux (Ceramiales, Delesseriaceae). *Biomacromolecules* **10**, 2998-3008
48. Li, J., Jiang, Z., Wu, H., Liang, Y., Zhang, Y., and Liu, J. (2010) Enzyme–polysaccharide interaction and its influence on enzyme activity and stability. *Carbohydr. Polym.* **82**, 160-166
49. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402
50. Katoh, K., and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780
51. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725-2729
52. Groisillier, A., Herve, C., Jeudy, A., Rebuffet, E., Pluchon, P. F., Chevolot, Y., Flament, D., Geslin, C., Morgado, I. M., Power, D., Branno, M., Moreau, H., Michel, G., Boyen, C., and Czjzek, M. (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
53. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein. Expr. Purif.* **41**, 207-234
54. 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. Detection of picomolar quantities by an imaging system based on a cooled charge-coupled device. *Biochem. J.* **270**, 705-713
55. Starr, C. M., Masada, R. I., Hague, C., Skop, E., and Klock, J. C. (1996) Fluorophore-assisted carbohydrate electrophoresis in the separation, analysis, and sequencing of carbohydrates. *J. Chromatogr. A* **720**, 295-321
56. Herve, C., Simeon, A., Jam, M., Cassin, A., Johnson, K. L., Salmean, A. A., Willats, W. G., Doblin, M. S., Bacic, A., and Kloareg, B. (2016) Arabinogalactan proteins have deep roots in eukaryotes: identification of genes and epitopes in brown algae and their role in *Fucus serratus* embryo development. *New Phytol.* **209**, 1428-1441
57. Kidby, D. K., and Davidson, D. J. (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.* **55**, 321-325
58. Giuliani, A., Jamme, F., Rouam, V., Wien, F., Giorgetta, J. L., Lagarde, B., Chubar, O., Bac, S., Yao, I., Rey, S., Herbeaux, C., Marlats, J. L., Zerbib, D., Polack, F., and Refregiers, M. (2009) DISCO: a low-energy multipurpose beamline at synchrotron SOLEIL. *J. Synchrotron Radiat.* **16**, 835-841
59. Domon, B., and Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J* **5**, 397–409
60. Niedermeyer, T. H., and Strohalm, M. (2012) mMass as a software tool for the annotation of cyclic peptide tandem mass spectra. *PLoS One* **7**, e44913
61. Kabsch, W. (2010) Xds. *Acta Crystallogr. D* **66**, 125-132
62. Weiss, M. S., Sicker, T., Djinovic-Carugo, K., and Hilgenfeld, R. (2001) On the routine use of soft X-rays in macromolecular crystallography. *Acta Crystallogr. D* **57**, 689-695
63. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D* **66**, 486-501
64. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60**, 2184-2195
65. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-324

Table 1: Data collection and refinement statistics for the crystal structure of *Zg*AgaC

	ZgAgaC
Data collection	
Space group	P 2_1 2_1 2
Unit cell	
a, *b*, *c* (Å)	60.6, 101.5, 46.69
α, β, γ (°)	90, 90, 90
Resolution(Å)	42.42 - 1.3 (1.346 - 1.3)
R_{merge}	0.05 (0.39)
I/σI	11.66 (2.29)
CC(1/2)	0.99 (0.84)
Completeness (%)	98.71 (89.68)
Redundancy	3.74 (3.12)
Refinement statistics	
Resolution (Å)	1.3
Unique reflexion	70662(6310)
R/R_{free}	0.1493/0.1689
Atoms numbers	2535
Proteins	2230
Ligand/ion	21
Water	284
B factor	
Proteins	12.57
Ligand/ion	24.66
Water	25.07
R. m. s. deviation	
Bond lengths (Å)	0.028
Bond angles (Å)	2.15
PDB code	6HY3

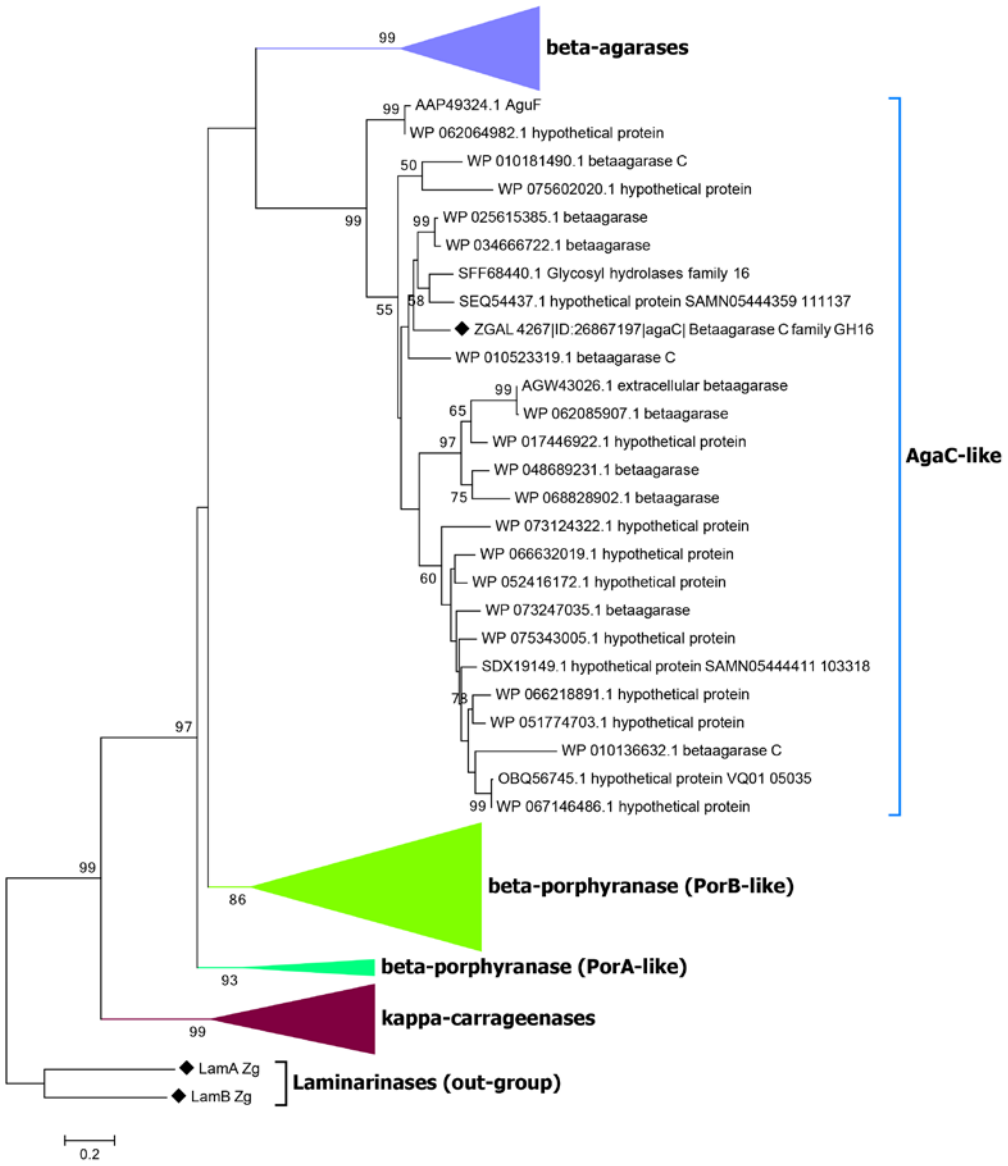


Figure 1. Phylogenetic tree of the galactanases of the GH16 family. The phylogenetic tree was generated using the Maximum Likelihood approach with the program Mega6 (51). Bootstrap numbers are indicated. The tree was rooted by the laminarinases *Zg*LamA and *Zg*LamB from *Z. galactanivorans*. The clades of the κ-carrageenases, the *Zg*PorA homologues, the ZgPorB homologues and the classical β-agarases have been collapsed. The sequence of *Zg*AgaC from *Z. galactanivorans* is marked with a black diamond.

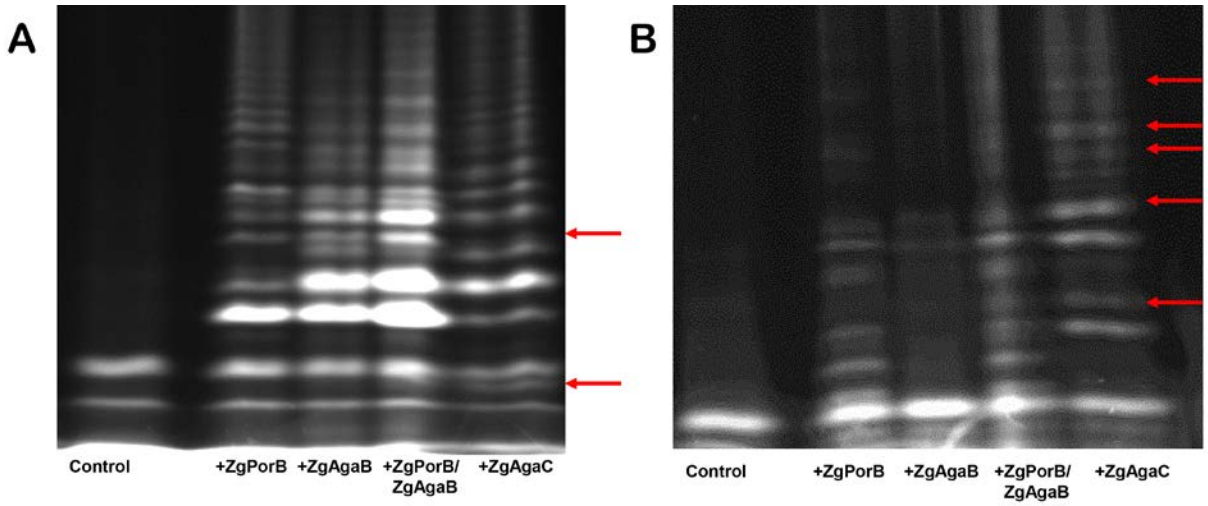


Figure 2. Degradation pattern of *Zg*PorB, *Zg*AgaB and *Zg*AgaC on the red algae *Polysiphonia elongata* (A) and *Osmundea pinnatifida* (B). Ground algae were resuspended with milliQ water (control) and incubated with the different GH16 galactanases. The degradation products were coupled with 8-aminonaphtalene-1,3,6-trisulfonate (ANTS) and analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE). The major differences in oligosaccharide banding patterns for *Zg*AgaC compared to *Zg*AgaB and *Zg*PorB are reported by red arrows.

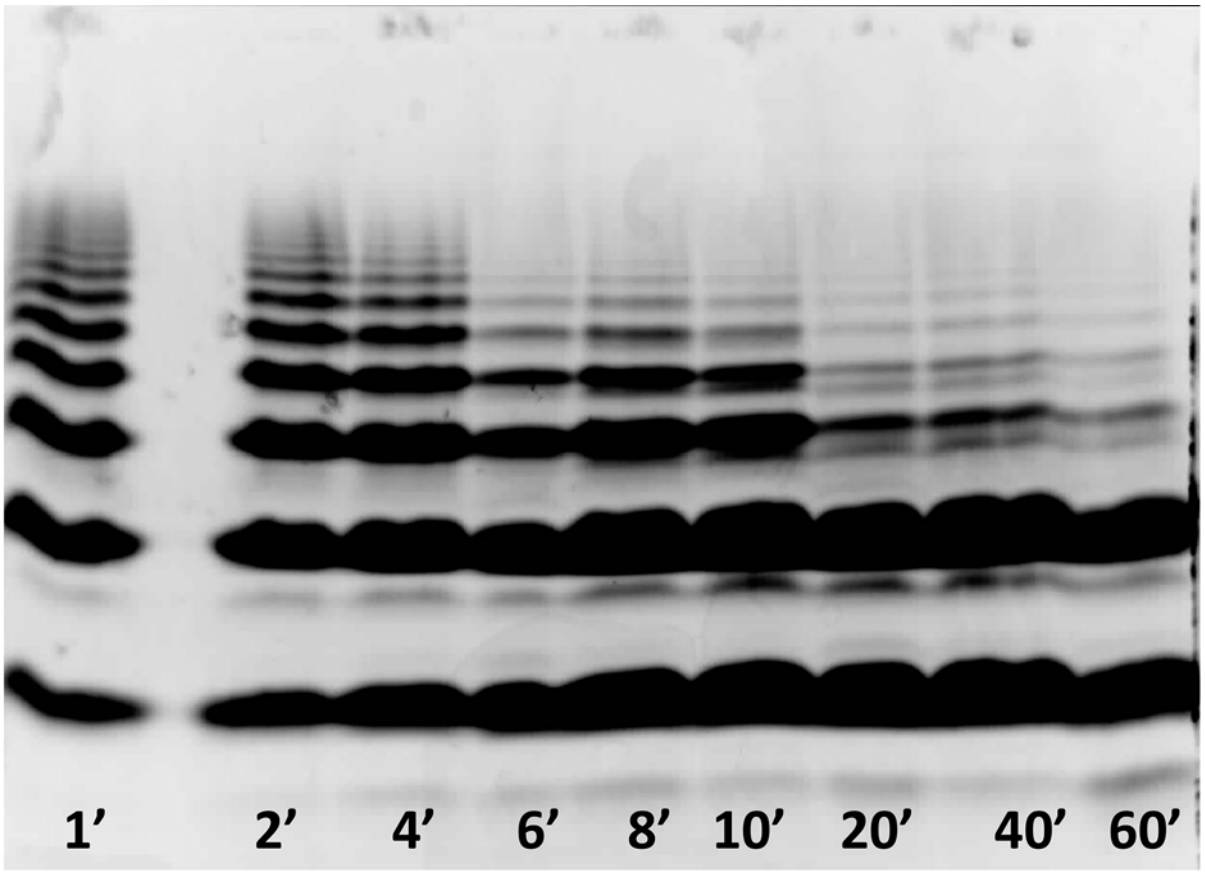


Figure 3. Hydrolysis of agarose by *Zg*AgaC monitored by fluorophore-assisted carbohydrate electrophoresis (FACE). The reaction products were labelled with the ANTS fluorophore and migrated onto a 37% polyacrylamide gel. The time points are shown in minutes.

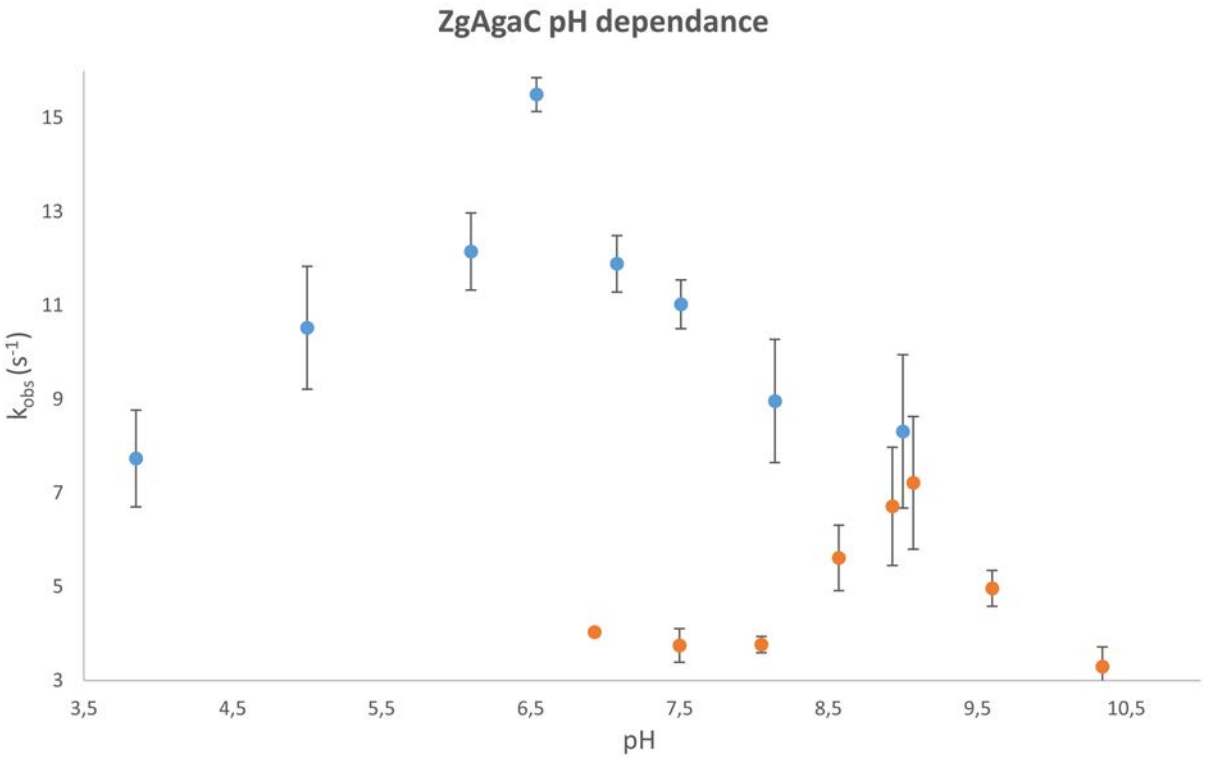


Figure 4. pH dependence of *Zg*AgaC. The pH dependence of *Zg*AgaC was determined for the neutral polysaccharide agarose (0.4% (w/v), blue dot) and for the sulfated agar from *Osmundea pinnatifida* (0.4% (w/v), orange dot). The rate of the reactions are expressed in s^{-1} (k_{obs}).

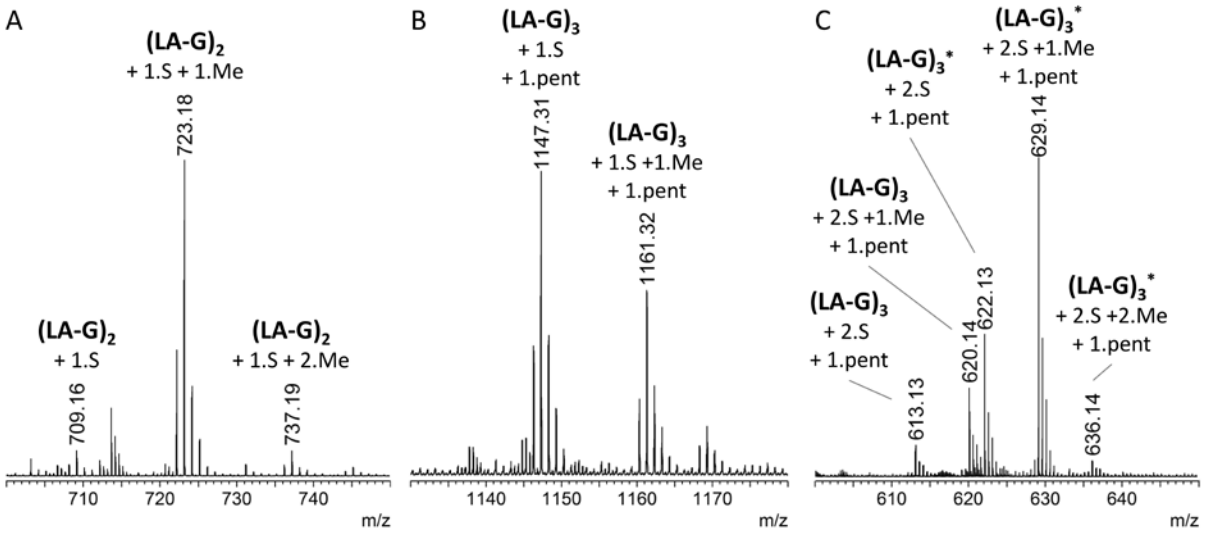


Figure 5: ESI-MS measurements of the three fractions (OP44, A; OP36, B; and OP30, C) which contained the smallest oligosaccharides identified by FACE gels. Annotations were deduced from the exact mass measurements. The black star indicate non-common species with one 3,6-anhydro-L-galactose replaced by a L-galactose in the regular moieties.

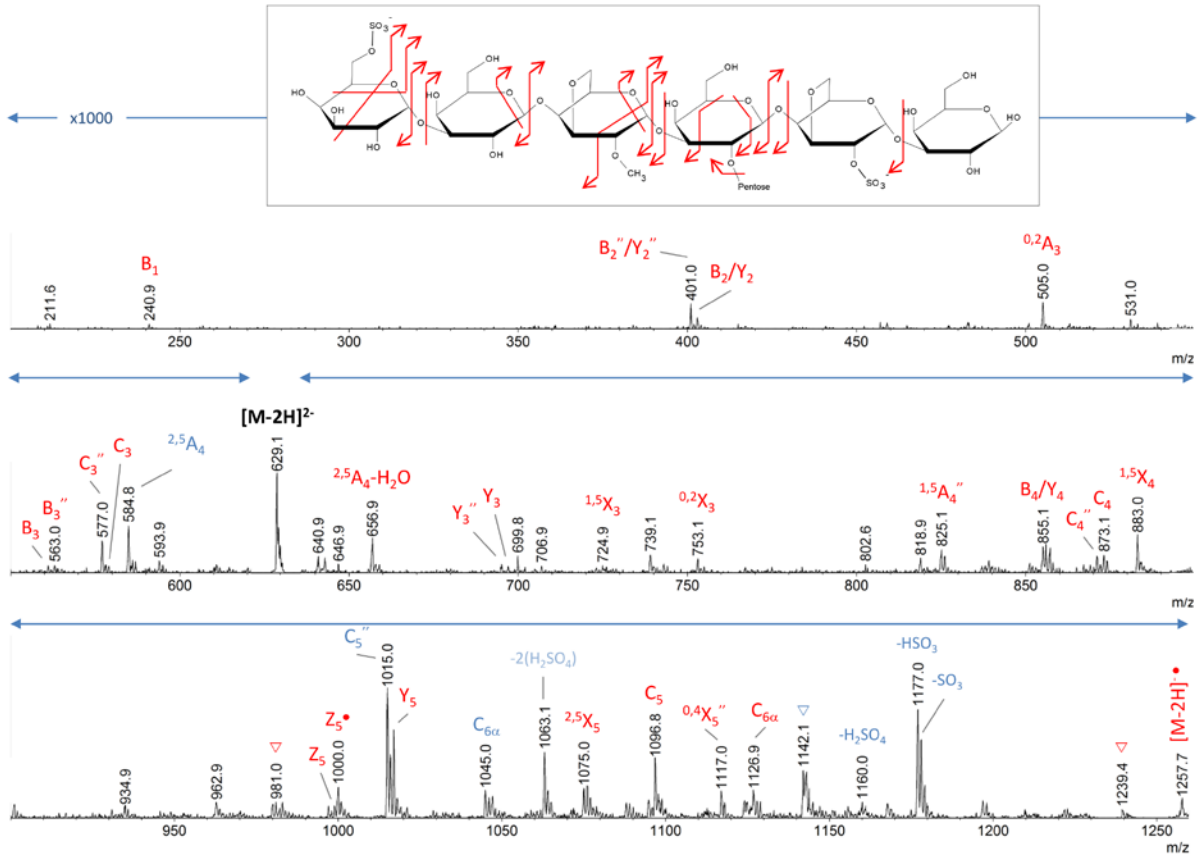


Figure 6: XUV-DPI tandem MS spectrum of the doubly charged species isolated at *m/z* 629.1 as a $[M-2.H]^{2-}$ ion. The spectrum was recorded following 18 eV-photon irradiation for 500 ms. For a better readability, the mass range was split in three. Red: fully sulfated fragments. Blue: fragments with one sulfate loss. Triangle: water losses. Signal was accumulated over 2 minutes.

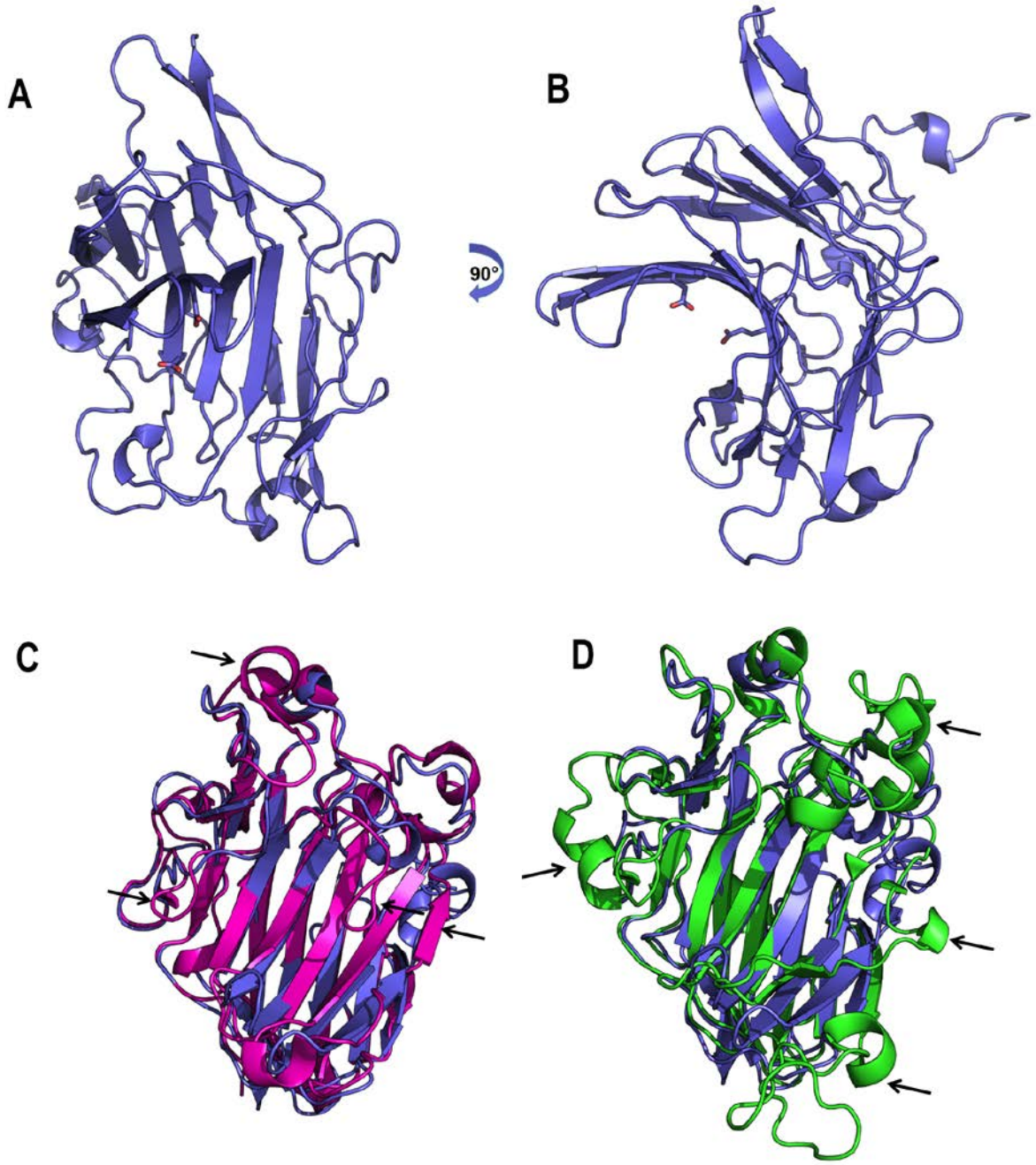


Figure 7. Comparison of the structure of *Zg*AgaC with the structures of the β-porphyranase *Zg*PorA and the β-agarase *Zg*AgaB. (A) *Zg*AgaC adopts a β-jelly-roll fold. The nucleophile and acid/base catalysts are shown in sticks. (B) View at 90° of the structure of *Zg*AgaC. (C) Superimposition of the structures of *Zg*AgaC (blue) and *Zg*PorA (pink). (D) Superimposition of the structures of *Zg*AgaC (blue) and *Zg*AgaB (green). The major differences between the structures are reported by arrow.

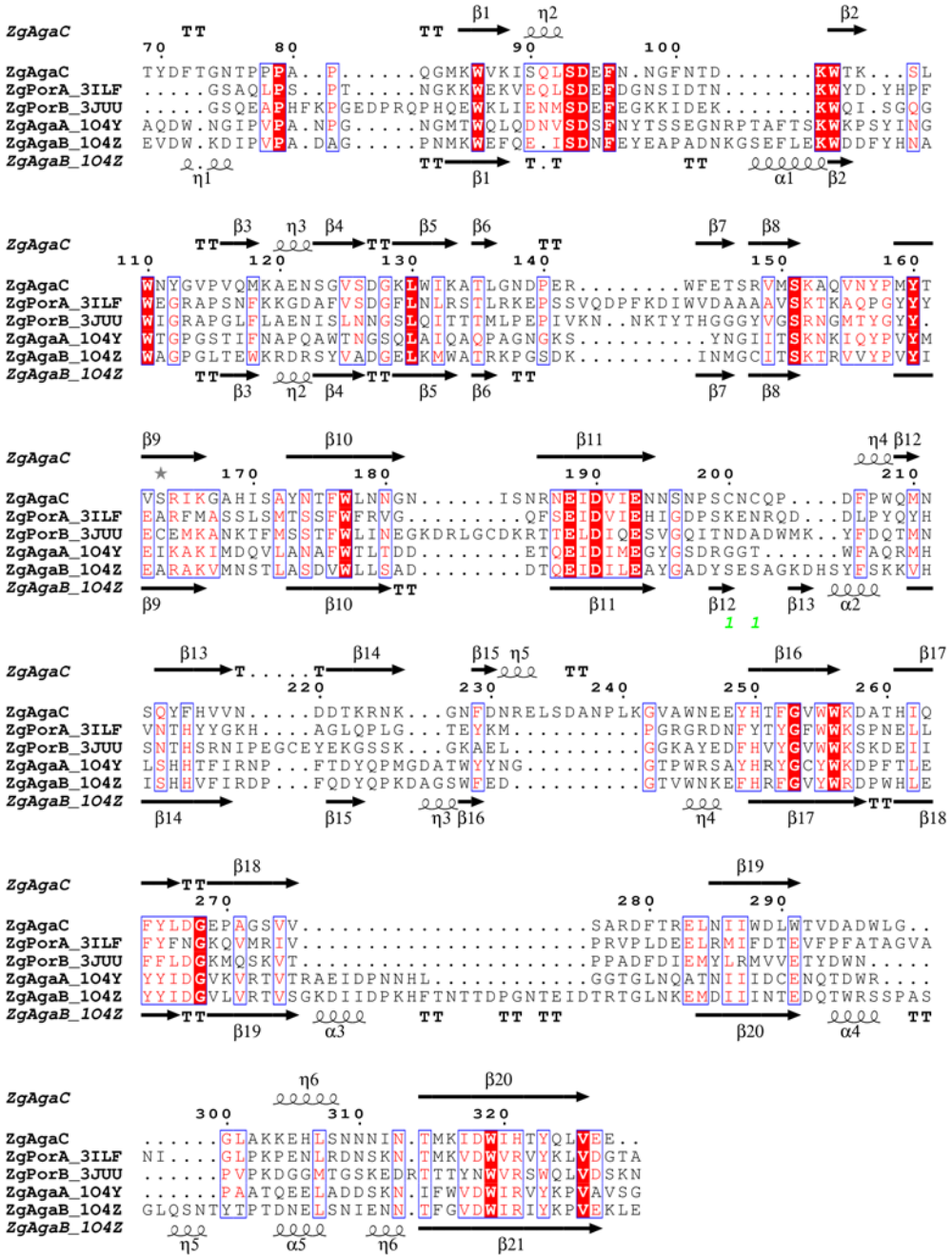


Figure 8. Structure-based sequence alignment of *Zg*AgaC with characterized β-agarases and β-porphyranases. Alpha helices and beta strands are represented as helix and arrows respectively, and beta turn are maked with TT. This sequence alignment was created using the following sequences from the Protein Data Bank; the β-agarases *Zg*AgaA (1O4Y) and *Zg*AgaB (1O4Z) and the β-porphyranases *Zg*PorA (3ILF) and *Zg*PorB (3JUU). All these GH16 enzymes originate from *Zobellia galactanivorans*. Dark shaded boxes enclose invariant positions, and light shaded boxes show positions with similar residues. The figure was created with ESPript 3.0 (65).

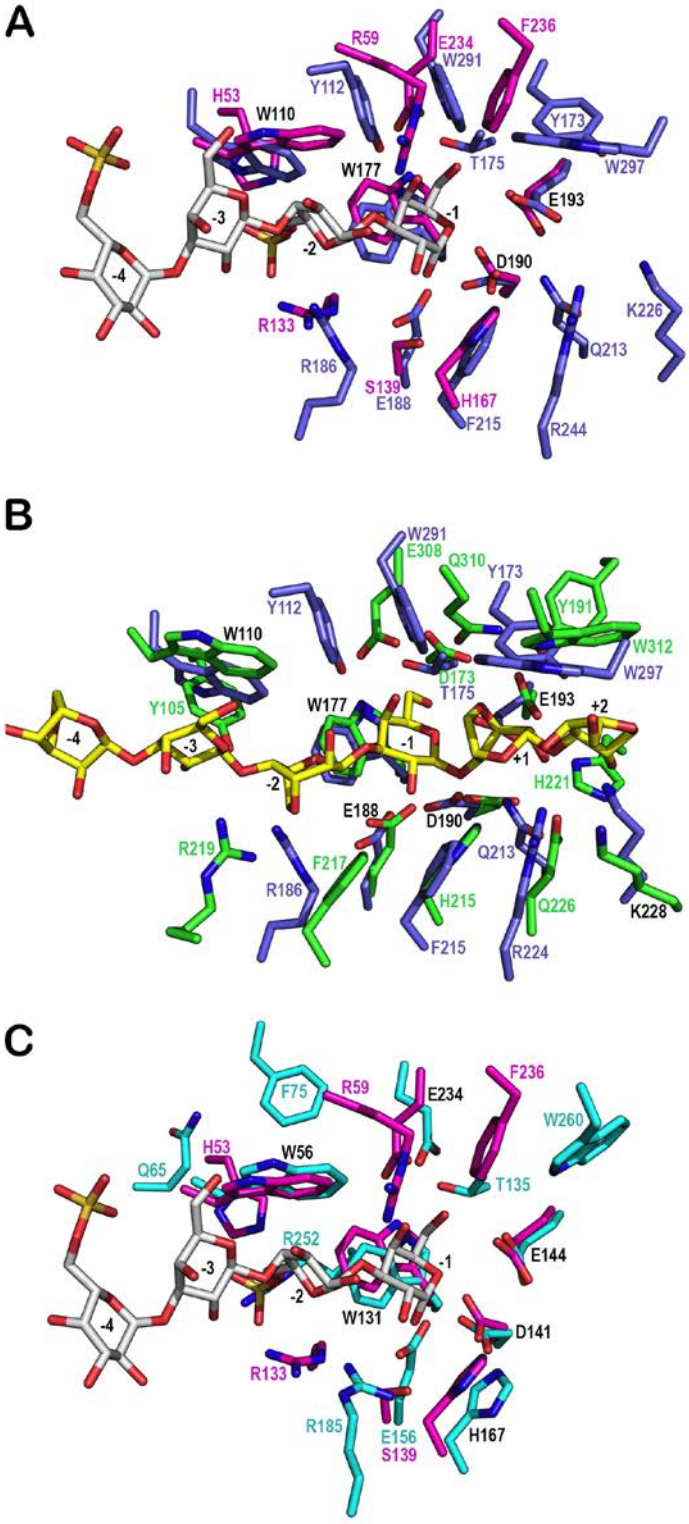


Figure 9. Comparison of the active site of *Zg*AgaC, of the β-porphyranase *Zg*PorA and of the β-agarase *Zg*AgaB. (A) Superimposition of *Zg*PorA (PDB: 3ILF, colored in pink) in complex with a porphyran tetrasaccharide (colored in grey) and *Zg*AgaC (PDB: 6HY3, colored in blue). (B) Superimposition of *Zg*AgaB (PDB: 4ATF, colored in green) in complex with an agarose octasaccharide (colored in yellow, six moieties are shown) and *Zg*AgaC (colored in blue). (C) Superimposition of ZgPorA (colored in pink) in complex with a porphyran tetrasaccharide (colored in grey) and *Zg*PorB (PDB: 3JUU, colored in cyan).

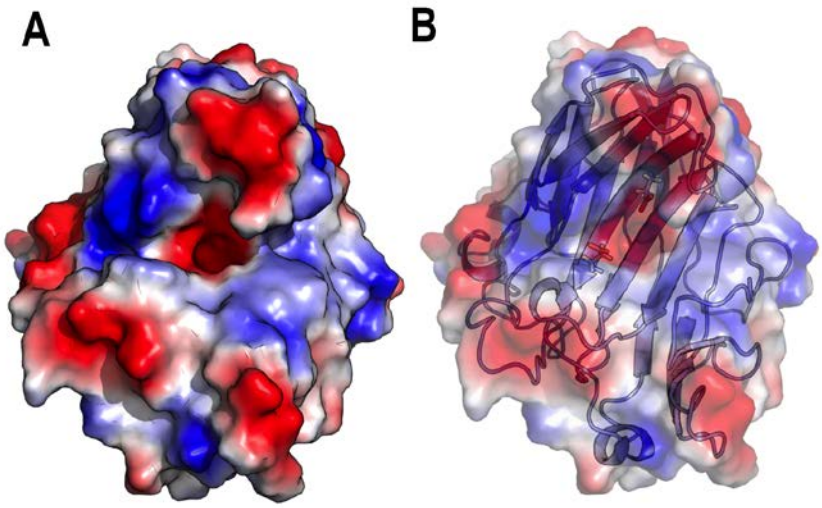


Figure 10. Molecular surface of *Zg*AgaC. (A) Molecular surface of *Zg*AgaC colored according to electrostatic potential ranging from deep blue, +, to red, -. (B) Structure of *Zg*AgaC represented in cartoon, with its electrostatic surface shown in transparent view.

Chapitre IV : Caractérisation de la sulfatase *Zg*4660

IV. 1 Introduction

Dans le but d'explorer la dégradation de l'agar par *Zobellia galactinovorans* Dsij^T, il est important d'identifier les enzymes responsables de l'élimination des substitutions présentes sur ce polysaccharide. Parmi ces substitutions il y a les sulfatations qui sont éliminées par des sulfatases. A ce jour, aucune sulfatase agissant sur les agars n'a été caractérisée. Néanmoins, des éléments d'information permettant d'identifier de potentiels candidats étaient disponibles. Notamment, un opéron, appelé PUL49 (Barbeyron et al., 2016b), a été prédit comme agissant sur l'agar chez *Z. galactanivorans* (**Figure 33**). En effet, la transcription de la majorité des gènes de cet opéron est induite en présence d'agar dans le milieu (14 gènes sur 17) (Thomas et al, 2017). De plus, il contient le gène *zgahgA* codant, une 1,3-α-3,6-anhydro-L-galactosidase permettant de cliver les liaisons α-1,3 à partir des extrémités non réductrices des oligo néo-agaroses (Rebuffet et al., 2011). Cet opéron contient dix-sept gènes codant, notamment, pour un régulateur transcriptionnel (*Zg*4664), des transporteurs localisés dans la membrane cytoplasmique (*Zg*4654 et *Zg*4656), une épimérase (*Zg*4657), mais aussi une sulfatase de la famille S1_19 (*Zg*4660). Pour rappel, à notre connaissance, seules les sulfatases de la famille S1 ou sulfatases formylglycine dépendante peuvent désulfater des carbohydrates.

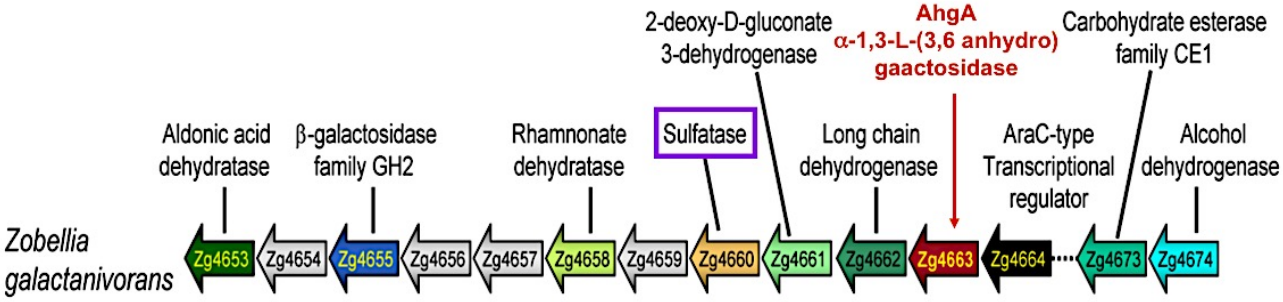


Figure 33 : Présentation du PUL 49 de *Z. galactanivorans* (Tiré de Rebuffet et al., 2011)

Ce PUL contient 17 gènes (de *zg*4653 à *zg*4669), les deux autres gènes présentés sur la figure zg4673 et zg4674 ne sont pas inclus dans le PUL49 mais peuvent jouer un rôle synergétique avec les protéines de l'opéron. Parmi les gènes de cette opéron putatif, deux transporteurs sont également présents : *Zg*4656 (symporteur de L-rhamnose) et *Zg*4654 (transporteur de D-galactonate). Un homologue de *Vej*ABG, correspondant à une GH2 permettant de libérer du D-galactose en clivant la liaison β(1,4), est également présent dans ce PUL, il s'agit de *Zg*4655 (Lee et al., 2014). D'autres enzymes sont présentes comme une épimérase (*Zg*4657), une lactate déshydrogénase (*Zg*4659), indiquant que ce PUL est plutôt impliqué dans les étapes terminale de la dégradation de l'agar et à l'insertion des produits de dégradation dans le métabolisme central du carbone, que dans les premières étapes du catabolisme de l'agar.

La présence de *zgahgA* dans cet opéron ainsi que les données de transcriptomiques indique qu'il est très probable que *Zg*4660 soit une sulfatase active sur les agars. Le problème générique rencontré lors de toutes les études portant sur les sulfatases est de réussir à les obtenir sous forme active quand elles sont exprimées en système hétérologue (par *Escherichia coli* par exemple). En effet, pour être actives les sulfatases de la famille S1 nécessitent d'être modifiées de façon co-traductionnelle. Cette modification, assurée par la *formylglycine generating enzyme* (FGE) chez *Z. galactanivorans*, consiste en l'oxydation d'une sérine ou d'une cystéine, présente dans le site catalytique de cette famille de sulfatases, en formylglycine. La contrainte expérimentale est que l'équivalent de la FGE ainsi que les systèmes de régulation chez *E. coli*, ne sont pas forcément adaptés à des protéines hétérologues et ne permettent donc pas de générer la modification de façon efficace.

Des tests de clonage et d'expression de différentes sulfatases de *Z. galactanivorans* chez *E. coli* avaient été préalablement réalisés par le Dr. T. Barbeyron et R. Larocque (communication personnelle). Ces tests ont permis d'obtenir la protéine *Zg*4660 exprimée soluble chez *E. coli*, mais l'enzyme était très peu active sur le substrat artificiel 4-methylumbelliferyl-sulfate (MUS), suggérant une conversion post-traductionnelle en formylglycine de cette enzyme très peu efficace chez *E. coli*.

Les défis de ce projet étaient donc d'obtenir (i) une sulfatase recombinante active et (ii) et de définir son substrat naturel.

IV. 2 Résultats

IV. 2. A Contexte génétique de la sulfatase *Zg*4660

Comme cela a été présenté dans l'introduction, la sulfatase *Zg*4660 fait partie d'un opéron comportant dix-sept autres gènes dont la transcription est induite lorsque de l'agar est présent dans le milieu de culture de *Z. galactanivorans* (Thomas et al., 2017). La protéine *Zg*4660 présente une portion N-terminale de 24 acides aminés prédite comme étant un peptide signal d'adressage périplasmique. Après clivage de ce peptide signal, la protéine est composée de 438 résidus. *Zg*4660 a été clonée de façon à comporter une étiquette de six histidines en position N-terminale, produisant une protéine recombinante de 454 résidus avec une masse moléculaire de 51,2 kDa et un point isoélectrique calculé de 6,96. D'après la base de données SulftAtlas (Barbeyron et al., 2016a), *Zg*4660 fait partie de la famille 1 (sulfatase formylglycine dépendante), sous-famille S1_19. Le motif catalytique consensus (C/S)xPxR est ici retrouvé sous la forme : ^{74}CGPSR^{78} (numérotation de la protéine native). La sous-famille S1_19 est connue comme contenant les activités endo-4S-kappa carraghénase et endo-4S-iota carraghénase (Préchoux et al., 2013; Préchoux et al., 2016). Un de ces membres a d'ailleurs été récemment caractérisé chez *Z. galactanivorans* : il s'agit de l'endo-4S-iota carraghénase (*Zg*3145 ou *Zg*CgsA) présenté dans l'introduction (Ficko-Blean et al., 2017). Un alignement réalisé avec ClustalW révèle, entre les deux enzymes, une identité de séquence de 34,3 %.

IV. 2. B Production chez *Escherichia coli* d'une sulfatase recombinante active et caractérisation cinétique avec un substrat artificielle

Après expression de la protéine *Zg*4660 recombinante dans *E. coli*, une seule étape de purification par chromatographie d'affinité permet d'obtenir une protéine très largement majoritaire qui, sur gel de polyacrylamide dénaturant, migre à environ 50 KDa, ce qui est cohérent avec la masse moléculaire théorique (51,2 kDa). Il est à noter que lorsque la protéine *Zg*4660 est produite dans un milieu auto-inductible à la place du milieu LB, nous observons une augmentation significative du nombre de protéines contaminantes. Un test de stabilité thermique de la protéine *Zg*4660, réalisé par

diffusion dynamique de la lumière (DLS, Malvern Instrument), a révélé qu'elle était stable jusqu'à 40°C (**Figure 34**).

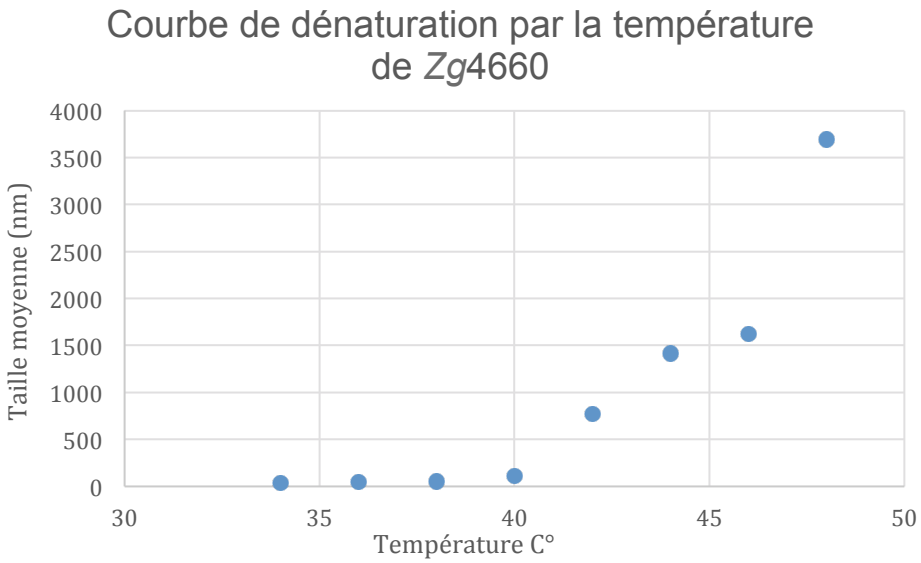


Figure 34 : **Evaluation de la thermostabilité de *Zg*4660 par DLS**
Cette mesure réalisée par DLS (Malvern Instrument) montre que nous avons une augmentation nette du rayon hydrodynamique de la protéine dès 42°C.

Une difficulté majeure lors de la production des sulfatases sous forme recombinante est d'obtenir la modification co-traductionnelle. Comme je l'ai signalé au début du chapitre, la sulfatase *Zg*4660 recombinante surexprimée chez *E. coli* est très peu active sur le substrat MUS. A titre de comparaison, à concentration égale il faut moins de cinq minutes à la sulfatase *Zg*3515 (voir chapitre V ; k_{cat}/K_M= 83,54 M^{-1} s^{-1}) pour obtenir une fluorescence maximale lors d'une incubation avec de la MUS tandis que pour *Zg*4660 il faut plus d'une journée (k_{cat}/K_M= 41,55 M^{-1} s^{-1}). Deux hypothèses peuvent expliquer cette faible activité : soit que la MUFS n'est pas un substrat reconnu par cette enzyme, soit que la conversion en formylglycine du site actif est très peu efficace pour cette protéine chez *E. coli*.

Pour vérifier cette deuxième hypothèse, la première idée a été de co-exprimer chez *E. coli* le gène de la sulfatase *Zg*4660 en présence du gène codant pour l'enzyme FGE de *Z. galactanivorans*. Avant ma thèse, le gène *zgfge* de *Z. galactanivorans* avait été cloné dans le plasmide pET-Duet qui possède un gène de résistance à la kanamycine et une origine de réplication compatible avec les vecteurs d'expression de la série pET (Dr. E. Ficko-Blean, communication personnelle). J'ai donc pu utiliser ce vecteur pour tester mon hypothèse. Les cinétiques avec de la MUS ont montré une augmentation significative de

l'activité de 67 fois par rapport à la protéine seule (k_{cat}/K_M=2802,33 $M^{-1}.s^{-1}$), validant ainsi l'hypothèse d'une conversion incomplète de la cystéine catalytique en formylglycine chez *E. coli*. Ce premier résultat nous a conforté dans l'idée que d'autres améliorations étaient possibles. Une seconde amélioration a donc été d'induire des protéines chaperonnes pour ralentir le processus de repliement et ainsi laisser le temps à la FGE de réaliser la modification. A partir de la même souche transformée par les deux vecteurs d'expression, j'ai réalisé une culture au cours de laquelle de l'éthanol a été ajouté à la concentration finale de 3 % (v/v) avant une incubation à 4°C. L'hypothèse est que ce double choc, thermique et alcoolique, permettrait probablement de freiner la traduction et de produire des chaperonnes susceptibles de maintenir, au cours de la traduction, la sulfatase sous forme non repliée, permettant à la FGE de réaliser la modification de l'acide aminé catalytique. Le gel illustrant la purification de la sulfatase *Zg*4660 (**Figure 35**) à partir de ces cultures met en évidence une seconde protéine qui migre entre 37,5 kDa et 25 KDa, masse moléculaire correspondant à celle de la chaperonne DnaJ de *E. coli* (De Marco, 2007). Cette dernière co-élue probablement avec la sulfatase lors de la purification.

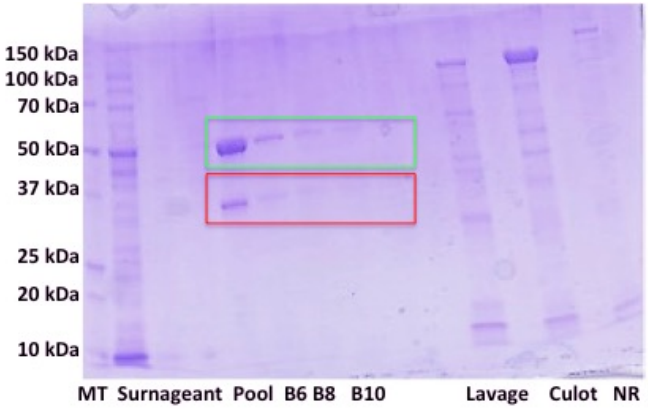


Figure 35 : Analyse par gel de la purification de *Zg*4660 co-exprimé avec *Zg*FGE suite à un choc thermique et alcoolique.

La bande aux environ de 50 kDa correspond à *Zg*4660 (encadré vert), tandis que la bande d'environ 30 kDa (sous la bande de 37 kDa du marqueur de taille et dans l'encadré rouge) est probablement une protéine chaperonne. **MT** : Marqueur de taille ; **Surnageant** : extrait protéique soluble après lyse cellulaire **; Pool** : toutes les fractions collectées correspondant au pic de sortie de la colonne Histrap ; **B6, B8, B10** : fait référence aux noms des fractions après l'étape de Nickel ; **Lavage** : protéines non fixées sur la colonne Histrap **; Culot** : extrait cellulaire insoluble ; **NR** : fraction non retenue sur la colonne Histrap lors de l'injection de l'extrait protéique soluble. Gel de polyacrylamide dénaturant 12%, révélé au bleu de Coomasie.

Les résultats des cinétiques ont montré que ce dernier protocole permettait d'obtenir une enzyme 1251 fois plus active que la sulfatase exprimée seule (**Tableau 5**). Ce protocole a été utilisé par la suite pour l'étude de l'activité sulfatase.

Traitements / Paramètres cinétiques	*Expression seul*	*Zg*4660+ Choc éthanol	*Zg*4660+ *Zg*FGE	*Zg*4660+ *Zg*FGE Choc éthanol
Efficacité catalytique (k_{cat}/K_M) $s^{-1}.M^{-1}$	41,55	10,46	2802,33	52017,72
Augmentation de l'activité comparé à l'expression seul	0	- 3,97	67,44	1251,93
Rendement (mg de protéine par litre de culture)	400	36	81	20,5

Tableau 5 : Activité catalytique de *Zg*4660 en fonction du protocole employé pour la produire

IV. 2. C Criblage de l'activité sulfatase sur des oligosaccharides issus de diverses algues rouges agarophytes

Pour tenter d'identifier le substrat naturel de cette sulfatase, nous avons mis au point une nouvelle technique de criblage. Dans le chapitre III sur *Zg*AgaC, il a été présenté la technique d'électrophorèse des carbohydrates assistées par fluorophore (FACE). C'est cette technique que nous avons adaptée en utilisant un fluorophore neutre, postulant que seuls les oligosaccharides chargés migreront et seront visualisés. Lors de l'action de la sulfatase, une charge négative est éliminée, entrainant une modification du profil de migration des oligosaccharides substrats. Nous avons donc utilisé un marquage à l'AMAC pour tous nos tests et visualisé les activités sulfatases avec des gels FACE (**Figure 36**). Pour réaliser les tests, nous avons repris la collection d'algues rouges générée lors des études réalisées sur la recherche de la spécificité de *Zg*AgaC.

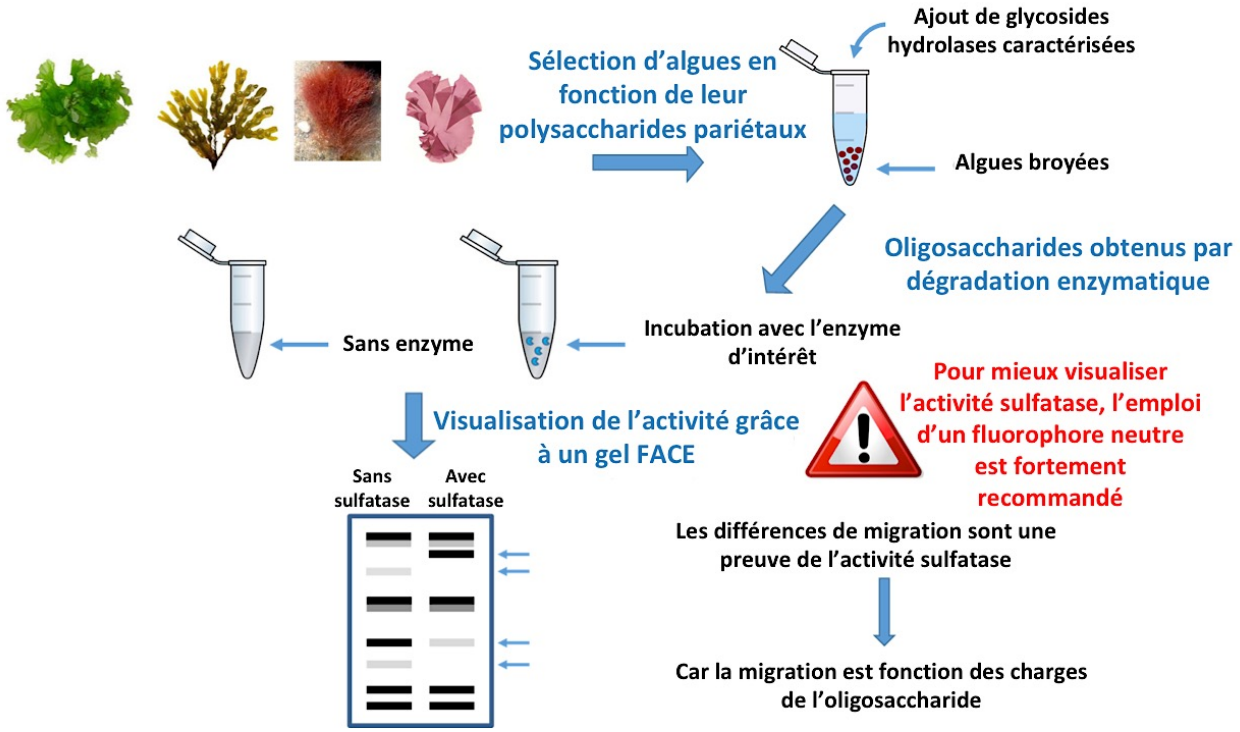


Figure 36 : Stratégie de criblage pour la détection de l'activité sulfatase.

La sulfatase *Zg*4660 a été testée sur dix algues agarophytes (cf **Tableau 5**), chacune ayant été traitée avec trois glycoside hydrolases différentes (*Zg*AgaB, *Zg*PorB, *Zg*AgaC). Il en résulte 30 fractions oligosaccharidiques différentes qui ont été testées avec *Zg*4660.

GH / Algue	*Zg*AgaB	*Zg*PorB	*Zg*AgaC
Ceramium rubrum	-	+	-
Chondria dasyphilla	+	+	-
Deleisseria sanguinea	-	-	-
Dumontia contorta	-	-	-
Osmundea pinnatifida	-	-	-
Polysiphonia brodei	-	-	-
P. elongata	++	-	-
P. simulans	-	-	-
Rhodomella sp.	-	-	-
Vertebrata lanosa	-	-	-

Tableau 6 : Activité sulfatase de *Zg*4660 sur différents oligosaccharides provenant d'algues rouges agarophytes

A gauche sont indiquées les algues dont proviennent es oligosaccharides, en haut est indiqué la GH16 utilisée pour générer les oligosaccharides. Un (-) indique que la sulfatase est inactive, un (+) indique que le traitement à la sulfatase perturbe la migration d'un oligosaccharide minoritaire. Deux (++) signifie que le traitement à la sulfatase perturbe de façon importante le profil électrophorétique des oligosaccharides.

Ce premier criblage a révélé que *Zg*4660 était particulièrement active sur les oligosaccharides générés à partir de *Chondria dasyphylla* et *Polysiphonia elongata* (**Tableau 6** et **Figures 37** et **38**). L'activité la plus importante a été détectée sur les oligosaccharides générés par *Zg*AgaB à partir de *P. elongata*. La double digestion *Zg*AgaB/*Zg*PorB n'apporte pas de différence significative par rapport à la digestion réalisée par une seule enzyme. La **Figure 38** qui compare les marquages ANTS (fluorophore chargé) et AMAC, montre que le marquage à l'ANTS ne permet pas de visualiser l'impact de la sulfatase. L'une des raisons est l'augmentation du bruit de fond due à la migration des oligosaccharides neutres. L'autre raison est que l'ANTS pourrait compenser (donc masquer) la perte de charge due à l'action de la sulfatase. Une fois le criblage réalisé, il était nécessaire de définir les conditions les moins dénaturantes possibles lors de l'extraction en grande quantité de l'agar de *P. elongata* pour pouvoir, par la suite, caractériser le substrat de *Zg*4660.

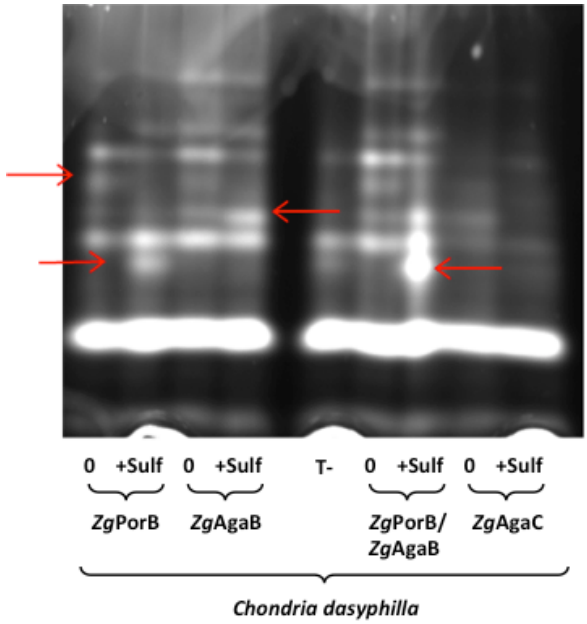


Figure 37: Activité sulfatase de *Zg*4660 sur les oligosaccharides générés à partir de *Chondria dasyphilla*

Les différentes pistes correspondent aux différents oligosaccharides générés par les glycosides hydrolases suivantes : *Zg*AgaB, *Zg*AgaC, *Zg*PorB et le mélange *Zg*AgaB/*Zg*PorB. Chacune de ces digestions a été incuber dans du tampon sans (0) ou avec la sulfatase active (+Sulf). La piste T- correspond à l'extrait pariétal non traité par les glycosides hydrolases. Sur le gel, les flèches indiquent les oligosaccharides dont la mobilité électrophorétique est modifiée suite à l'action de la sulfatase. Le marquage a été réalisé avec de l'AMAC.

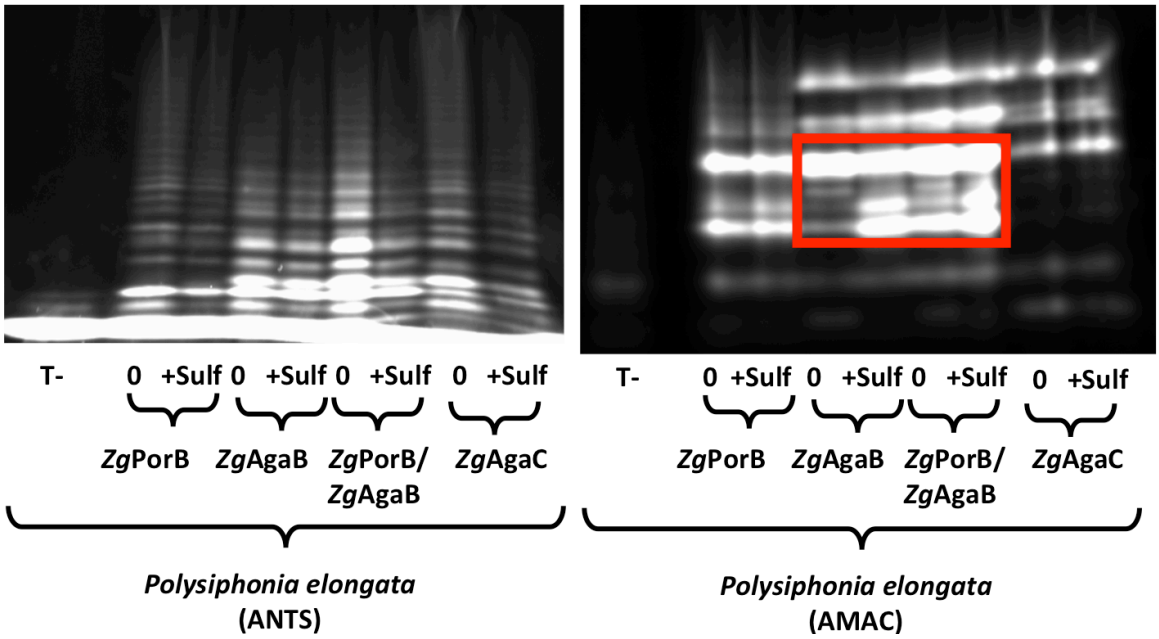


Figure 38 : Activité sulfatase de *Zg*4660 sur les oligosaccharides générés à partir de *Polysiphonia elongata*

Les différentes pistes correspondent aux différents oligosaccharides générés par les glycosides hydrolases suivantes : *Zg*AgaB, *Zg*AgaC, *Zg*PorB et le mélange *Zg*AgaB/*Zg*PorB. Chacune de ces digestions a été incuber dans du tampon sans (0) ou avec la sulfatase active (+Sulf). La piste T- correspond à l'extrait pariétal non traité par les glycosides hydrolases. Le fluorophore ayant permis le marquage des oligosaccharides est indiqué sous le gel. L'encadré rouge montre les pistes où se situent les différences entre le contrôle et l'incubation avec la sulfatase.

IV. 2. D Extraction de l'agar de *Polysiphonia elongata*

Les algues, une fois séchées, sont traitées de façon à obtenir un résidu insoluble à l'alcool (AIR). Ce traitement permet d'éliminer une grande partie des lipides, pigments et protéines, qui endommageraient les colonnes lors des purifications. Deux extractions aqueuses des polysaccharides ont été réalisées. La première consiste à faire bouillir l'échantillon plusieurs fois deux heures à 95 °C et la deuxième à l'autoclaver 5 minutes à 120 °C. Dans les deux cas, un rendement en polysaccharides de 9% (w/w) était obtenu. Les polysaccharides issus des deux protocoles ont été digérés par la GH16 *Zg*AgaB et les oligosaccharides produits ont été analysés par FACE. La différence majeure entre les deux protocoles réside dans le nombre plus important de petits oligosaccharides chargés obtenus à partir de la préparation autoclavée par rapport à ceux obtenus par ébullition (**Figure 39**). Ainsi, le polymère obtenu par autoclavage contient probablement plus de modification que celui préparé par l'autre méthode. Ceci est probablement dû au fait que l'extraction étant plus rapide, 5 minutes à 120°C au lieu de deux heures à 95°C, les modifications sont mieux préservées. Sauf mention contraire, les tests de caractérisation ont tous été réalisés à partir du polysaccharide extrait par autoclavage.

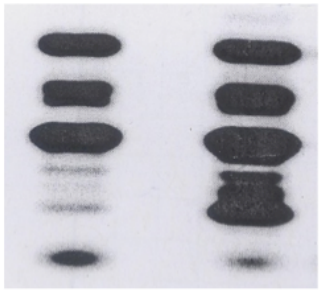


Figure 39 : Extrait pariétal total de *Polysiphonia elongata* digéré par *Zg*AgaB
Le marquage a été réalisé à l'AMAC, ainsi seuls les oligosaccharides chargés sont visibles. Aucune des fractions n'a subi d'ultrafiltration avant électrophorèse.

IV. 2. E Purification des oligosaccharides issus de *Polysiphonia elongata* et tests d'activité sulfatases

Il est très difficile de purifier des oligosaccharides provenant d'agar complexe à cause des nombreuses modifications et isomères produits lors de la dégradation enzymatique.

Ainsi, une seule chromatographie par exclusion de taille n'est pas suffisante pour isoler un oligosaccharide. Afin de permettre leur caractérisation structurale par spectrométrie de masse et éviter que la présence d'isomères ne perturbe l'identification des fragments, les oligosaccharides doivent être les plus purs possibles. Comme nous recherchions des oligosaccharides sulfatés, une purification par chromatographie sur résine échangeuse d'anions semblait être la meilleure option pour séparer les différents oligosaccharides sulfatés.

Sur la base de tests préliminaires réalisés par le Dr. Diane Jouanneau sur les carraghénanes (communication personnelle), j'ai choisi de tester différentes matrices commerciales : QFF, ANX, DEAE, CaptoDEAE et CaptoAdhere (GE Healthcare). La résine qui permettait la meilleure séparation s'est avérée être la CaptoAdhere. En effet, pour un gradient identique (**Figure 40**) de NaCl 1M, j'ai obtenu, en fonction de la matrice choisie, des profils electrophorétiques très différents. Ainsi, à titre d'exemple, avec la matrice DEAE, nous avons obtenu une élution des oligosaccharides qui commence à environ 1 M de NaCl (**Figure 40**), alors qu'avec la matrice de type QFF l'élution des oligosaccharides nécessite au moins 2 M de NaCl. Pour ces phases stationnaires, le fait que les oligosaccharides étaient élués à haute concentration de sels et faiblement résolus a disqualifié l'utilisation de ces matrices.

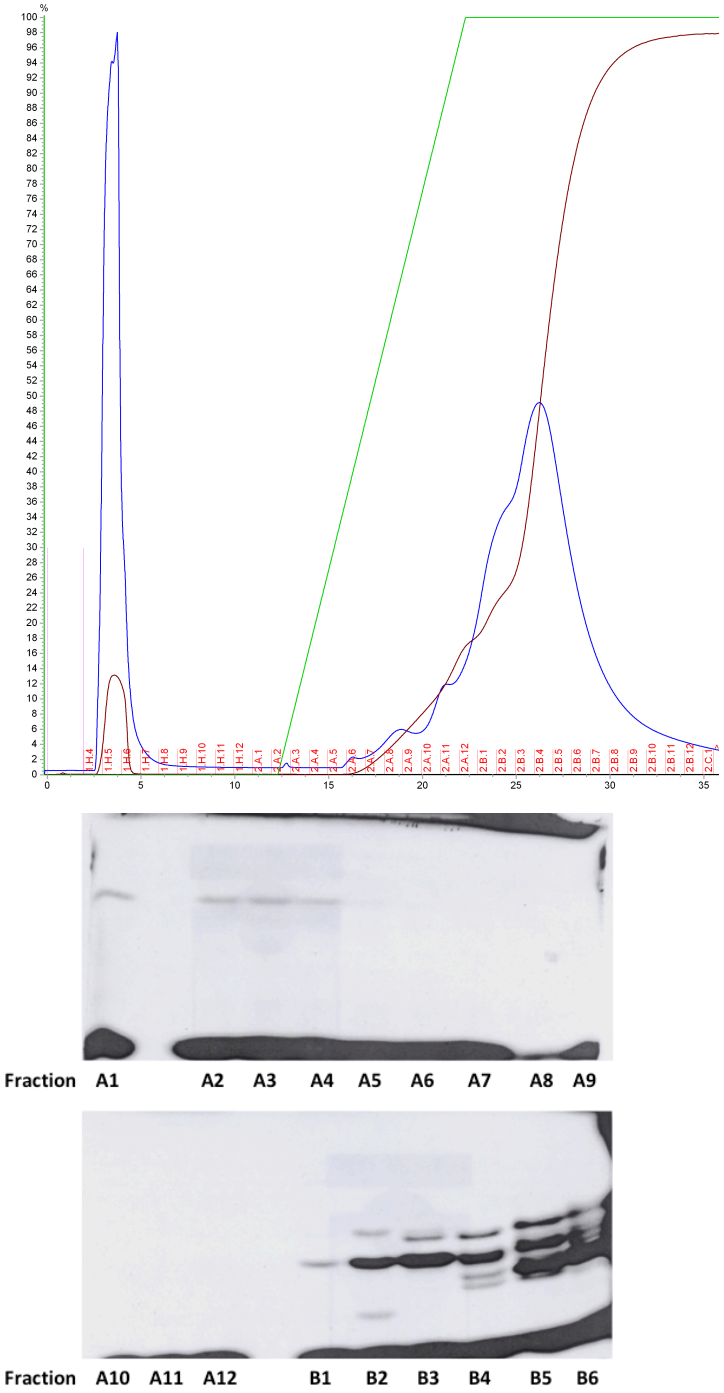


Figure 40 : Gradient d'élution et visualisation de la composition en oligosaccharides des différentes fractions de purification des oligosaccharides issus de *Polysiphonia elongata* sur une résine DEAE de 1 mL.

Le gradient est montré dans le panneau supérieur. Le gradient en sels est symbolisé par la courbe en vert (avec les concentrations en sels exprimées en pourcentage au niveau de l'axe des ordonnés). Le signal UV est enregistré et symbolisé par la courbe bleue (le signal maximum UV atteint 14 mAU). La conductimétrie est mésurée et symbolisée par la courbe marron. Le volume des fractions collectées était de 1 mL. La numérotation des fractions correspond aux numéros des plaques 96 puits utilisés pour le fractionnement, le numéro des fractions est indiqué en bas des gels et en bas du gradient en rouge. Le fluorophore utilisé pour le marquage des oligosaccharides est l'ANTS (fluorophore chargé négativement). La ligne noire en bas des gels correspond à la migration de l'ANTS non lié.

De plus, la présence de haute concentration de NaCl (jusqu'à 2 M) dans les fractions d'oligosaccharides perturbe les électrophorèses FACE, les chromatographies sur couche mince (CCM) et est également incompatible avec la spectrométrie de masse. Nous avons donc cherché à éliminer les sels par différentes méthodes.

Une dialyse avec des membranes de 300 Da a tout d'abord été testée. Il s'est avéré que ce n'était pas une solution envisageable pour des purifications préparatives, due à la perte importante de matériel lors du processus, mais aussi parce que cette technique ne permettait pas un dessalage suffisant de l'échantillon après trois jours de dialyse.

La seconde option a été d'adapter un protocole utilisé pour le dessalage de l'héparine (Powell et al., 2010). Ce dernier consiste à mettre en série deux colonnes de dessalage 26/10 (HiPrep Desalting, GE). Cependant, nous avons été dans l'incapacité de récupérer l'échantillon à la sortie des colonnes, probablement parce qu'il était trop dilué pour être détecté par électrophorèse FACE ou par CCM.

Devant la difficulté technique que représente l'étape de dessalage, nous avons remplacé, lors de l'élution, le chlorure de sodium par le carbonate d'ammonium. En effet, ce dernier a l'avantage de se sublimer entièrement quand il est lyophilisé. Il est cependant nécessaire de travailler à 4 °C lors des purifications à haute concentration de carbonate d'ammonium, pour éviter un dégagement de CO_2 dans les colonnes de chromatographie.

Une première purification d'oligosaccharides de *P. elongata* avait été réalisée par chromatographie échangeuse d'anions sur CaptoAdhere (**Figure 41**). Le polysaccharide utilisé pour ce premier test a été extrait par ébouillantage.

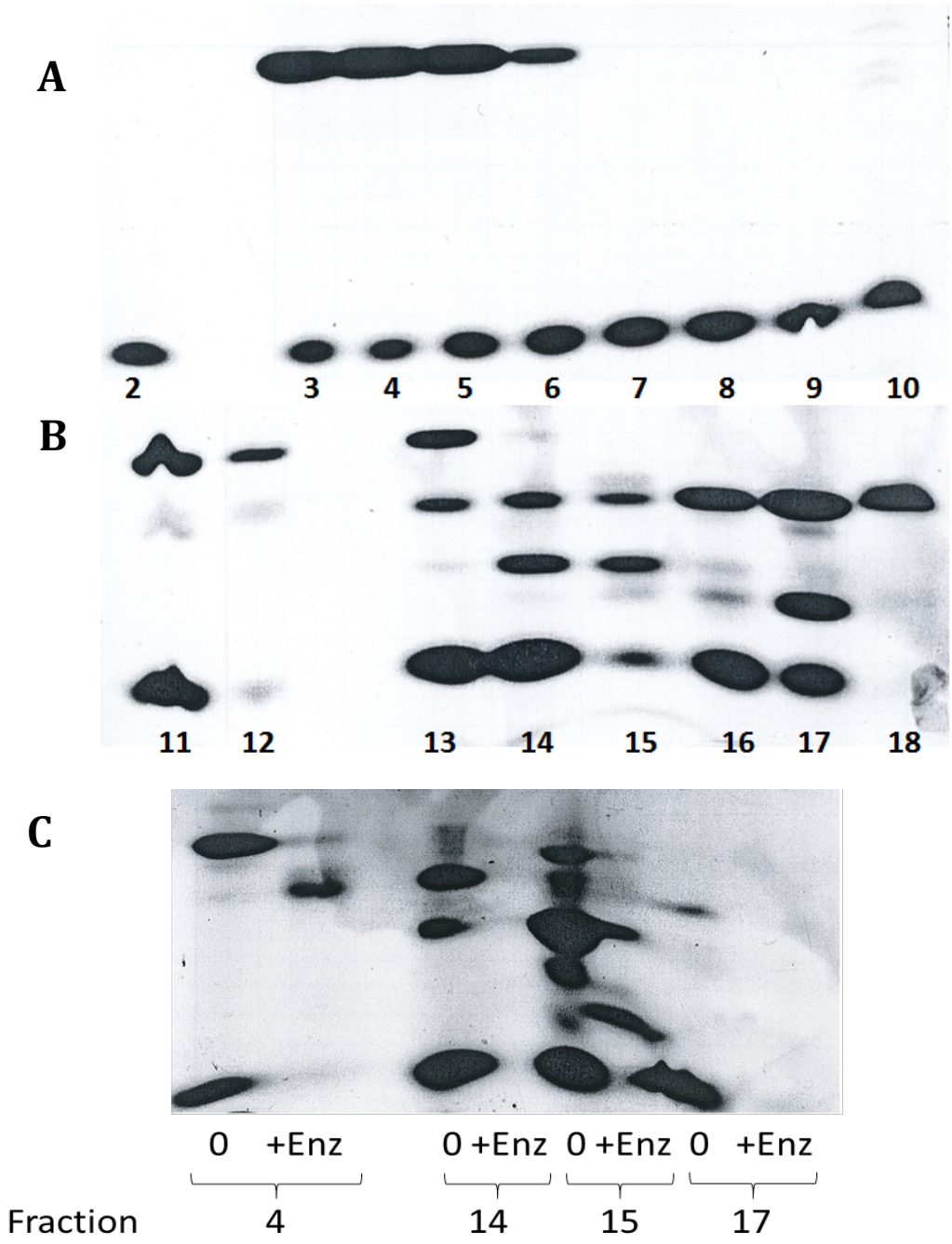


Figure 41: Récapitulatif de la première purification des oligosaccharides issus de *Polysiphonia elongata* par chromatographie d'échange d'anions en utilisant les colonnes CaptoAdhere et premier test d'activité de *Zg*4660.

Les différentes pistes correspondent aux différentes fractions éluées par un gradient de concentration croissant en carbonate d'ammonium (Figure A et B). La figure C correspond au test sulfatase sur les fractions 4, 14, 15 et 17 issues de la purification par CaptoAdhere. Les pistes en dessous de lesquelles est indiquées O ont été incubées avec l'enzyme bouillie, les pistes en dessous de lesquelles est indiqué +Enz ont été incubées avec la sulfatase *Zg*4660. Tous les oligosaccharides ont été marqués avec le fluorophore neutre AMAC.

Les oligosaccharides ainsi obtenus ont permis de réaliser un premier test d'activité de la sulfatase *Zg*4660. Une activité sur les oligosaccharides des fractions 3 à 6 (qui seront ensuite regroupés et nommées F3) et aussi sur la fraction 15 (**Figure 42**) avait été observée.

Suite à ce test, j'ai voulu améliorer le gradient pour mieux séparer les fractions 11 à 20. Ces oligosaccharides étant élués entre 6 % et 15 % de carbonate d'ammonium, j'ai divisé par deux la pente du gradient (**Figure 43**). Cette seconde purification a été réalisée avec un polysaccharide issu de l'extraction optimisée ce qui peut expliquer une partie des différences visibles. Avec la partie non modifiée du gradient, nous retrouvons le même profil.

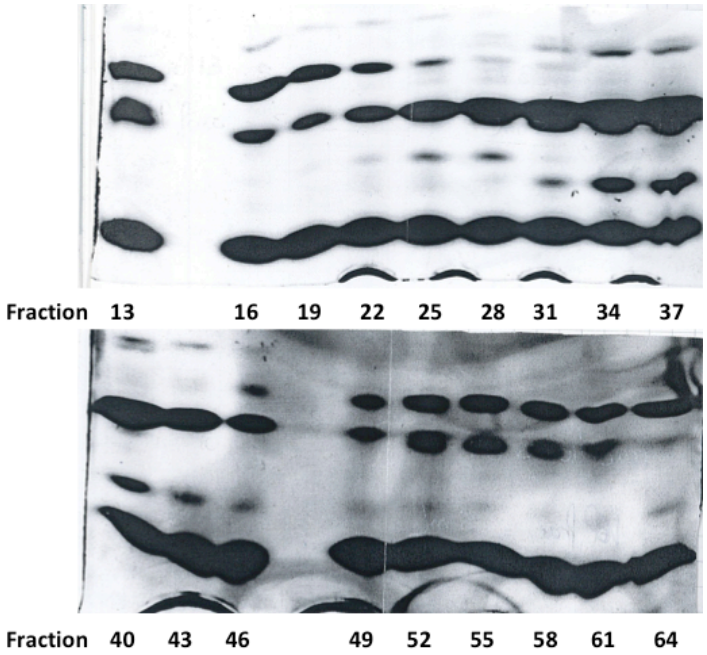


Figure 43 : Amélioration du gradient d'élution des oligosaccharides extraits de *Polysiphonia elongata.*

Pour cette purification optimisée, les fractions sont de taille plus réduite: 5 mL au lieu des 15 mL du premier test. La fraction 13 correspond à la première fraction éluée avec la pente modifiée par rapport à la première purification. Le marquage a été réalisé à l'AMAC.

A l'issue de ces tests et de l'accumulation des oligosaccharides F3, j'ai voulu confirmer les résultats de l'activité sulfatase par HPAEC. Pour cela, j'ai dû réaliser de nouvelles cultures d'*E. coli* dans le but de préparer à nouveau la sulfatase *Zg*4660. Cependant, des difficultés d'expression de la protéine sous forme active ne m'ont pas permis de tester ces fractions d'oligosaccharides.

IV. 2. F Tests pour reproduire l'enzyme *Zg*4660 sous forme active

Après avoir re-séquencés les plasmides stocks, une co-transformation fraiche, a été refaite car une dérive des plasmides est toujours possible, mais sans succès. Face à cette échec, j'ai cherché à développer un nouveau protocole où les chaperonnes ne seraient pas induites. En effet, avec l'éthanol nous induisons des chaperonnes dont les identités et les proportions sont inconnues. De plus, notre système d'expression impliquant déjà deux plasmides, réussir à y inclure un troisième codant pour des chaperonnes qui soit compatible avec les deux autres, était très improbable.

Notre idée a donc été de récréer chez *E. coli* le système de modification co-traductionnelle des sulfatases de *Z. galactanivorans*. Comme expliqué dans l'introduction, cette modification semble se dérouler dans le périplasme chez de *Z. galactanivorans*.

Dans ce but, j'ai préparé de nouveaux plasmides. Le gène *zgfge* a été cloné, avec son peptide signal, en aval d'un promoteur arabinose. L'avantage de ce dernier est que l'expression du gène *zgfge* sera dépendante de la concentration en arabinose. Cela permettra d'adapter les quantités de *Zg*FGE à ceux de la sulfatase co-produite et d'éviter une surproduction inutile de *Zg*FGE. L'autre avantage de ce promoteur est que nous avons ainsi découplé les productions des deux enzymes. La sulfatase a également été reclonée sous sa forme entière avec une étiquette histidine à son extrémité C-terminale.

Quatre tests ont été tentés, dont le but était de déterminer les conditions permettant une production optimum de sulfatase. Ainsi, deux concentrations d'arabinose de 0,1 % et 0,2 % et deux temps d'induction ont été testés pour mettre au point le taux d'expression de *Zg*FGE.

Les tests de production et de purification ont permis de mettre en évidence que la sulfatase était bien produite. Cependant, certaines conditions de culture et d'expression utilisées (induction arabinose à 0,3 de DO_{600} par exemple) n'ont permis qu'une expression très faible. La production maximum de protéine a été obtenue à la concentration finale d'arabinose de 0,2%, ajouté à 0,6 de DO_{600nm}. Malheureusement, la sulfatase ainsi obtenue n'était pas active sur substrat artificiel.

IV. 2. G Analyse des biais de codons présents au sein des séquences codant pour les sulfatases

L'une des premières questions à laquelle nous avons été confrontés au début de nos travaux sur les tests d'expression des sulfatases chez *E. coli* a été : pourquoi certaines sulfatases sont exprimées actives et d'autres non ? Bien évidemment, le problème potentiel de la reconnaissance des séquences des différentes sulfatases de *Z. galactanivorans* par le système de modification post-traductionnelle des sulfatases d'*E. coli* a été prise en compte. Cependant, les tests effectués en présence d'éthanol et donc permettant l'expression de chaperonnes, qui ont permis d'obtenir une sulfatase active donc modifié par le système *d'E. coli*, plaidaient plutôt pour un problème de cinétique de traduction et/ou de repliement que de compatibilité de séquence.

Or l'un des moyens de jouer sur la cinétique de traduction sont les pauses ribosomiques induites par la présence de codons rares dans l'ARNm.

Le modèle putatif d'expression des sulfatases chez *Z. galactanivorans* implique un début de traduction suivi d'une pause pour permettre d'initier la translocation du peptide vers le périplasme probablement *via* le système Sec (Driessen et al., 1998), où le peptide est modifié par la FGE (Cf. Introduction). Si la traduction est trop rapide, la FGE (quelle que soit son origine) peut ne pas avoir le temps de modifier la sulfatase avant son repliement, d'où la nécessité d'une pause traductionnelle. Celle-ci laisserait le peptide en cours d'élongation à distance compatible de la membrane cytoplasmique pour que la FGE, qui y est probablement ancrée, puisse agir.

Si ce raisonnement est valable chez *Z. galactanivorans*, on peut supposer qu'il le soit également chez *E. coli*. En effet, si nous exprimons de façon recombinante chez *E. coli* une séquence sans ou avec trop peu de codons rares, nous n'aurons pas la pause traductionnelle nécessaire à la modification. Deux questions se posent alors : est-ce que les sulfatases naturellement actives chez *E. coli* comprennent des codons rares ? Cela expliquerait pourquoi certaines sont actives et d'autres non. La deuxième question est : est-ce que ce système de contrôle cinétique est bien celui utilisé par *Z. galactinovorans* ?

Pour répondre à la première question, nous avons réalisé une étude des biais de codons sur les séquences de sulfatases de *Z. galactanivorans* déjà exprimées chez *E. coli* (Sharp et al., 1988). La notion d'activité, très relative ici, correspond à la vitesse d'observation de fluorescence dans le cadre de tests d'expression. Il est alors apparu que les sulfatases très actives avec le substrat MUS (*Zg*3527, *Zg*3461 et *Zg*3515) ou

moyennement actives sur ce même substrat (*Zg*3548, *Zg*3635, *Zg*3469, *Zg*3503, *Zg*3509, *Zg*3578 et Zg4660) ont toutes des codons rares dans les séquences situées 20 à 30 acides aminés en aval du site catalytique. Pour les résidus du motif consensus (CXPXR), l'arginine est souvent un codon rare pour les sulfatases très actives ou moyennement actives. La catégorie la plus active a généralement 0,66 codon rare en position + 10/ + 20 et +20/+30 par rapport à la séquence consensus, et au moins 1 codon peu fréquent dans la région +20 et + 30. Ces statistiques sont plus basses avec la catégorie des sulfatases moyennement actives : 0,57 codon rare en position + 10/ + 20 et 0,43 codon rare en position +20/+30 acides aminés. Ces différences suggèrent une importance particulière des codons rares dans la portion d'ARNm codant pour les acides aminés éloignés de 20 à 30 positions du site catalytique.

Cela est également confirmé pour les sulfatases peu actives (*Zg*3642, *Zg*3643) et très peu actives (*Zg*2198). En effet, dans la portion de séquences d'ARNm correspondant aux positions +20/+30, aucun codon rare n'est retrouvé. Pour les séquences classées peu actives, il y a 1 codon peu fréquent dans la portion codante située entre les positions + 10 et +20 après la séquence consensus.

IV. 2. H Cristallisation de *Zg*4660

Les différents tests de cristallisation ont montré que la concentration en protéine (20 mg/mL) était compatible avec la cristallisation de la sulfatase *Zg*4660. En effet, plusieurs précipités microcristallin ont été observés. Néanmoins, nous avons eu peu de conditions exploitables. Une seule a pu être reproduite et optimisée par A. Jeudy, il s'agit d'une condition retrouvée au sein du crible PACT (Quiagen). Cependant, les cristaux n'apparaissent, au mieux, qu'après trois semaines et n'apparaissent que en boite Greiner 96 puits (gouttes assises). La condition donnant les cristaux les plus réguliers est 0,15 M ammonium, 24 % de PEG3350 (**Figure 44**), mais ces cristaux sont très fragiles.

Figure 44 : Cristaux de Zg4660 après optimisation

Les cristaux ont été obtenus dans les conditions suivantes : 0,15 M ammonium, 24 % de PEG3350.

Les cristaux ainsi obtenus ont été testés sur la ligne ID30 de l'ESRF. Ils diffractent au mieux à 6 Å. Les clichés de diffraction présentent une diffraction avec un motif hexagonal répété. Néanmoins, il a été impossible d'indexer ces clichés. Un test de déshydratation ou bien de pontage par le glutaraldéhyde pour diminuer le pourcentage de solvant et améliorer la diffraction a été envisagé, mais la fragilité des cristaux a empêché la première option. La taille des gouttes et l'utilisation de plaques de type Greiner, ont rendu la deuxième option trop complexe à réaliser.

Au vu de ces résultats, nous avons décidé de tester une nouvelle approche. Au cours de la thèse de D. Martinez Zapien, avaient été mis au point des vecteurs comprenant une séquence codant pour une *maltose binding protein* (MBP) dont des résidus de surface ont été modifiés. Ces modifications permettent de réduire l'entropie de surface et ainsi de favoriser la formation de réseau cristallin (Moon et al, 2010). Le second avantage de cette fusion est de permettre facilement le phasage en utilisant les structures existantes de la MBP. La séquence d'acides aminés entre la MBP et la protéine clonée est composée de plusieurs alanines, notamment trois dans le cas de la protéine Rack1 (qui est une protéine dont la cristallisation et la résolution structurale a été mené à bien *via* cette stratégie) (PDB 3DMO) (Ullah et al., 2008). Dans le cas du clonage des sulfatases en fusion avec la MBP, pour ne pas avoir trop de flexibilité, j'ai décidé de conserver ce motif polyalanine en testant deux degrés de liberté différents (3 alanines et 6 alanines) plutôt que d'utiliser une séquence composée d'une alternance de prolines et de glycines. Ces

constructions ont permis d'obtenir plusieurs précipités micro-cristallins que nous cherchons actuellement à optimiser.

IV. 3 Conclusions et Perspectives

Au début de cette étude, deux questions majeures se posaient : comment obtenir l'enzyme active et comment identifier son substrat ?

Pour répondre à la première, je suis partie du postulat que la FGE d'*E. coli* n'est pas optimisée pour les séquences de *Z. galactanivorans*. Pour remédier à ce problème, la sulfatase *Zg*4660 a été surexprimée avec la *Zg*FGE. Ceci a permis d'obtenir une augmentation de l'activité de la sulfatase de 67 fois par rapport à la protéine exprimée seule. Un second test a consisté à exprimer les deux protéines en induisant des chaperonnes. L'expression de chaperonnes par la présence d'éthanol dans le milieu de culture a permis d'augmenter l'activité de la sulfatase *Zg*4660 de 1252 fois. L'implication supposée de ces chaperonnes suggère que la traduction se fait trop rapidement pour permettre la modification chez *E. coli*. Or comment dans une cellule peut-on avoir un contrôle cinétique de la traduction ? L'une des réponses possibles est le rôle éventuel des biais de codons. Pour confirmer ou infirmer cette hypothèse, une étude des biais de codons présents dans la séquence a été réalisée et corrélée avec l'activité de la protéine obtenue chez *E. coli*. Bien que cette étude ne soit pas exhaustive, il est apparu que les séquences permettant d'obtenir les sulfatases les plus actives sont celles où l'on trouve le plus de biais de codons dans la région codante pour les 30 acides aminés situés en aval de la séquence catalytique consensus. Néanmoins, pour conclure, il faudrait réaliser une étude systématique des biais de codons des séquences utilisées pour savoir si ces biais sont responsables ou non des variations d'expression observées.

Une fois l'enzyme obtenue sous forme active en co-expression avec *Zg*FGE et une induction de chaperonnes, il était nécessaire de mettre au point une stratégie de criblage efficace pour identifier le substrat de *Zg*4660. La stratégie mise au point dans ce travail repose sur la dégradation d'agars de différentes origines par des β-agarases et β-porphyranases de la famille GH16 de spécificités connues. Ainsi, nous avons obtenu un très large panel d'oligosaccharides présentant différentes modifications. Ces derniers ont été incubés avec la sulfatase *Zg*4660 et les éventuels désulfatations ont été mise en évidence par électrophorèse FACE en comparant, avant et après incubation, les profils électrophorétiques des oligosaccharides marqués à l'aide d'un fluorophore neutre. Ainsi la migration des oligosaccharides ne se fait qu'en fonction des charges naturellement présentes et si l'une d'entre elles est éliminée alors la migration est clairement impactée.

Ce criblage a permis d'identifier que le substrat de *Zg*4660 est un oligosaccharide généré par *Zg*PorB à partir de l'agar produit par *Polysiphonia elongata*. Pour caractériser ce dernier il était nécessaire d'optimiser l'extraction de l'agar de *P. elongata* et de purifier les oligosaccharides. Pour réaliser la purification des oligosaccharides j'ai testé différentes matrices d'échangeuses d'anions et celle qui a été retenue est la matrice CaptoAdhere car elle est plus résolutive que les autres. De plus, nous avons mis au point la purification avec un nouvel éluant qui nous permet d'éviter une étape de dessalage : le carbonate d'ammonium. Cette technique a permis d'accumuler suffisamment d'oligosaccharides pour envisager une caractérisation mais malheureusement avant d'avoir pu réaliser cette étape nous avons été dans l'incapacité de reproduire l'enzyme sous forme active.

Pour remédier au problème, nous avons commencé la mise en œuvre d'un système reproduisant le système de maturation des sulfatases de *Z. galactanivorans* chez *E. coli*. Ce système de maturation repose sur l'expression de *Zg*FGE dans le périplasme puis l'expression de la sulfatase dans le même compartiment. L'idée étant que *Zg*FGE devrait plus facilement modifier la sulfatase exportée sous forme dépliée via un transporteur Sec. Malheureusement, ce test ne s'est pas révélé concluant.

Les pistes à explorer pour améliorer ce système sont les suivantes. Tout d'abord il faudra finir le travail sur les biais de codons pour vérifier si le contrôle de la modification comporte bien une part cinétique. Il sera nécessaire d'effectuer également ce même travail pour l'expression chez *Z. galactanivorans* afin de confirmer ou d'infirmer que c'est une partie du système de régulation natif chez cette bactérie. Les biais de codon de *Z. galactanivorans* ont déjà été calculés par le Dr. M.M. Perrineau (communication personnelle). Il reste à réaliser l'analyse statistique des biais de codons chez toutes les sulfatases de *Z. galactanivorans*.

Le système d'expression des sulfatases au sein du périplasme doit être amélioré notamment en déterminant le peptide signal le plus efficace permettant la translocation des sulfatases.

La deuxième piste concerne une étude de la localisation correcte de *Zg*FGE et de son insertion dans la membrane. Une des méthodes pouvant être employée à cette fin est l'expression de *Zg*FGE avec la protéine rapporteuse *Green Fluorescent Protein* (GFP) pour pouvoir localiser la protéine-fusion. L'autre avantage d'utiliser une fusion de ce type sous le contrôle d'un promoteur arabinose est de déterminer plus facilement la

concentration d'arabinose et la densité optique de culture pour avoir une expression optimale de *Zg*FGE sans épuiser les ressources de la bactérie. En effet, nous pourrions calculer approximativement la production de la fusion en se basant sur la fluorescence émise par la GFP. L'idée qui émerge de cette étude est de disposer d'un système de modification des sulfatases de multiples origines, « *ready to use* » chez *E. coli*. Ce projet est extrêmement complexe et nécessite un travail de fond important, il n'est donc pas envisagé ici pour clore le projet *Zg*4660 dans un futur proche.

Afin de résoudre le problème de répétabilité de la production de *Zg*4460 active, et ainsi finir le projet, il est envisagé une surexpression chez *Z. galactanivorans* utilisé comme système homologue de surexpression. Mais ce projet est, lui aussi, loin d'être trivial, en effet il est extrêmement complexe de réaliser des modifications génétiques au sein de *Z. galactanivorans*. En effet, les outils génétiques fonctionnels pour les bactéries Gram positif ou les proteobactéries Gram négatif sont généralement inutilisables chez les *Bacteroidetes* (McBride and Kempf, 1996). Cependant, il a été récemment mis au point une méthode de conjugaison bactérienne pour permettre l'insertion et la délétion de gènes chez *Z. galactanivorans* (Zhu et al., 2017). Dans le cadre de cette étude, le gène de la GFP a été placé sous le contrôle du promoteur du gène *ompA*, dans une zone dite neutre du génome, ce qui a abouti à la production de la protéine GFP dans cet organisme (Zhu et al., 2017). Actuellement des travaux sont réalisés sur la transformation de *Z. galactanivorans* afin d'obtenir sous forme active la sulfatase *Zg*4660, à laquelle une étiquette histidine a été ajoutée pour faciliter sa purification.

Pour conclure l'étude de *Zg*4660, je rappelle que cette dernière fait partie de la sous-famille S1_19 des sulfatases formylglycine dépendante (Barbeyron et al., 2016a). Or, un des membres de cette famille, *Zg*CgsA (*Zg*3145) est une 4S-iota-carraghénane sulfatase récemment caractérisée (Ficko-Blean et al., 2017). L'appartenance de *Zg*4660 et *Zg*CgsA à la même sous-famille de la famille S1 des sulfatases, suggère que les deux sulfatases pourraient agir sur un carbone sulfaté en position identique présent à la fois dans le ι-carraghénane (C4 du D-galactose) et dans l'agar de *Polysiphonia elongata* (C4 du L-galactose). Si cela était confirmé, cela ferait de *Zg*4660 la première 4S-agar sulfatase caractérisée.

Chapitre V : Couple *Zg*3515 (sulfatase S1_20) et *Zg*3516 (GH30)

V. 1 Introduction

L'un des enjeux de cette thèse est d'explorer les carbohydrates sulfatases (CS) actives sur les polysaccharides marins. A l'heure actuelle, seules les sulfatases de la famille S1 sont actives sur les carbohydrates (Barbeyron et al., 2016a). Le génome de *Zobellia galactanivorans* Dsij^T possède 71 sulfatases et 22 loci d'utilisation des polysaccharides (PUL) putatifs comprenant au moins une sulfatase (Barbeyron et al., 2016b), ce qui fait de cet organisme un modèle pour l'étude des CS.

La difficulté majeure rencontrée lors de l'étude de ces enzymes est de les obtenir sous forme actives. Le résidu catalytique est une formylglycine issue d'une modification co-traductionnelle consistant en une oxydation d'une cystéine ou d'une sérine (Hanson et al., 2004). Ainsi, quand la protéine est produite chez *E. coli* la modification peut ne pas avoir lieu car les séquences du site catalytique des sulfatases hétérologues ne sont pas ou sont imparfaitement reconnues par la machinerie d'*E. coli*.

Pour cette raison, si nous souhaitions continuer à étudier les carbohydrate sulfatases de *Z. galactanivorans*, il était crucial d'avoir une sulfatase naturellement modifiée par *E. coli*. Heureusement, lors des premiers tests d'expression et d'activité des sulfatases réalisés par le Dr. T. Barbeyron et R. Larocque, une sulfatase exprimée chez *E. coli* s'est révélée particulièrement active sur le substrat MUS : *Zg*3515. Cette dernière se trouve au sein du PUL 37 (**Figure 45**) (Barbeyron et al., 2016b), où se trouve une glycoside hydrolase de la famille GH30. Cette dernière était annotée comme étant une endo-1,4-β-xylanase du fait de sa proximité de séquence avec les xylanases déjà caractérisées. Cependant, aucune donnée expérimentale n'ayant confirmé cette annotation ce PUL n'est pas considéré comme étant un PUL xylanolytique. Si au sein des plantes terrestres, la présence de xylane sulfaté n'a jamais été démontrée ce n'est pas le cas chez les algues. La présence de xylanes sulfatés au sein de la paroi de *Palmaria palmata* a en effet été rapportée (Deniaud Boüet, 2002). Au sein du PUL 37, deux autres sulfatases *Zg*3517 et *Zg*3513, ainsi qu'un récepteur TonB dépendant (*Zg*3511) et une esterase de la famille GDSL (*Zg*3514) sont présents. Les sulfatases *Zg*3517 et *Zg*3513 appartiennent aussi à la sous-famille S1_20 pour laquelle aucune activité n'a encore été décrite, ni de structure disponible (Barbeyron et al., 2016a). Ainsi, la caractérisation d'une de ces sulfatases permettrait une avancée dans la description de cette sous-famille.

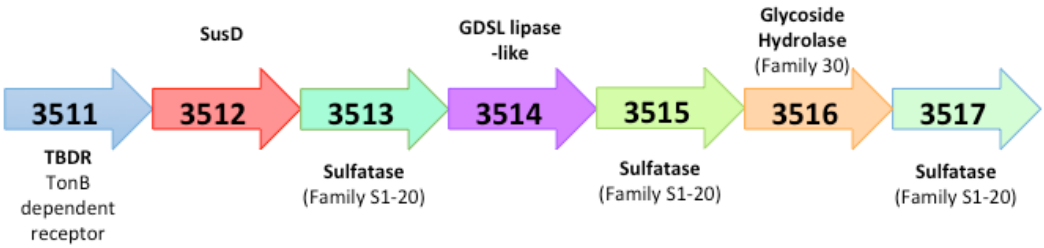


Figure 45 : Représentation schématique du PUL 37 de *Zobellia galactanivorans* DsijT

Pour appréhender le fonctionnement du PUL 37, nous avons cherché les localisations putatives de ces enzymes au sein de la bactérie. Les sulfatases *Zg*3513 et *Zg*3517 ont été annotées comme étant des lipoprotéines ancrées au niveau de la membrane externe, tandis que *Zg*3515 et *Zg*3516 sont proposées comme étant périplasmiques, *Zg*3514 étant ancrée à la membrane cytoplasmique. Nous pouvons ainsi émettre des hypothèses sur la nature du substrat de ce PUL et sur le fonctionnement de ce dernier. Le substrat est probablement fortement sulfaté puisque trois sulfatases différentes sont présentes. L'agar substrat est aussi probablement acétylé par l'acide pyruvique ou bien comporte des acides galacturoniques. En effet, *Zg*3514 est une lipase de type GDSL qui sont des lipases ou des estérases comprenant le motif GxSxG et qui sont actives sur une très grande gamme de substrats (Akoh et al., 2004), dont VmA une lipases de type GDSL active sur le *p*-nitrophényl butyrate (Shaw et al., 1994). Le butyrate ressemble au pyruvate, qui présente une fonction cétone supplémentaire par rapport au butyrate. C'est pourquoi il est très probable que *Zg*3514 permette l'élimination d'un pyruvate (sauf si les pyruvyl sont sous forme d'acétal) ou l'élimination du carboxylate de l'acide galacturonique.

Concernant l'ordre séquentiel, nous aurions un substrat qui serait partiellement désulfaté extracellulairement par les sulfatases *Zg*3513 et *Zg*3517 puis qui serait importé dans périplasme via le couple de protéines de type SusC/SusD de la membrane externe codées par *Zg*3511 et *Zg*3512. Le substrat serait complètement désulfaté dans le cytoplasme par la sulfatase *Zg*3515 où il serait finalement désestérifié/décarboxylé par la lipase de type GDSL *Zg*3514.

V. 1. A La famille 30 des glycoside hydrolases (GH30)

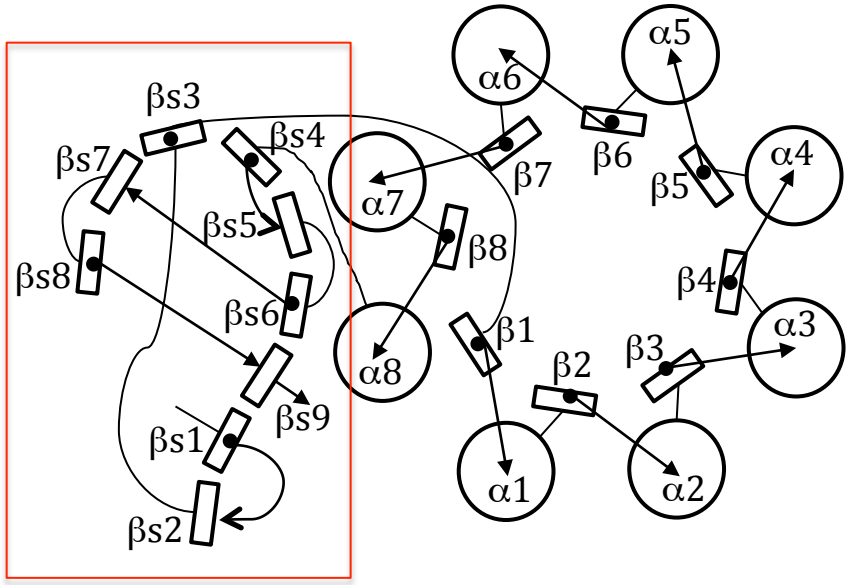
La classification de cette famille a été revue en 2010 (St John et al., 2010), en effet il était alors apparu qu'un certain nombre d'enzymes de la famille GH5 était en réalité plus proche des GH30. Aujourd'hui, dans la base CAZy, la famille GH30 est subdivisée en neuf sous-familles. Les activités présentes dans la famille GH30 sont des β-glucosylceramidases et β-glucosidases (EC 3.2.1.45, et EC 3.2.1.21, GH30_1), des xylanes 1,4-β-xylosidases (EC 3.2.1.37, GH30_2), des glucanes endo-1,6-β-glucosidases (EC 3.2.1.75, GH30_3), des β-D-fucosidases (EC 3.2.1.38, GH30_4), des galactanes endo-1,6-β-galactosidases (EC 3.2.1.164, GH30_5), des endo-1,4-β-xylanases (EC 3.2.1.8, GH30_7 et GH30_8), des glucuronoarabinoxylan endo-1,4-β-xylanases (EC 3.2.1.136, GH30_8) et des β-glucuronidases (EC 3.2.1.31, GH30_9) (St John et al., 2010).

Les enzymes de cette famille fonctionnent par un mécanisme de rétention de configuration du carbone anomérique de double déplacement décrit par Koshland (Miao et al., 1994).

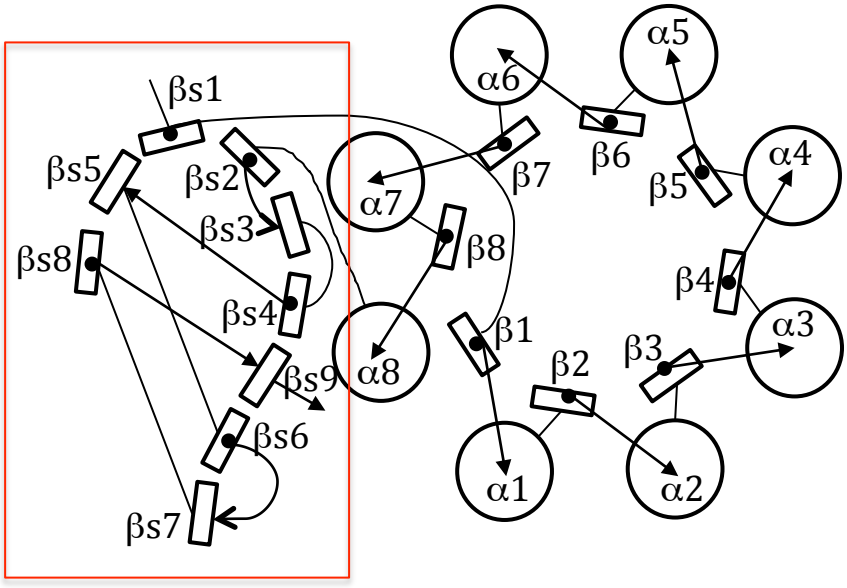
Le nucléophile catalytique, un glutamate, a été identifié pour la première fois en 1994 (Miao et al., 1994) chez la β-glucocerebrosidase humaine en piégeant l'intermédiaire réactionnel. La famille GH30 est composée aussi bien d'enzymes ayant un mode d'action endolytique, qui clivent le polysaccharide au milieu de la chaine, qu'exolytique qui clivent le polysaccharide via une des extrémités. Une classification supplémentaire des enzymes exolytiques est possible selon que la liaison osidique hydrolysée se situe du côté réducteur ou non réducteur.

Au niveau de leur structure tertiaire, les GH30 ont une structure composée d'un tonneau TIM $(\alpha/\beta)_8$ avec les deux résidus catalytiques acide/base et nucléophile situés respectivement au niveau des extrémités C-terminales des brins β4 et β7. Un deuxième domaine, de fonction inconnue, caractérise les GH30. Il s'agit d'un β-sandwich qui est fusionné avec des éléments conservés des extrémités N et C-terminales et forme une interface hydrophobe avec l'extrémité du tonneau TIM où se trouve le site actif. Un rôle proposé pour ce domaine est qu'il serait conservé pour étendre la diversité de substrats reconnus. En effet, les structures primaires de cette région sont très peu conservés, au contraire des structures secondaires, et cette variabilité des acides aminés pourrait permettre une acquisition plus rapide de nouvelles fonctions (St John et al., 2010).

Par ailleurs l'organisation des différents brins composant le β-sandwich permet de subdiviser en 2 classes les glycosides hydrolases de la famille 30 (**Figure 46**).



Sous-Famille 1



Sous-Famille 2

Figure 46 : Présentation des deux sous familles structurales de GH30 (Tiré de St John et al., 2010)

Les structures secondaires des deux sous-famille sont symbolisées ici selon le code suivant : un rectangle est un brin β, un rond est une hélice α. La succession des structures en fonction des extrémités N-terminale et C-terminale est indiquée par les fléches dont la tete indique l'extrémité C-terminale du linker considéré. Le β-sandwich est entouré par un encadré rouge et les brins constitutifs de ce dernier sont indiqués avec un s. Les deux sous-familles se distinguent par l'organisation de leur domaine β sandwich.

Les GH30 n'ont pas été uniquement subdivisées en fonction de ces sous-familles structurales mais également selon leur phylogénie au sein de laquelle nous retrouvons pour chaque clade une activité enzymatique de référence (St John et al., 2010).

V. 1. B Les divers polysaccharides retrouvées chez les algues brunes et vertes

Dans ce chapitre de nombreux polysaccharides, autres que les agars, ont été testés. Nous présentons donc ici une courte description de chacun d'entre eux. Le choix de ces polysaccharides a été en partie dicté par les données de transcriptomique dont nous disposions. En effet, la plupart des gènes du PUL 37 sont induits en présence de porphyrane ou de κ-carraghénane, à l'exception de la GH30_4 *Zg*3516 et de la sulfatase *Zg*3517 qui, elles, ne sont induites qu'en présence de κ-carraghénane (Thomas et al., 2017).

Les carraghénanes, les fucoïdanes et les ulvanes sont tous des polysaccharides pariétaux sulfatés extraits, respectivement, des algues rouges (carraghénophytes), brunes et vertes. Les xylanes sont aussi des polysaccharides pariétaux, très probablement sulfatés, présents chez certaines algues rouges non carraghénophytes (Popper et al., 2011).

En fonction de leur degré de sulfatation, les carraghénanes sont appelés κ-, ι-, et λ-caraghénane. Ce sont des polysaccharides dont l'unité répétitive est formé de D-galactose lié par une liaison β-1,4 à un ose bicyclique, le 3,6-anhydro-D-galactose. Le κ-carraghénane qui forme des gels dur et cassant est uniquement sulfaté en position 4 du D-galactose. Le ι-carraghénane qui gélifie en donnant des gels mous est sulfaté en position 4 de l'unité D-galactose et en position 2 de l'unité 3,6-anhydro-D-galactose (Van de Velde et al., 2002). Enfin, Le λ-carraghénane est sulfaté en position 2 du D-galactose et en position 2 et 6 d'un ose qui, du fait de la sulfatation sur le carbone 6, perd le pont 3,6-anhydro. Le λ-carraghénane est donc composé d'un D-galactose 2-sulfate lié à un D-galactose 2,6-disulfate et qui, de ce fait, ne gélifie pas. Les polysaccharides testés ont été préparés de façon industrielle et ils sont donc très proches des motifs décrits plus haut, qui sont des motifs idéaux (**Figure 47**).

Les fucanes et les fucoidanes forment un ensemble de polysaccharides appelés FCSP (pour *Fucose-Containing Sulfated Polysaccharides* ; Deniaud-Bouët et al., 2017). Les fucanes sont des polymères ramifiés dont la chaine principale est constituée quasiment

exclusivement de L-fucose sulfaté. En général, ils sont présents dans la matrice extracellulaire d'animaux marins comme les échinodermes. Les fucoïdanes se distinguent des fucanes par la présence d'autres sucres qui accompagnent le fucose, celui-ci pouvant être minoritaire (Deniaud-Bouët et al., 2017). Les fucoïdanes sont des polysaccharides présents dans les parois cellulaires des algues brunes. Le fucoïdane désigné sous le nom de FS28 est un homofucane sulfaté provenant de *Pelvetia canaliculata,* dont le principal dissaccharide de répétition est le 4-α-L-fucose 2,3-disulfate-(1,3)-α-L-fucose 2-sulfate (Descamps et al., 2006 ; Colin et al., 2006). Les fucoïdanes de *Fucus serratus, de F. vesiculosus* et d'*Ascophylum nodosum* présentent aussi le même disaccharide de répétition (Berteau and Mulloy, 2003 ; Deniaud-Bouet et al., 2017).

Chez les algues vertes, le composé acide majoritaire est l'ulvane. Ce polysaccharides est constitué de la répétition d'un disaccharide composé de β-D-acide uronique lié par une liaison α(1-4) à du L-rhamnose-3-sulfate (Lahaye and Robic, 2007). Selon la nature de l'acide uronique, on distingue l'ulvane de type A qui contient de l'acide glucuronique et l'ulvane de type B dans lequel l'acide glucuronique est remplacé par de l'acide iduronique (Lahaye and Robic, 2007). De plus, dans certains ulvanes, l'acide uronique peut être remplacé par du xylose qui peut, à son tour, être sulfaté ou non. Enfin, d'autres monosaccharides neutres (glucose ou galactose) peuvent être présents.

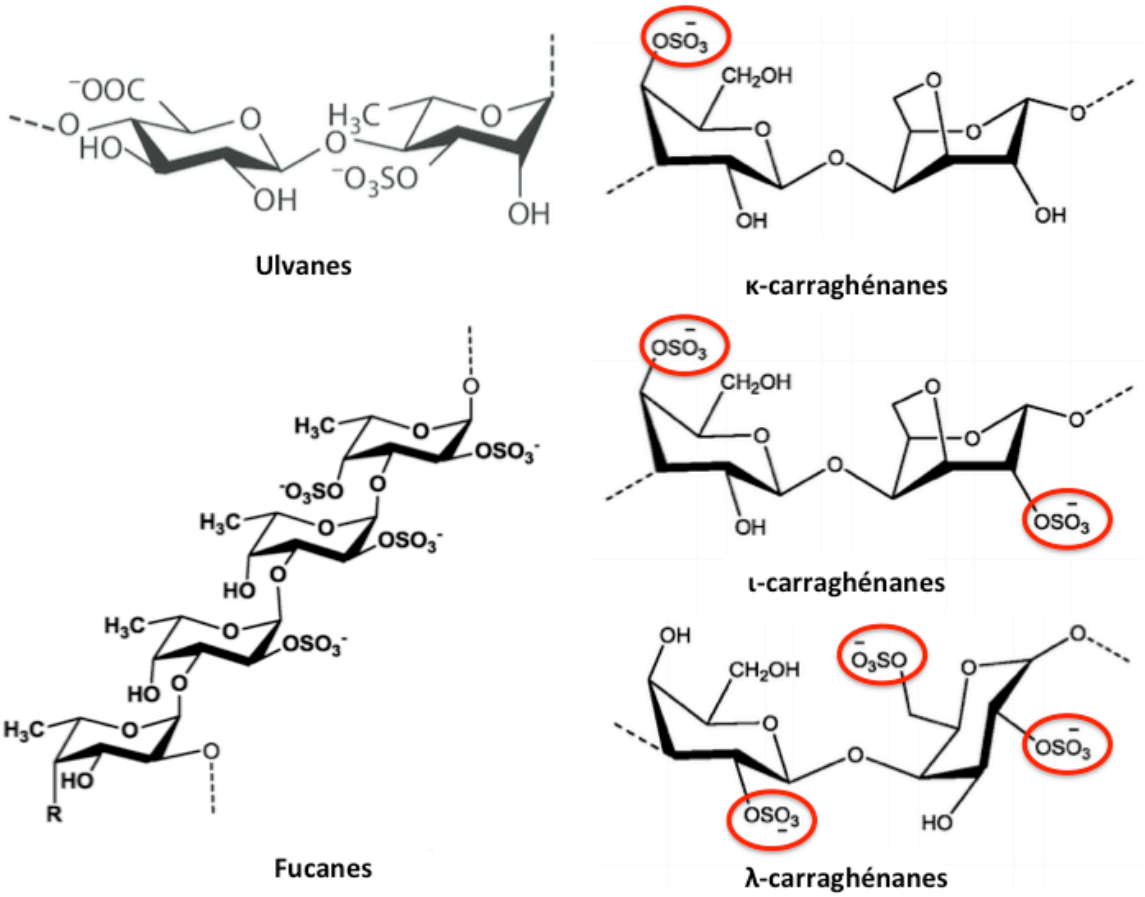


Figure 47 : Schéma des différentes unités de répétition des polysaccharides utilisés.

Les sulfatations essentielles à la définition des carraghénanes ont été entourées. Le dimère représenté pour l'ulvane est le suivant : 4-β-D-Glucuronosyluronic acide-(1-4)-α-L-Rhamnose-3-sulfate avec des ions sodium comme contre ions. Le fucane est constitué de fucose lié par des liaisons α (1-3), des sulfatations et des méthylations sont apparentes en position 2 et 5 respectivement.

Le xylane utilisé provient de *Palmaria palmata* et est composé de xyloses liés entre eux par une alternance de liaison β-1,3/β-1,4 (Lahaye et al., 2003). La thèse du Dr. E. Deniaud (2002) a mis en évidence la présence de sulfate provenant de la paroi de *Palmaria palmata*. Néanmoins, la question de l'origine des sulfates reste ouverte : sont-ils issus des polysaccharides ou des protéines constitutives de la paroi ? Cependant, comme il est très probable que le xylane issu de *Palmaria palmata* soit sulfaté, ce substrat a également été testé.

Un des moyens pour connaître le substrat catabolisé par le PUL 37 de *Z. galactanivorans*, est de caractériser le substrat de la GH 30 *Zg*3516 ou celui de la sulfatase *Zg*3515 (qui devrait être le même). La forme recombinante produite dans *E. coli* de la sulfatase *Zg*3515 étant active, nous avons étudié le comportement vis-à-vis des

polysaccharides, à la fois de la sulfatase *Zg*3515 et et de la GH30_4 *Zg*3516. L'idée était que si nous découvrions le substrat de la glycoside hydrolase, alors nous aurions une idée plus précise du rôle joué par la sulfatase.

Cette étude avait un autre objectif. Comme *Zg*3515 est une sulfatase produite sous forme active par *E. coli*, le but était de pouvoir réaliser sa caractérisation enzymologique. Ces connaissances permettraient d'envisager de réaliser de l'ingénierie protéique sur les sulfatases pour que ces dernières deviennent des *trans*-sulfatases. L'idée étant que ces enzymes permettraient une sulfatation sur des substrats spécifiques mais surtout que nous pourrions utiliser un donneur de sulfate moins instable et moins couteux à produire que le 3'-Phosphoadenosine-5'-phosphosulfate (PAPS), qui est le donneur de sulfate utilisé par de nombreuses sulfotransférases (Chapman et al., 2004).

V. 2 Résultats

V. 2. A Etude de la glycoside hydrolase Zg3516

V. 2. A. a Purification et contexte génétique de la GH30_4 *Zg*3516

La protéine *Zg*3516 appartient à la famille 30 des glycosides hydrolases, sous-famille 4 (GH30_4). La protéine native contient un signal d'adressage périplasmique. Il est composé de 19 acides aminés qui ont été éliminés et remplacés par une étiquette de 6 histidines.

La glycoside hydrolase *Zg*3516 a les caractéristique calculée suivant : une masse moléculaire x de 56,9 kDa, un coefficient d'absorption molaire (ε_{280}) de 120320 et un point isoélectrique de 6,99. Sa purification de s'est déroulée en tampon Bicine car la présence conjuguée de tampon Tris et d'imidazole avait conduit à une précipitation très importante de la protéine pendant l'étape de concentration. De plus, le tampon Tris est parfois un inhibiteur des glycoside hydrolases, (Kitamura et al., 2008). Après élimination de l'imidazole, la protéine a été concentrée facilement et était stable à 4°C. Le tampon Bicine a, par conséquent, été utilisé pour toutes les expériences postérieures aux premiers tests. Le rendement de purification de la protéine *Zg*3516 se situe entre 80 mg/mL et 100 mg/mL selon les préparations. L'électrophorèse en gel dénaturant de la protéine concentrée révèle la présence de plusieurs contaminants suite à l'étape de purification sur colonne de nickel (**Figure 48**). Sa température de dénaturation thermique a été mesurée en tampon Tris 50 mM pH 8, NaCl 400 mM à 40 °C par diffusion dynamique de la lumière (DLS, Malvern Instrument) (**Figure 49**).

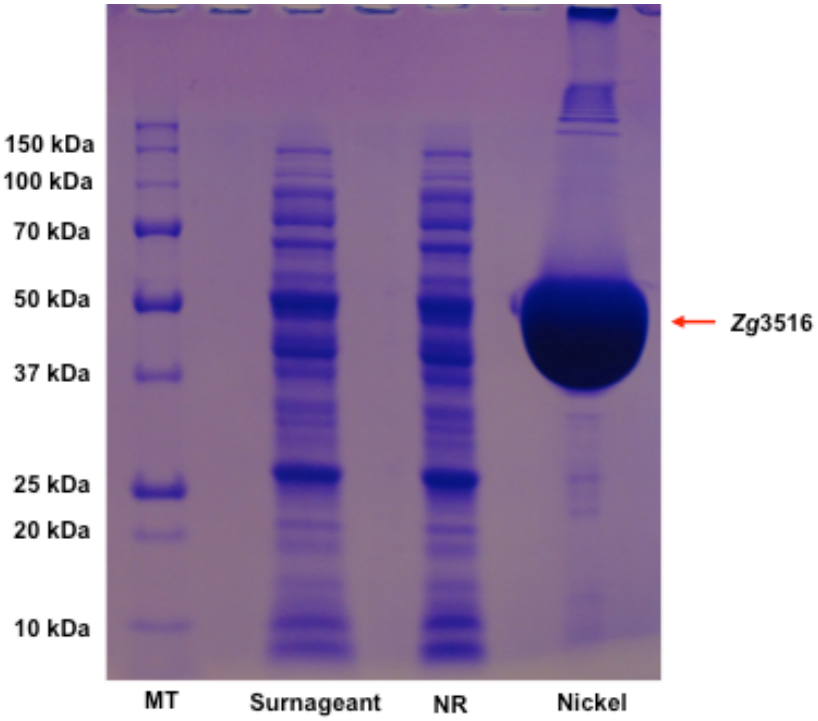


Figure 48 : Electrophorèse SDS-PAGE de *Zg*3516 après purification par chromatographie d'affinité.

Selon le marqueur de taille (MT), la GH30_4 *Zg*3516 migre au environ de 50 kDa. Surnageant : extrait protéique brut. NR : fraction non retenue par la colonne lors l'injection de l'extrait protéique brute. . Nickel : Protéine obtenu après élution par un gradient d'imidazole. Electrophorèse réalisée en gel d'acrylamide 12%, révélé au bleu de Coomasie.

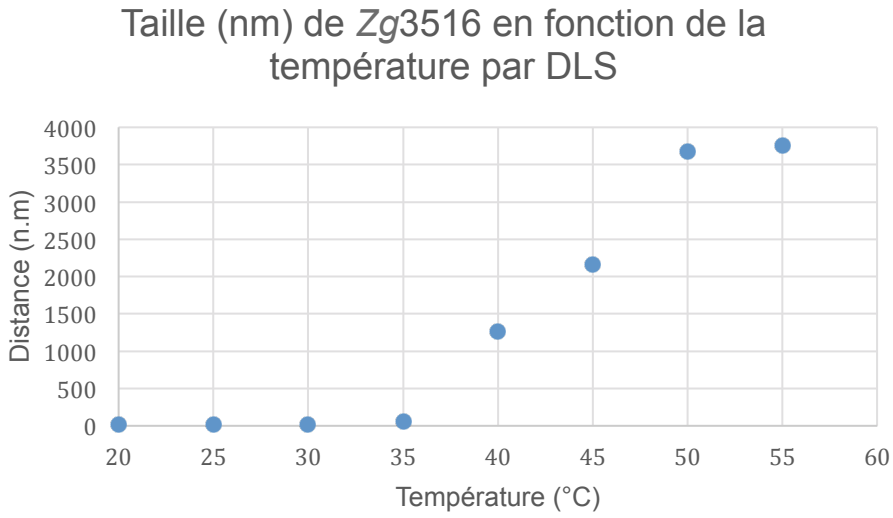


Figure 49 : Profil de dénaturation thermique de *Zg*3516 mesuré par diffusion dynamique de la lumière

Le profil a été obtenu par DLS (Malvern instrument), une augmentation nette du rayon hydrodynamique de la protéine (axe nommée distance) est visible dès 40°C mais une très légère augmentation est déjà visible à 35°C.

V. 2. A. b Recherche des substrats naturels de *Zg*3516

Afin d'identifier un substrat pour la GH30_4 *Zg*3516, nous avons d'abord testé sur série de substrats synthétiques (cf. Matériel et méthode, I) (p-nitrophényl-β-D-xylobioside, p-nitrophényl-α-L-fucopyranoside, p-nitrophényl-α-D-glucopyranoside et 4-méthylumbelliferyl-β-D-glucopyranoside). Malheureusement, aucune activité n'a pu être détectée avec ces substrats. Cette absence d'activité sur des substrats simples (notamment sur le p-nitrophényl-β-D-xylobioside) suggère que l'enzyme pourrait ne reconnaître que des oligosaccharides plus grands qu'un disaccharide.

Pour les essais sur les polysaccharides naturels, j'ai d'abord utilisé les polysaccharides de *Palmaria palmata*, car cette algue était connue contenir des xylanes sulfatés (Deniaud-Bouet, 2002). Cependant, différents polysaccharides ont été testés en parallèle, à savoir des carraghénanes (ι, κ et λ), des fucoïdanes de différentes origines (*Ascophyllum nodosium*, *Laminaria hyperborea* et FS28) et du porphyrane. Malheureusement, aucune activité n'a été détectée, avec la technique des gels d'oligosaccharides FACE, quel que soit le polysaccharide utilisé comme substrat.

Nous avons alors élargi le panel d'oligosaccharides, en sélectionnant des algues agarophytes qui avaient été utilisées pour les études concernant *Zg*AgaC. De plus, nous avons aussi récolté des algues présentes sur l'estran de Roscoff : *Ulva* sp. (ulvane), *Fucus spiralis*, *F. vesiculosus*, *F. serratus*, *Saccharina latissima* (fucoïdane) et *Chondrus crispus* (carraghénanes). Au total, ce sont 22 algues dont les polysaccharides ont été incubés en présence de la GH30_4 *Zg*3516 par électrophorèse sur gels FACE. Là encore, aucune activité glycoside hydrolase n'a pu être détectée. Deux possibilités peuvent expliquer cet échec. La première est que le substrat de *Zg*3516 ne soit pas un des polysaccharides testés. La deuxième est que l'enzyme ne serait pas endolytique mais exolytique, générant que des petits oligosaccharides, tous de la même taille, invisibles car masqués par le fluorophore.

Nous avons donc réalisé des chromatographies sur couche mince (CCM) qui utilise un mode de séparation basé sur la solubilité différentielle dans un mélange de solvant. Ces expériences ont permis de mettre en évidence des différences de migration en présence

de *Zg*3516 pour certains polysaccharides extraits d'algues agarophytes : *Osmundea pinnatifida*, *Cryptopleura ramosa*, *Acrosorium ciliolatum*, *Vertebrata lanosa*, *Polysiphonia elongata* et *P. simulans*. Mais c'est avec les extraits de *Porphyra* sp. et *Gracillaria* sp. que nous avons observé les différences les plus importantes, révélant ainsi une activité de *Zg*3516 sur les polysaccharides de ces algues (**Figure 50**). Disposant déjà de porphyrane digéré par différentes β-agarases et β-porphyranases de la famille GH16, nous avons aussi incubé *Zg*3516 avec ces différentes fractions et analysé les résultats par électrophorèse FACE (**Figure 51**). Nous avons ainsi mis en évidence une activité de *Zg*3516 sur les oligosaccharides produits par *Zg*PorB et *Zg*AgaB.

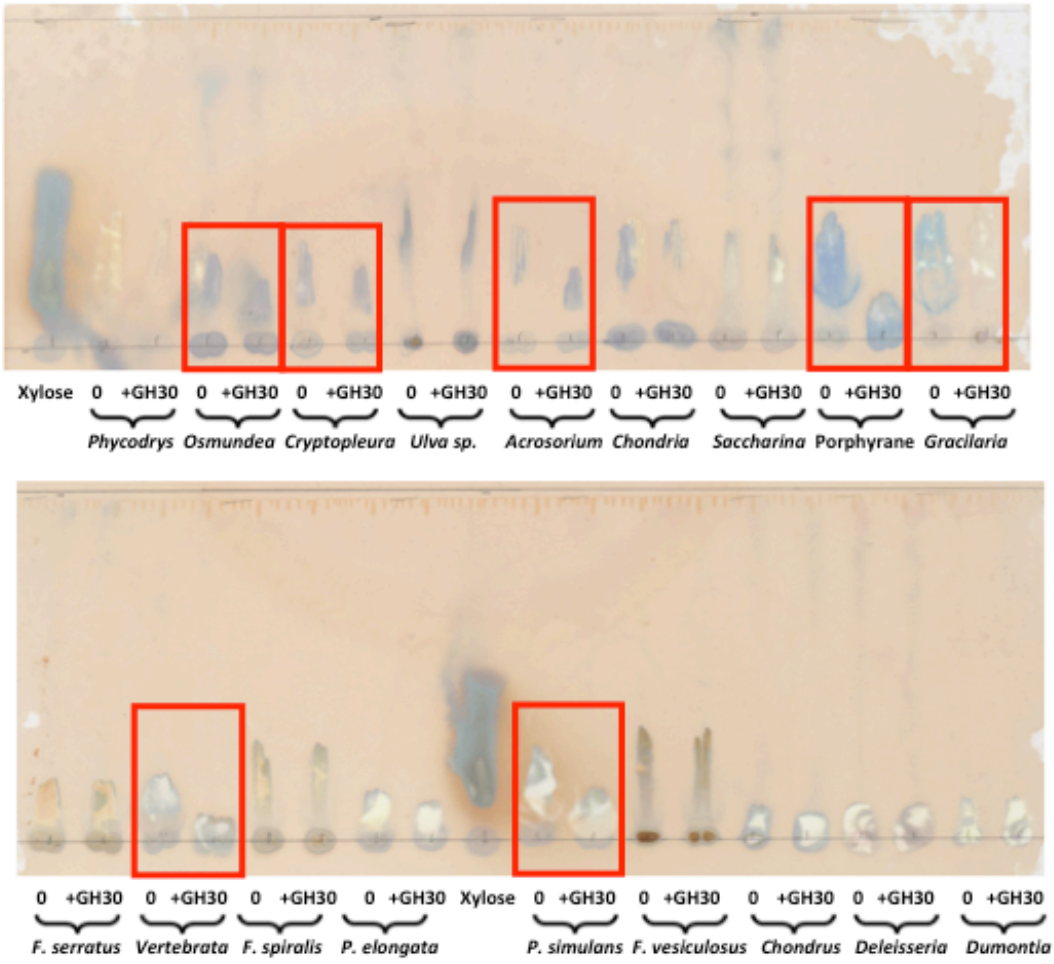


Figure 50 : Criblage de l'activité de *Zg*3516 par CCM

Les oligosaccharides avant (0) et après incubation avec la GH30_4 *Zg*3516 sont séparés par CCM où la phase mobile est constituée d'un mélange d'éthanol, de butanol et d'eau (5 :5 :3). Les algues sur lesquelles l'enzyme a été testé sont indiquées sous les pistes. Dans un souci de clarté seuls les ordres ont été indiqués, sauf pour les *Fucus* et les *Polysiphonia* où l'espèce a été indiquée. Les algues utilisées sont les suivantes : *Phycodrys rubens*, *Osmundea pinnatifida*, *Cryptopleura ramosa*, *Ulva* sp., *Acrosorium uncinatum*, *Chondria dasyphylla*, *Saccharina latissima*, *Gracillaria* sp., *Vertebrata lanosa*, *Chondrus crispus*, *Deleisseria sanguinea* et *Dumontia contorta*. Les pistes où les différences entre absence et présence de *Zg*3516 sont les plus notables sont encadrées en rouge.

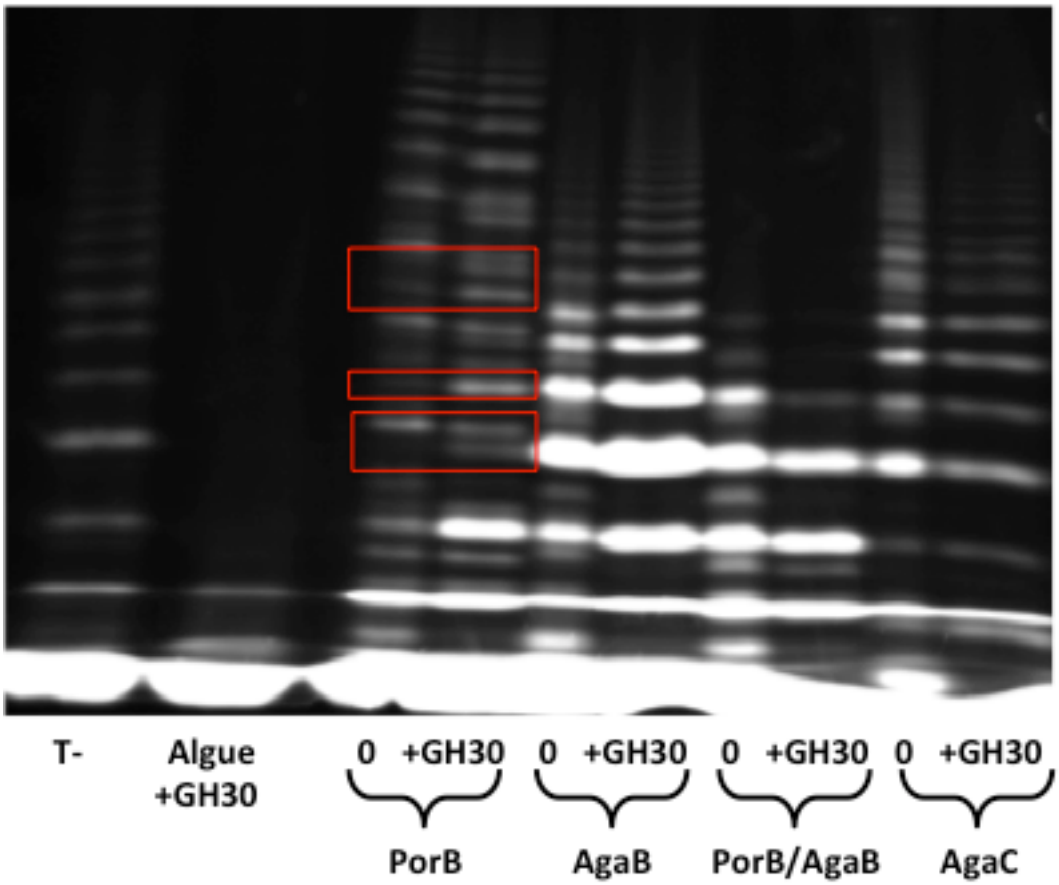


Figure 51 : Gels d'électrophorèse FACE mettant en évidence l'activité de *Zg*3516 sur des oligo-porphyranes.

L'enzyme a été incubé en présence des différents oligoporphyranes générés par *Zg*PorB, *Zg*AgaB, *Zg*AgaB/*Zg*PorB ou par *Zg*AgaC (l'enzyme utilisée sur le porphyrane extrait de *Porphyra* sp. est indiquée sous les pistes). L'action de *Zg*3516 (nommée GH30 sur la figure) est particulièrement visible sur les oligosaccharides générés par *Zg*PorB et *Zg*AgaB. En effet, sur les pistes concernées, une disparition de bandes intervient après ajout de l'enzyme. Le témoin correspond à une incubation de *Porphyra* sp. cryo-broyés dans du tampon. La piste algue + GH30 correspond au témoin incubé en présence de *Zg*3516. Le marquage a été réalisé avec de l'ANTS. Les encadrés rouges signalent les différences les plus importantes entre le contrôle et le test avec *Zg*3516.

V. 2. A. c Caractérisation de l'activité de *Zg*3516.

Pour comprendre l'action de la GH30_4 *Zg*3516, nous l'avons incubée sur une collection d'oligoagaroses (DP4, DP6, DP8-DP10 et DP10) et d'oligo-porphyranes (DP4 et DP8) semi-purifiés (NB : la dénomination correspond à l'oligosaccharide majoritaire ; des oligosaccharides minoritaires de taille proche étant présent). Ces expériences ont révélé que *Zg*3516 n'avait pas d'activité sur les oligo-porphyranes ni sur les oligo-agarotetraose (DP4). Au contraire, quand *Zg*3516 est incubé en présence d'oligo-agarohexaose (DP6), il apparaît deux bandes de tailles inférieures dont une migre à la même taille que l'oligo-agarotetraose (DP4), la deuxième bande étant probablement

unoligo-agarobiose (DP2)(**Figure 52**). Avec les oligo-agarooctaoses (DP8) et oligo-agarodecaose (DP10), les mêmes bandes apparaissent. Des oligosaccharides de tailles supérieures, qui n'étaient pas présents dans le témoin, sont visibles dans ces dernières expériences. Il est probable que ces oligosaccharides de grande taille soient présents en faible quantité dans les témoins et non visibles car retenues dans les puits. Ces données semblent suggérées que *Zg*3516 génère des oligo-agarobioses (DP2) comme produits à partir d'oligosaccharides de tailles supérieures au DP4. Ces résultats montrent que *Zg*3516 est active sur oligo-agaroses et est probablement une-exo-β-agarase (β-agarosidase) libérant essentiellement (voire exclusivement) du DP2, l'action hypothétique de Zg3516 est schématisée **Figure 53**. Cette activité n'avait encore jamais été reportée pour la famille 30 des glycosides hydrolases.

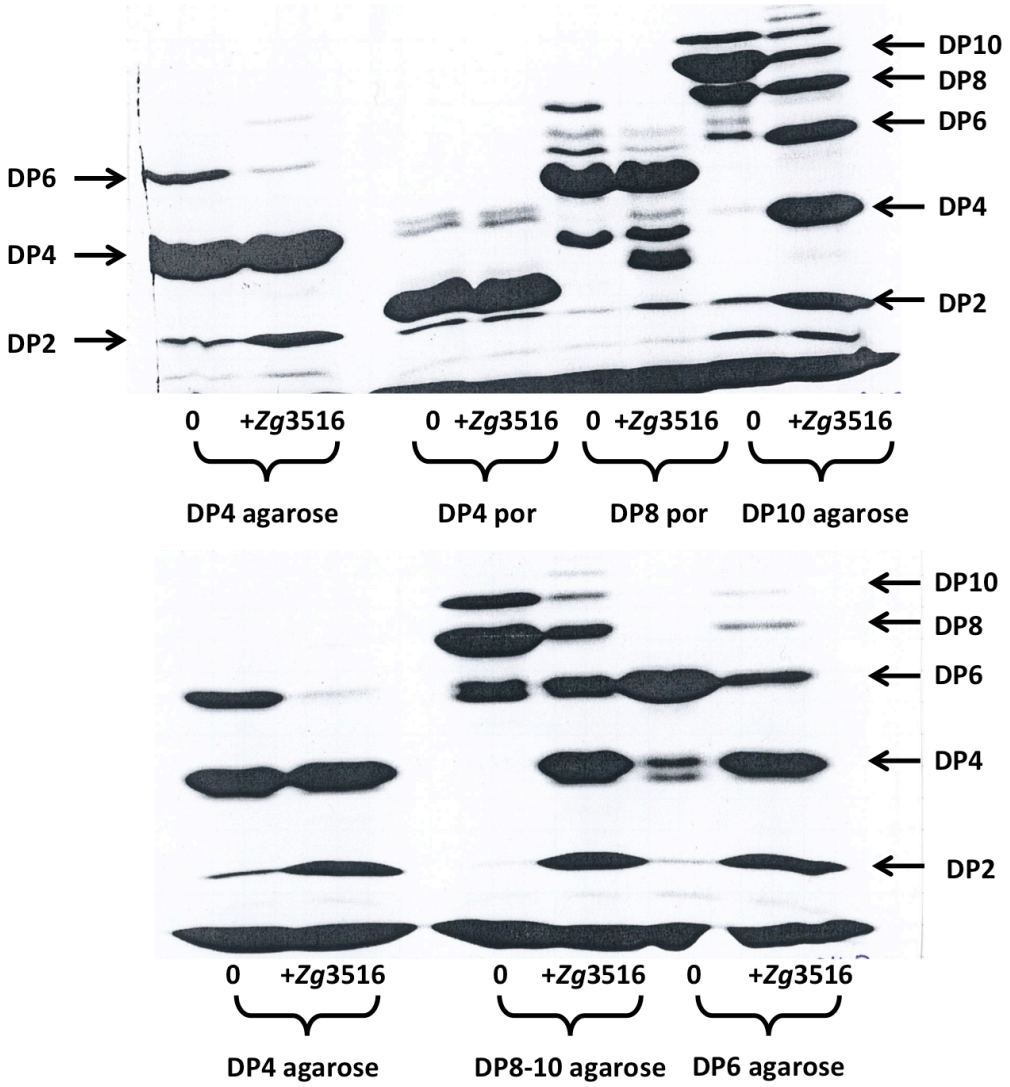


Figure 52 : Gels d'électrophorèse FACE de *Zg*3516 sur différents oligo-porphyranes et oligo-agars.

Les différents degrés de polymérisation des oligosaccharides sont indiqués sur le côté. Ces valeurs ne sont pas valables pour les oligo-porphyranes (nommées par l'abréviation por) car ils possèdent des groupes sulfates qui altèrent la migration des oligo-porphyranes. Les degrés de polymérisation indiqués au bas de chaque piste correspondent à l'espèce majoritaire de la fraction purifiée.

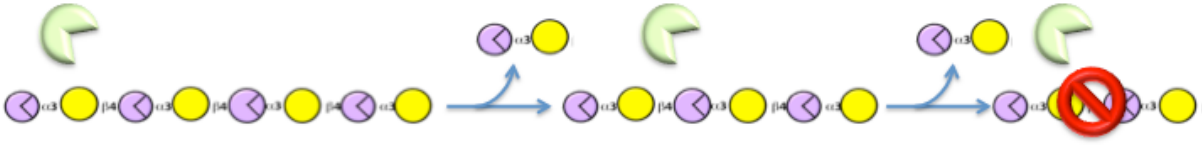


Figure 53 : Action putative de *Zg*3516 sur des oligoagaroses

Les oligoagaroses sont schématisés selon le code suivant : le rond jaune symbolise le D-galactose et le rond violet symbolise le L-galactose, le triangle au niveau du L-galactose indique la présence d'un pont 3,6 anhydro. Zg3516 aurait pour substrat de long oligoagarose, l'enzyme agirait à partir de l'une des extrémités et libérerait un agarobiose. L'enzyme ne serait pas capable d'agir sur un DP4.

L'ordre d'action de la GH30_4 *Zg*3516 et de la sulfatase *Zg*3515 a été recherché en utilisant une fraction d'oligo-porphyrane comme substrat, purifié par chromatographie

d'exclusion. Cette fraction est substrat de la sulfatase *Zg*3515, il s'agit de la fraction 50 (cf. V. 2. B. d). L'éléctrophorèse FACE révèle que la *Zg*3516 génère des oligosaccharides que si le substrat a d'abord été incubé avec la sulfatase. Ainsi, il semble que les oligosaccharides doivent d'abord être désulfatés avant que *Zg*3516 puisse agir (**Figure 54**).

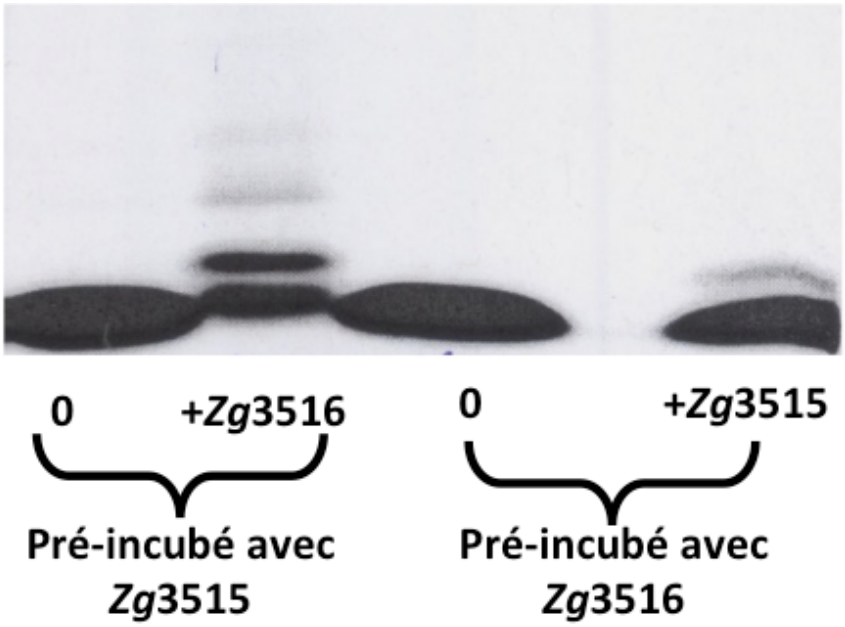


Figure 54 : Mise en évidence de l'ordre d'action de *Zg*3515 et *Zg*3516
Electrophorèse FACE réalisées avec un marquage ANTS. La fraction F50 de porphyra purifiée par chromatographie d'exclusion a été incubée soit avec *Zg*3515 soit avec *Zg*3516, après avoir été pré-incubée avec *Zg*3516 ou *Zg*3515. Dans le cas où les oligosaccharides ont été pré-incubés avec *Zg*3515, plusieurs petits oligosaccharides apparaissent.

V. 2. A. d Structure de *Zg*3516

Des cristaux de *Zg*3516 en boite Linbro selon la technique de la goutte suspendue ont été obtenu dans les conditions suivantes : 0,1 M d'acétate de sodium pH 6, 0,01 M de $ZnCl_2$, 21% de PEG3350 (w/v). Les gouttes ont été réalisées avec 2 µL de protéine à 30 mg/mL et 1 µL de précipitant au-dessus d'un réservoir contenant 500 µL de précipitant. L'optimisation pour l'obtention des cristaux a été réalisée par A. Jeudy. Les données ont été collectées sur la ligne ID29 de l'ESRF à Grenoble par le Dr. M. Czjzek. La structure de *Zg*3516 a été obtenue à une résolution de 1,9 Å dans le groupe d'espace $P6_122$. Deux molécules sont présentes dans l'unité asymétrique. Le phasage a été compliqué par le fait que ces deux molécules sont reliées par une opération de translation/rotation. Cette translation/rotation supplémentaire rendait problématique la recherche des fonctions de translation et rotation permettant de placer le modèle lors de l'utilisation des

programmes de remplacement moléculaire. En ajustant les paramètres disponibles dans le programme Molrep et en utilisant un modèle tronqué (le tonneau TIM d'une GH30 putative de *Bacteroides fragilis* NCTC 9343^T (PDB 3CLW, *Structural genomics project*), j'ai pu obtenir une solution partielle où il manquait au modèle environ 500 résidus, qui correspondent aux deux domaines C-terminaux (les β-sandwich) du dimère cristallographique de *Zg*3516.

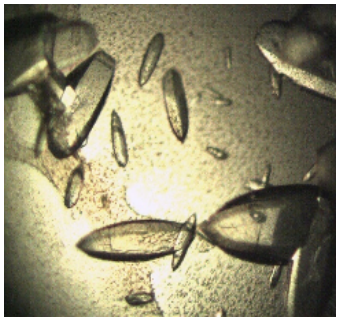


Figure 55 : Cristaux optimisé de la GH30_4 *Zg*3516

Les cristaux ont été obtenus grâce à la technique de la goutte suspendue dans les conditions suivantes : 0,1M d'acétate de sodium pH6, 0,01 M de $ZnCl_2$, 21 % de PEG3350 (w/v). Ce sont les plus petits cristaux qui permettent la meilleure diffraction.

	***Zg*3516**
Collection des données	
Groupe d'espace	P 6_1 2 2
Dimensions de l'UA	
a, b, c (Å)	97.42 97.42 432.14
α, β, γ (°)	90 90 120
Résolution (Å)	49.82 - 1.93 (1.999 - 1.93)
R_{merge}/R_{pim}	0.23(1.0)/0.06(0.30)
I/σI	9.9 (2.1)
CC(1/2)	0.99 (0.74)
Complétude (%)	99.94 (99.99)
Redondance	16.5 (13.7)
Affinement	
Résolution (Å)	1.93
Nombre de réflexions	92453 (9054)
R/R_{free}	0.1883/0.2243
Nombre d'atomes	8306
Protéines	7806
Ligand/ion	58
Eau	442
Facteur B	
Protéines	30.99
Ligand/ion	46.39
Eau	34.66
R. m. s. déviations	
Longueur de liaisons (Å)	0.013
Angles (°)	1.58

Tableau 7 : Statistiques des données de collecte et d'affinement pour le cristal ayant permis l'obtention de la structure de *Zg*3516

Les valeurs entre parenthèse concernent les données pour la plus haute résolution.

Pour la construction du modèle, les hélices α ou les brins β manquants ont été placés dans la densité électronique sous forme de polyalanine. A la suite de quoi le logiciel ARP/wARP a été utilisé pour améliorer l'ajustement à la densité électronique. Les alanines ont alors été remplacées par les résidus de la séquence peptidique lorsque leur chaine latérale était visible. J'ai répété cette stratégie jusqu'à l'obtention d'un modèle le plus complet possible. Ce modèle a ensuite été utilisé pour phaser le jeu de données originel. La dernière carte ainsi obtenue a permis d'obtenir le modèle final avec les facteurs de qualité suivants R= 0,22 et R_{free}= 0,26. Actuellement, seule une boucle incluse dans le domaine C-terminale de la chaine B est incomplète, en effet, il manque trois résidus entre la S465 et la S469. Le repliement de *Zg*3516 est similaire à celui des autres structures déterminées au sein de la famille GH30. Le domaine N-terminal est composé

d'une structure de type $(\alpha/\beta)_8$ ou tonneau TIM et d'un domaine C-terminal plus petit composé de deux feuillets β antiparallèles superposés formant un repliement de type β-sandwich. Le tonneau TIM se présente sous la forme d'un feuillet de 8 brins β parallèles formant un tunnel central entouré par huit hélices α. Quant au domaine C-terminal, son feuillet au contact du solvant contient 3 brins β, tandis que celui positionné contre le tonneau TIM comprend 7 brins β, et contient également une hélice α constituée de 6 résidus, qui n'avait pas encore été observée chez les autres structures de GH30 disponibles. Ce domaine C-terminal est positionné sur le côté du tonneau TIM et est relié à ce dernier *via* le brin β1 et l'hélice α8. D'après la topologie des structures secondaires, cette GH30_4 ne fait partie d'aucun des deux groupes structuraux définis selon St John et collaborateurs (2010). Nous suggérons donc de former un troisième groupe en plus de ceux déjà décrits précédemment (**Figure 56**). Le travail présenté ici montre que la phylogénie de 2010, qui reposait sur l'identité de séquence globale, est potentiellement biaisée par le tonneau TIM. En effet, ce dernier est très conservé et est constitué par une séquence continue située au centre de la structure primaire de la protéine. Ainsi les extrémités N et C-terminales, qui constituent le domaine β-sandwich, ont pu être éliminées lors de l'alignement et donc biaiser la définition des sous-familles.

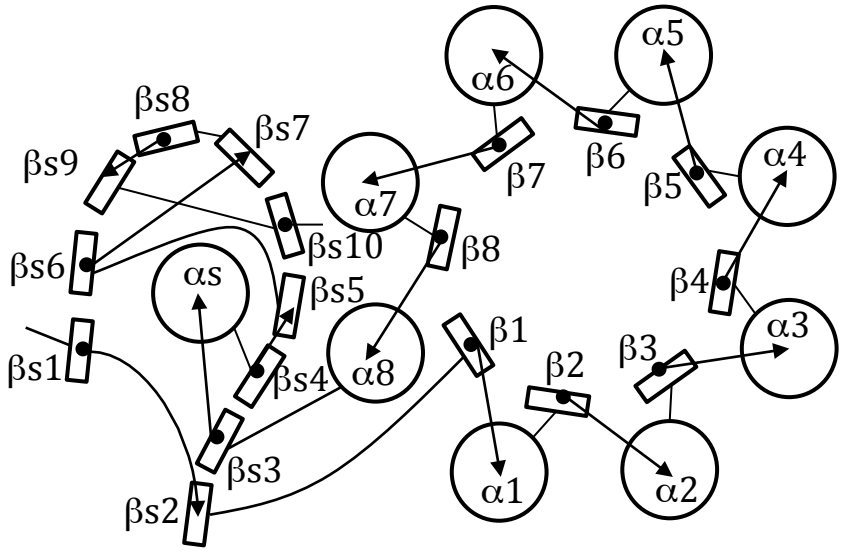


Figure 56 : Organisation des structures secondaires chez *Zg*3516

Comme montré sur la figure, l'organisation du domaine β sandwich est différente de ce qui est observable dans les deux groupes précédemment décrites. En effet le β-sandwich est relié au tonneau TIM par le brin β2 et non par les brins β1 ou 3 comme pour les autres groupes. Autre particularité remarquable, la présence d'une hélice α au sein du β-Sandwich et un brin supplémentaire à ce dernier.

Les résidus catalytiques se trouvent au niveau de l'extrémité C-terminale des brins β4 (E204) et β7 (E327) du tonneau TIM comme pour les autres GH30 (**Figure 57**). Ce site

actif a été comparé avec celui de son plus proche homologue ayant été cristallisé avec son substrat : la glucuronoxylanase de *Dickeya chrysanthemi* D1(anciennement *Erwinia chrysanthemi* (PDB : 2Y24)(Urbániková et al., 2011). Entre ces deux sites actifs, outre les résidus catalytiques, sont conservés uniquement 4 résidus sur les 11 impliqués dans la reconnaissance du substrat (Urbániková et al., 2011). Nous avons donc vraisemblablement une grande divergence dans la reconnaissance des substrats. L'analyse de la surface à proximité ne révèle pas de grande gorge mais une légère dépression au fond de laquelle se trouve la paire catalytique. De part et d'autre de cette dernière se trouvent deux zones chargées positivement dues à la présence, d'une part de la lysine 103 et, d'autre part des lysines 338 et 340 (**Figure 58**).

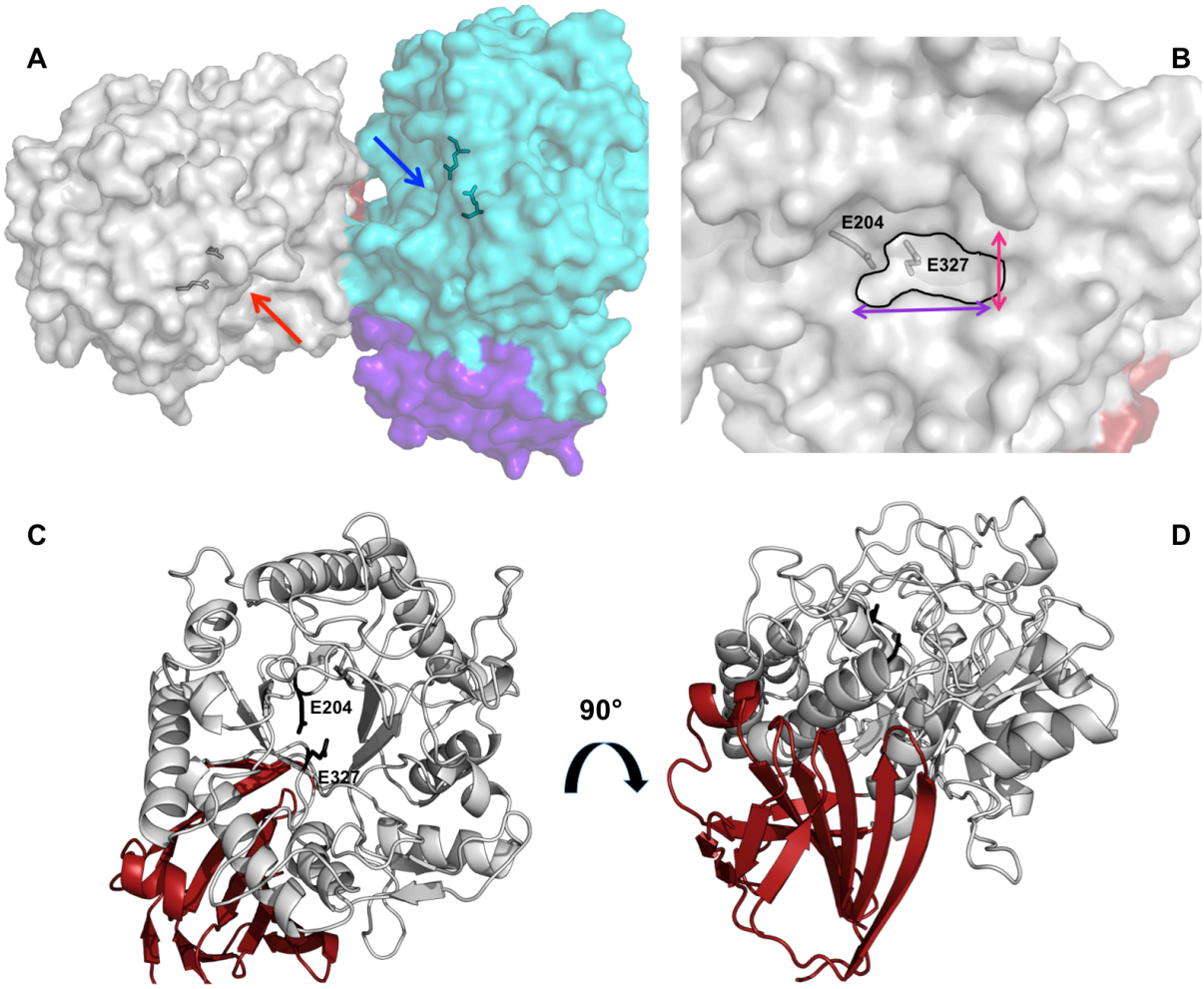


Figure 57 : Les caractéristiques structurales de *Zg*3516

A : Dimère cristallographique de *Zg*3516, le monomère A (en bleu) a son site actif (flèche bleu) partiellement obstrué par le monomère B (en blanc) qui lui a son site actif beaucoup plus accessible (flèche rouge). Les monomères sont reliés entre eux par une fonction de translation/rotation.

B : Visualisation de la surface atomique au niveau du site actif du monomère B. Les résidus catalytiques sont représentés en bâton. La poche du site actif est soulignée par un trait noir, sa longueur (double flèche violette) est de 15,6 Å, la hauteur (double flèche rose) de 8,2 Å et la profondeur de 6,2 Å.

C : Visualisation du tonneau TIM (en blanc) et des résidus catalytiques (en représentation bâton et en noir).

D : Vues de profil de *Zg*3516. Le sandwich β (en rouge) comprend une petite partie structurée en hélice.

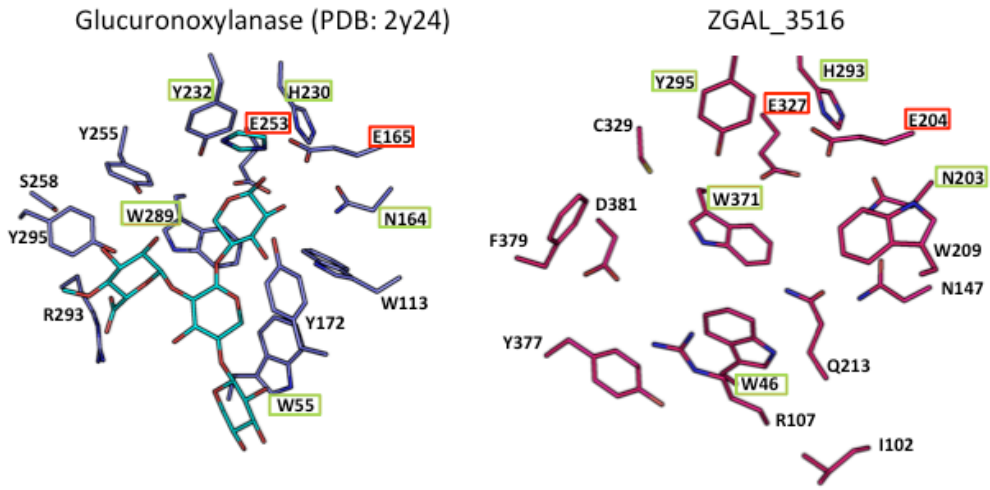


Figure 58: Comparaison des sites actifs de *Zg*3516 avec celui de la glucuronoxylanase de *Dickeya chrysanthemi* D1.(PDB : 2Y24)

La structure de la glucuronoxylanase de *Dickeya chrysanthemi* D1 est la structure en complexe la plus proche de *Zg*3516 (25 % d'identité sur 65% de la séquence). Les acides aminés encadrés en vert sont les résidus conservés au sein des deux structures et les acides aminés encadrés en rouge sont les résidus catalytiques des enzymes. Tous les résidus affichés correspondent aux acides aminés nécessaires pour la fixation du substrat de la glucuronoxylanase et leur équivalent chez *Zg*3516.

Plusieurs boucles se situent à proximité du site actif et pourraient éventuellement intervenir dans la reconnaissance du substrat, mais les facteurs b sont de l'ordre de 20 à 30 indiquant une faible mobilité des atomes par rapport à leur position moyenne. Ceci suggère une faible mobilité de ces boucles au sein des cristaux. Cependant, la conformation protéique est contrainte dans le cristal et seule une comparaison entre la protéine seule et un complexe nous permettrait de définir le rôle de ces boucles.

V. 2. A. e Obtention d'un complexe *Zg*3516-oligoporphyranes

Les oligo-porphyranes n'etant pas substrats, et notamment les oligo-porphyranotetraoses (DP4), nous avons décidé de les utiliser pour obtenir la protéine native en complexe avec l'oligosaccharide. Par trempage, nous avons obtenu un complexe à la résolution de 2,4 Å. A proximité des résidus du site actif (E204 et E327), des densités électroniques qui n'étaient pas présentes dans la structure de l'enzyme seule ont été observées. Un motif porphyranobiose a pu être correctement placé au sein de la chaine B. Il est difficile de placer le second motif porphyranabiose dans la continuité du précèdent, car la densité est moins nette. A ce stade de l'affinement, seule une molécule de galactose a pu être placée au sein de la chaine A. Néanmoins, l'accès au

site actif du monomère A est compliqué par la proximité du monomère B ce qui peut expliquer cette observation. En effet, si l'on a moins d'oligo-saccharides qui ont pu diffuser jusqu'au site actif, l'occupation du site actif est partielle, donc la densité est mal définie.

L'orientation du motif porphyranabiose au sein du site actif montre, dans la poche, un sucre perpendiculaire à la surface, plaidant pour un mode d'action exolytique. Le porphyranobiose se trouve très près des résidus catalytiques, ce qui rend improbable que ce soit le positionnement adopté par un motif agarobiose. Néanmoins, le placement de ce DP2, nous permet de mieux visualiser l'organisation du site actif. Ainsi, nous pouvons émettre une première hypothèse sur les sites -1 et -2. En effet, situé à côté des acides aminés catalytiques (E204 et E327) se trouvent plusieurs résidus hydrophobes (W371, W46, L331 et F379) et deux acides aminés polaires (C329 et D381) qui pourraient permettre la stabilisation de l'ose. De plus, cette dépression est bloquée par la présence de la phénylalanine 379, ce qui fait que seul un DP2 pourrait y être stabilisé. Ces données vont dans le sens de ce que nous avons observé biochimiquement (**Figure 58**).

V. 2. B. Etude de la sulfatase *Zg*3515

V. 2. B. a Purification de la sulfatase *Zg*3515 et contexte génomique

La protéine *Zg*3515 appartient à la famille 1 des sulfatases, sous-famille 20 (S1_20) selon la base de données SulfAtlas. Elle possède un peptide signal d'adressage périplasmique. Lors du clonage, les 26 acides aminés de ce peptide signal ont été éliminés et remplacés par une étiquette 6 histidines.

La protéine *Zg*3515 sous sa forme recombinante a une masse moléculaire de 59,1 kDa, un coefficient d'absorption moléculaire (ε_{280}) de 58790 et un point isoélectrique de 6,43. La purification de *Zg*3515 s'est déroulée en deux étapes successives : une colonne d'affinité et une chromatographie d'exclusion de taille. Lors de la chromatographie d'exclusion de taille avec une colonne Hiload S200 16/60 (GE Healthcare), la protéine est élué au environ de 75 mL ce qui d'après la calibration de la colonne correspond à une taille approximative de 50 kDa, suggérant que la sulfatase se trouve sous forme monomérique. A partir d'un litre de culture, la quantité de protéine sature la colonne d'affinité et peu de contaminants sont visibles sur gel. Le rendement pour un litre de culture se situe aux environs de 80 mg/mL. Dans le cas de la protéine séléniée, le protocole de purification est le même et le rendement pour un litre de culture est de 60 mg/mL.

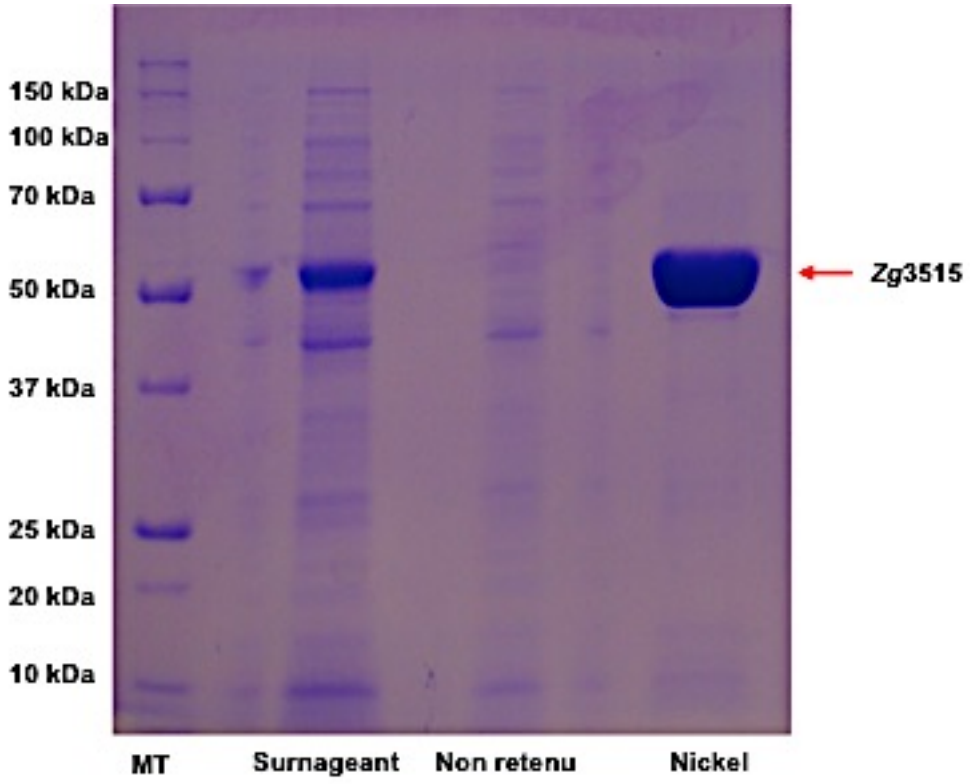


Figure 59 : Suivi par électrophorèse SDS-PAGE de la purification de *Zg*3515

La bande qui élue au niveau du marqueur de taille (MT) de 50 kDa correspond à *Zg*3515. Surnageant, correspond à l'extrait proteique brut. Non retenu, correspond à ce qui n'as pas été retenu lors de l'injection du de l'extrait proteique brut. Nickel, correspond à l'élution par un gradient d'imidazole. Electrophorèse en gel d'acrylamide 12%, révélé au bleu de Coomasie.

V. 2. B. b Caractérisation cinétique de *Zg*3515

Afin de caractériser la sulfatase *Zg*3515, nous avons voulu identifier le cation nécessaire à la catalyse. Cependant, après plusieurs essais pour chélater le cation, dont une incubation sur la nuit dans 100 mM d'EGTA et d'EDTA, l'activité de *Zg*3515 n'avait pas diminué par rapport à l'enzyme sans traitement. Ce resultat suggère que le cation est soit très fortement chélaté dans le site actif ou, plus improbable, qu'il n'est pas indispensable à la catalyse. La structure cristallogaphique (voir § V. 2. B. f) révèlera qu'un cation est bien présent dans le site actif.

Nous avons ensuite déterminé les paramètres cinétiques de *Zg*3515 en utilisant la MUS comme substrat à pH 8 (le pH de l'eau de mer est en moyenne de 8,3 à Roscoff). J'ai obtenu un k_{cat} de 4,593 e^{-2} s^{-1}. Le K_M était de 549,8 µM et l'efficacité catalytique de k_{cat}/K_M de 83,54 M^{-1}. s^{-1}.

La courbe de dépendance de l'activité de *Zg*3515 par rapport au pH indique que le pH optimum de *Zg*3515 avec la MUS utilisée comme substrat est de 7,3. Les tests pour obtenir les paramètres cinétiques de l'enzyme à ce pH n'ont pas encore pu être réalisés.

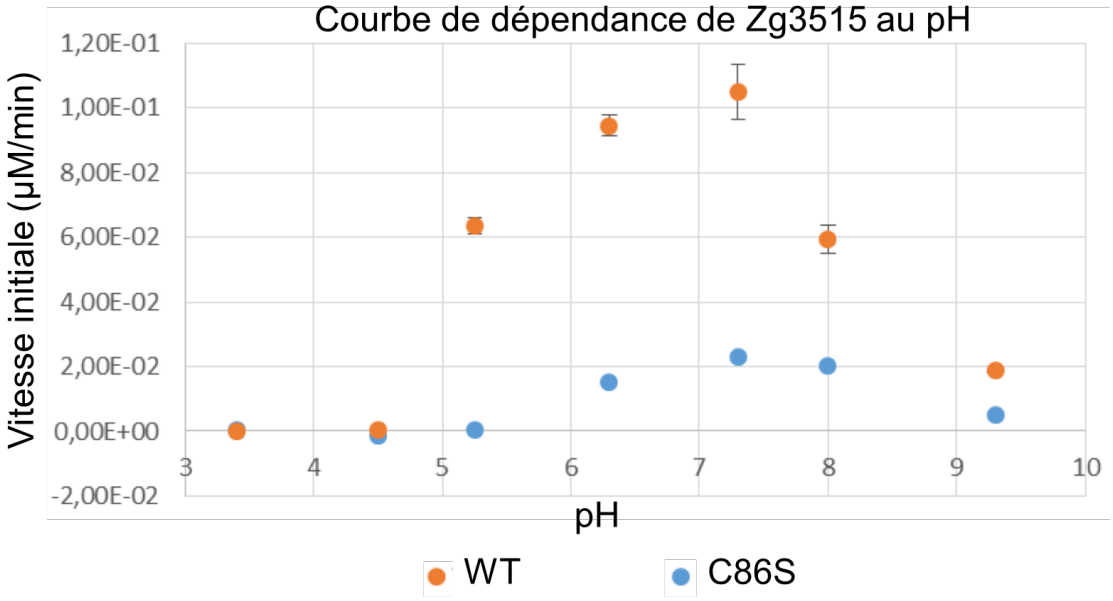


Figure 60 : Courbe de dépendance de *Zg*3515 et du mutant C86S par rapport au pH.

V. 2. B. c Première évidence de l'activité de *Zg*3515 sur substrats naturels

Suite aux résultats obtenus avec la GH30_4 *Zg*3516 sur des oligo-porphyranes, nous avons vérifié si la sulfatase pouvait désulfater ces oligosaccharides (**Figure 61**). La sulfatase *Zg*3515 est capable de désulfater les fractions générées par *Zg*PorB à partir du porphyrane. Pour caractériser cette activité, il était nécessaire de purifier les oligosaccharides substrats.

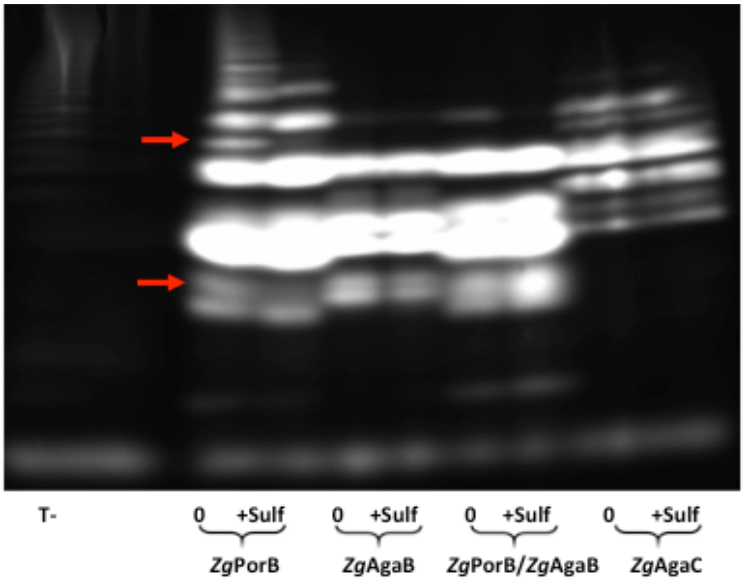


Figure 61 : Activité de *Zg*3515 sur oligo-porphyranes.

L'enzyme a été incubée avec différents oligo-porphyranes générés soit par *Zg*PorB, *Zg*AgaB, *Zg*AgaB/*Zg*PorB ou par *Zg*AgaC (l'enzyme utilisée est indiquée sous les pistes) à partir du porphyrane (extrait à partir de l'algue *Porphyra* sp.). La sulfatase est active sur les oligo-saccharides générés par *Zg*PorB, son action est soulignée par des flèches rouges sur le gel. Le marquage a été réalisé avec de l'AMAC.

V. 2. B. d Purification et caractérisation des oligo-porphyranes, désulfatés par *Zg*3515

Les oligosaccharides produits en incubant le porphyran de *Porphyra* sp. en présence de *Zg*PorB ont d'abord été purifiés par chromatographie d'exclusion de taille. Après électrophorèse, il apparaît que deux des fractions purifiés montrent nettement une activité sulfatase (**Figure 62**). Une seule de ces fractions, la fraction 50, a pu être obtenue en quantité suffisante pour une analyse en spectrométrie de masse.

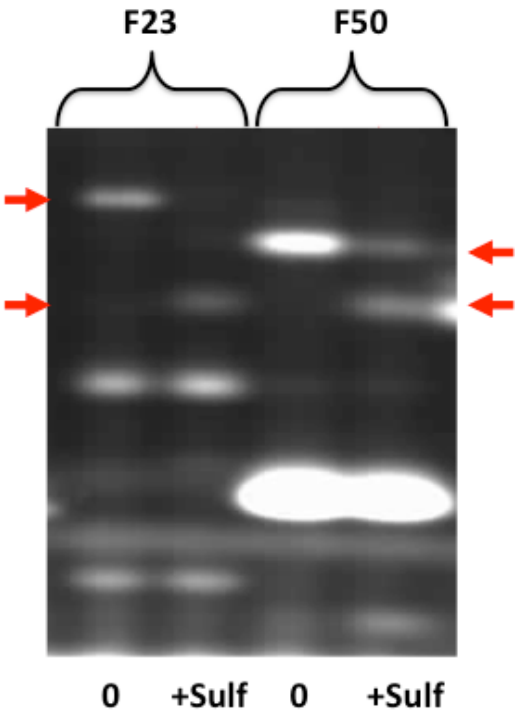


Figure 62 : Résultats des tests sulfatases sur des oligoporphyranes purifiés.
Les fractions 23 et 50 du premier test de purification ont permis d'observer, après ajout de la sulfatase, une différence de migration, ce qui indique que ces deux fractions sont substrat de *Zg*3515. Le marquage utilisé ici est de type AMAC. Les flèches rouges signalent les différences entre les pistes en présence de la sulfatase.

L'analyse par spectrométrie de masse a révélé que la fraction 50 comporte des DP4 et des DP6 composés de motifs agarobiose et porphyranobiose, présentant des sulfatations et parfois des groupements méthyles (**Figure 63**).

Après incubation avec la sulfatase, la composition de la fraction 50 contient les mêmes oligosaccharides DP4 et DP6 accompagnés de deux nouveaux oligosaccharides. Le premier est composé d'un motif porphyranobiose et d'un motif agarobiose avec un groupe méthyle, et correspondrait donc à un tetrasaccharide méthylé. Cet oligosaccharide est cependant très minoritaire. Le deuxième est également composé d'un motif porphyranobiose et d'un motif agarobiose, mais sans méthylation. Cependant, en comparant cet oligosaccharide avec ceux présents dans la fraction 50 avant l'action de la sulfatase, il semblerait que des activités enzymatiques contaminantes autres que l'activité sulfatase soit présentes. En effet, une contamination par une glycoside hydrolase pourrait expliquer l'élimination d'un motif agarobiose pour chacun des DP6. En conséquence, il apparaît que *Zg*3515 ne puisse probablement pas hydrolyser le groupement sulfate positionné sur le carbone 6 du L-galactose dans les oligoporphyranes. Cette sulfatation étant la modification majeure dans le porphyrane, si

la sulfatase avait agi sur le L-galactose-6-sulfate, nous aurions dû observer, en spectrométrie de masse, une diminution forte des pics correspondant aux motifs porphyranobiose. La présence de radicaux méthyles pourrait empêcher la désulfatation et expliquer pourquoi certains motifs porphyranobioses subsistent. Mais cette hypothèse est réfutée par la présence, après action de la sulfatase, d'un oligosaccharide mixte porphyranobiose/agarobiose sans groupe méthyle.

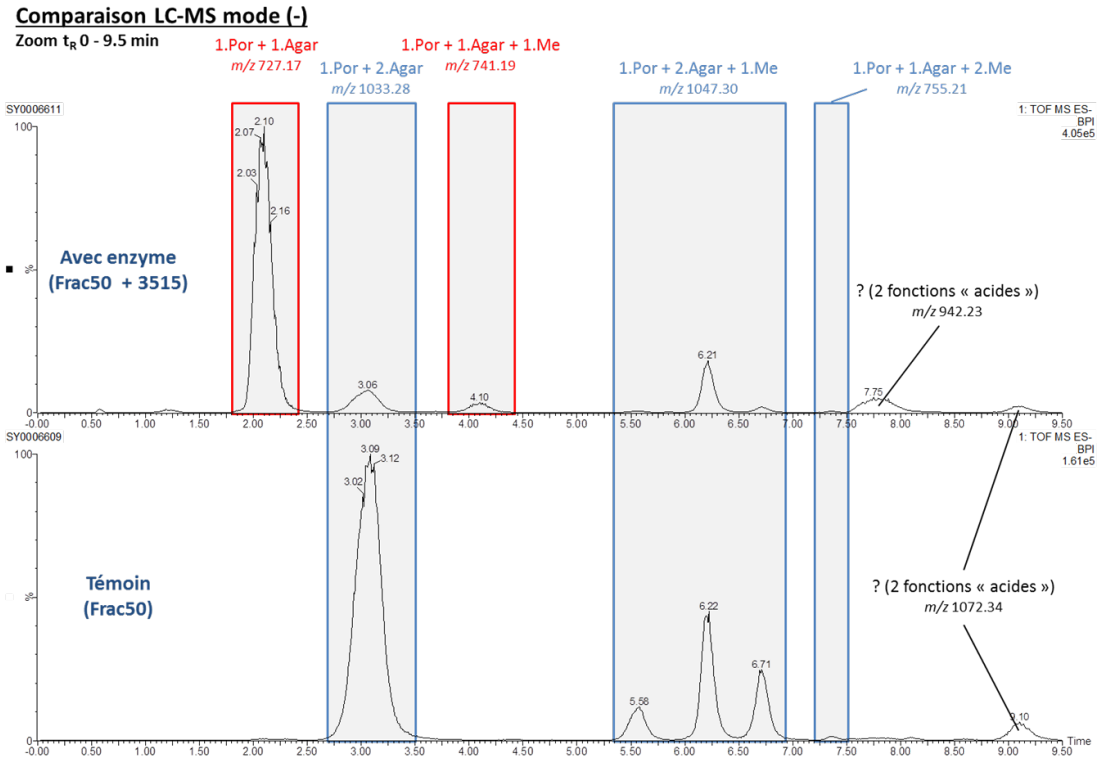


Figure 63 : Spectrogramme de masse LC-MS avant et après action de *Zg*3515

La comparaison des spectres montre qu'en présence de la sulfatase il y a ns une diminution de l'intensité des pics majoritaires présents dans la fraction 50 (en bleu) et l'apparition de nouvelles espèces (en rouge). agar : motif agarobiose (G-LA), por : motif porphyranabiose (G-L6S).

Dans le but d'améliorer la pureté des fractions d'oligoporphyrane, nous avons utilisé la technique de chromatopgraphie d'échangeuses d'ions (CaptoAdhere) mise au point pour caractériser les substrats de la sulfatase *Zg*4660 (**Chapitre IV**). Toutes les fractions ont été incubées avec la sulfatase *Zg*3515 et analysées par électrophorèse FACE (**Figure 64**. Malheureusement, aucune activité sulfatase n'a pu être détectée sur ces fractions. Au vu de ces résultats, et sachant que la caractérisation par spectrométrie de masse des oligoporphyranes générés par *Zg*PorA n'a pas révélé de sulfatations suppémentaires autre que le L-galactose 6-sulfate (Ropartz et al., 2015), il est possible que la sulfatase

*Zg*3515 agisse sur une sulfatation minoritaire dans le porphyrane. Nous avons donc recherché le substrat de cette sulfatase à partir d'une autre source d'algue.

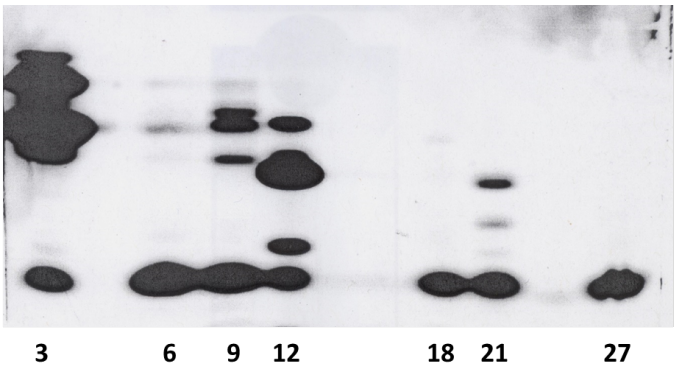


Figure 64 : Purification par CaptoAdhere des oligoporphyranes générés par *Zg*PorB.

Présentations des différentes fractions obtenues après purification par chromatographie échangeuse d'ions. Le gradient de carbonate d'ammonium est croissant de la fraction 3 à 27. Le marquage des oligosaccharides était réalisé avec de l'AMAC.

V. 2. B. e Détection et caractérisation de l'activité sulfatase de *Zg*3515 sur des oligosaccharides provenant de diverses algues agarophytes.

L'activité sulfatase a été criblée selon le même protocole que celui utilisé pour la sulfatase *Zg*4660 (**Chapitre IV**). Après action des β-agarases et β-porphyranases de *Zobellia galactanivorans* Dsij^T (*Zg*AgaB, *Zg*AgaC et *Zg*PorB), un grand nombre d'oligosaccharides ont été obtenus à partir des algues suivantes : *Osmundea pinnatifida*, *Deleisseria sanguinea*, *Chondria dasyphylla*, *Rhodomella sp.*, *Polysiphonia simulans*, *Polysiphonia brodiei* et *Dumontia contorta*. .

Nous avons pu ainsi identifier par électrophorèse FACE que les oligosaccharides provenant de *Chondria dasyphylla*, d'*Osmundea pinnatifida*, de *Deleisseria sanguinea* et de *Polysiphonia simulans* peuvent être des substrats potentiels pour la sulfatase *Zg*3515 (**Figure 65**).

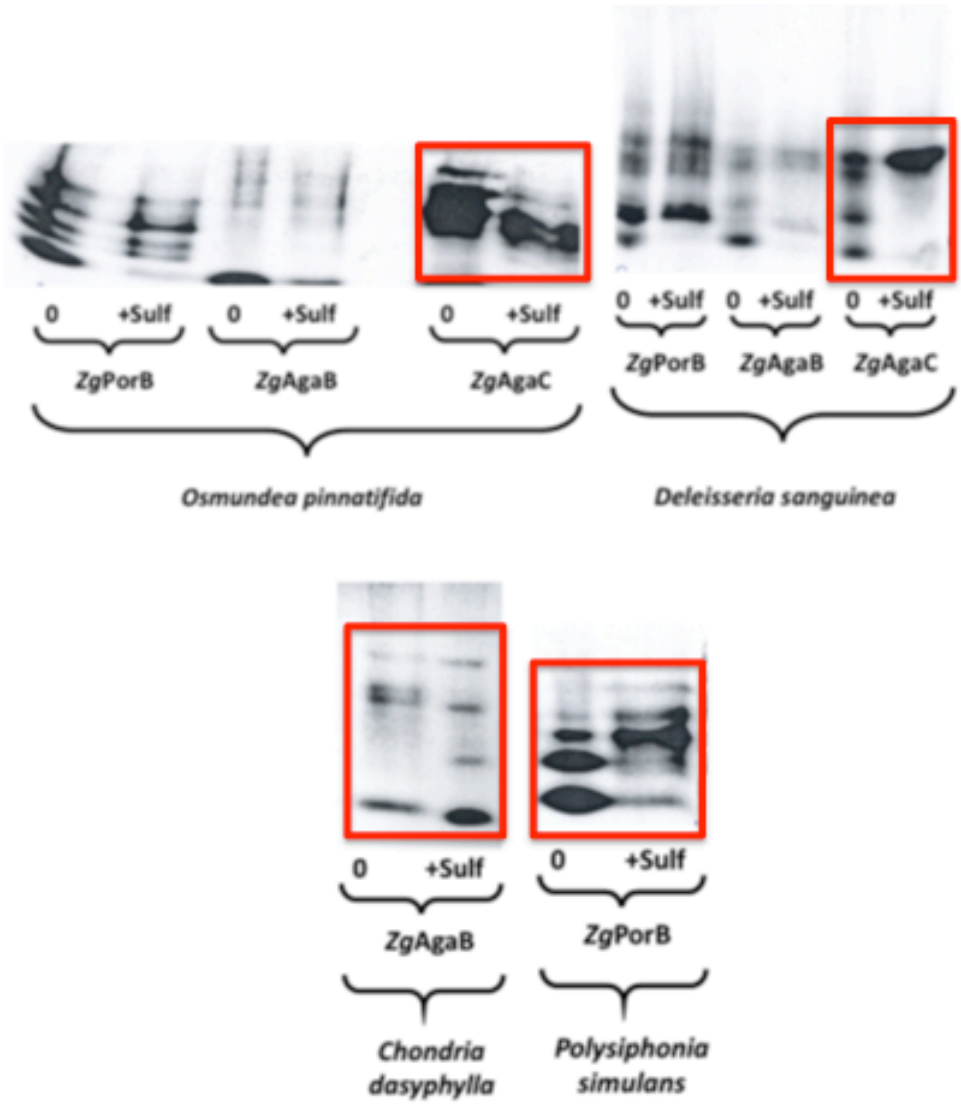


Figure 65 : Activité sulfatase de *Zg*3515 sur différents oligosaccharides issus d'algues agarophytes.

Sont présentés ici les profils de dégradation générés par différentes GH16 (*Zg*AgaB, *Zg*PorB et *Zg*AgaC) à partir de plusieurs algues (*Chondria dasyphylla*, *Osmundea pinnatifida*, *Deleisseria sanguinea* et *Polysiphonia simulans*). Les fractions substrats de *Zg*3515 sont encadrées en rouge. Le fluorophore utilisé pour le marquage est l'AMAC.

Pour caractériser l'activité sulfatase, nous avons choisi de purifier les oligosaccharides issus de *Chondria dasyphylla.* En effet, l'hydrolyse par *Zg*AgaB des polysaccharides pariétaux de cette algue, générant peu d'oligosaccharides, leur purification devait être potentiellement plus facile. Le rendement de l'extraction des polysaccharides de *Chondria dasyphylla* est de 43 %. Le protocole de purification par

CaptoAdhere a été adapté pour optimiser la séparation des oligosaccharides. Plusieurs purifications ont été réalisées pour accumuler suffisamment de substrat afin de réaliser des chromatographies analytiques par HPEAC. Les expériences ont été réalisées sur les fractions 2 et 4. La fraction 4 a révélé qu'elle contenait un oligosaccharide majoritaire qui, avant traitement avec la sulfatase *Zg*3515, est élué au bout de 11,5 minutes (**Figure 66**). Après incubation avec la sulfatase, la fraction 4 fait apparaître un pic d'oligosaccharide après 19 minutes d'élution. Un résultat identique était obtenu pour la fraction 2, mais avec une intensité des pics, et donc de quantité de d'oligosaccharide, plus faible. Finalement, les quantités d'oligosaccharide obtenues, trop faibles, ne nous ont pas permis de réaliser des analyses structurales par spectrométrie de masse.

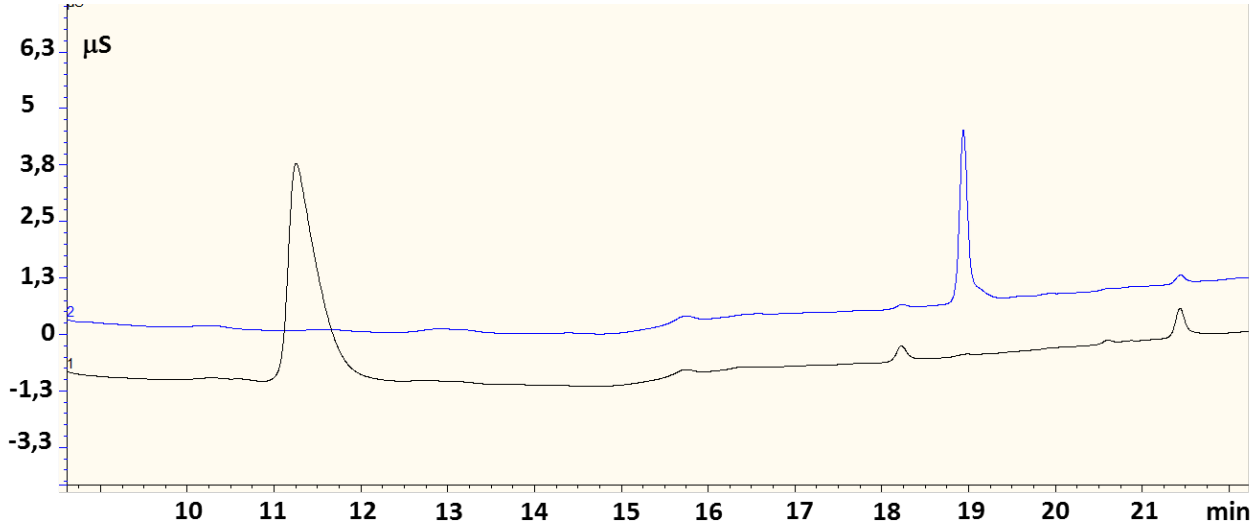


Figure 66 : Chromatographie HPEAC sur la fraction 4 d'oligosaccharides de *Chondria dasyphylla* avant et après traitement à la sulfatase *Zg*3515

La figure présente le profil chromatographique avant (noir) et après incubation avec la sulfatase (bleu). Le pic majoritaire avant incubation avec la sulfatase est élué à 11,5 minutes. Après incubation avec la sulfatase le pic majoritaire (en bleu) est élué à 19 minutes. Deux pics de faible intensité à 18,2 et 21,5 minutes semblent également diminués après action de l'enzyme.

Chondria dasyphylla étant très peu présente sur l'estran, nous avons décidé de tester les oligosaccharides produits par *Zg*AgaC à partir *d'Osmundea pinnatifida*. En effet, dans le cadre des études menées pour caractériser la spécificité de *Zg*AgaC, plusieurs purifications d'oligosaccharides d'*Osmundea pinnatifida* avaient été réalisées. Les chromatographies par HPAEC de ces fractions n'ont mis en évidence aucunes activités de *Zg*3515 sur ces oligosaccharides. Cela indique que la sulfatase n'agit probablement pas en présence d'autres modifications telles que certaines sulfatations ou méthylations. Cette hypothèse est cohérente avec les localisations cellulaires putatives

des enzymes au sein de *Z. galactanivorans*. En effet, *Zg*3515, qui se situerait au sein du périplasme, ne peut peut-être pas agir sur un oligosaccharide qui présenterait les sulfatations reconnues par *Zg*3514 et *Zg*3517 qui sont censées agir en premier car potentiellement localisées sur la membrane externe et exposées vers l'extérieur. De plus, la fraction oligosaccharidique caractérisée dans le chapitre consacré à *Zg*AgaC a été incubée en présence de la sulfatase *Zg*3515 sans succès ce qui indique qu'il est peu probable que les sulfatations du L-galactose 6-sulfate et le 3,6 anhydro-L-galactose 2-sulfate soient des substrats de *Zg*3515. Le polysaccharide d'*Osmundea pinnatifida* ayant été caractérisé (Bowker and Turvey, 1968a, b), nous pouvons proposer deux positions sulfatées qui peuvent être substrat de la sulfatase *Zg*3515 : le L-galactose 2-sulfate et le L-galactose 4-sulfate.

Actuellement, nos efforts se poursuivent sur les oligosaccharides provenant de *Deleisseria sanguinea* dont une importante récolte a été effectuée en Juin 2018. Une autre avancée a été réalisée en utilisant le reste des oligoporphyranes purifiés par chromatographie d'exclusion de taille (§ V. 2. A. c).

V. 2. B. f Cristallisation et résolution de la structure de *Zg*3515.

Lors des tests de cristallisation, les premiers cristaux sont apparus avec le kit de criblage commercial JCSG+ de Qiagen. La condition de cristallisation était acide malique D/L 2,1 M, pH7. J'ai reproduit ces cristaux en boite Linbro avec les conditions suivantes acide malique D/L (1,8/ 1,9 M) pH 7 avec un micro-ensemencement à la moustache de chat. Ces cristaux ont été congelés sous flux d'azote en présence de 15 % de glycérol comme cryoprotectant. Les données ont été collectées sur la ligne ID23-A du synchrotron ESRF. Un des cristaux a diffracté à la résolution de 1,3 Å.

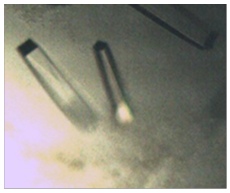


Figure 67 : Cristaux optimisé de la sulfatase *Zg*3515
Les cristaux ont été obtenus par la technique de la goutte suspendue dans les conditions suivantes : d'acide malique D/L (1,8/ 1,9 M) pH 7 avec micro-ensemencement à la moustache de chat.

	***Zg*3515**	**Se-*Zg*3515**
Collection des données		
Groupe d'espace	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Dimensions de l'UA		
a, b, c (Å)	48.85 146.35 146.38	48.76 146.00 146.16
α, β, γ (°)	90 90 90	90 90 90
Résolution (Å)	44.18 - 1.33 (1.378 - 1.33)	50.00 – 1.74 (1.79 – 1.74)
R_{merge}	0.07 (0.71)	0.16 (0.74)
I/σI	12.46 (2.67)	8.79 (2.20)
CC(1/2)	0.99 (0.78)	98.9 (80.5)
Complétude (%)	97.99 (98.16)	99.6 (97.6)
Redondance	5.71 (6.49)	3.39 (3.31)
Anom CC	ND	0.60 (à 2.0 Å)
Affinement		
Résolution (Å)	1.33	
Nombre de réflexions	235768 (23333)	
R/R_{free}	0.1613/0.1845	
Nombre d'atomes	9162	
Protéines	8021	
Ligand/ion	36	
Eau	1105	
Facteur B		
Protéines	13.83	
Ligand/ion	28.69	
Eau	28.75	
R. m. s. déviations		
Longueur de liaisons (Å)	0.028	
Angles (°)	2.38	

Tableau 8 : Statistiques des données de collecte et d'affinement pour le cristal ayant permis l'obtention de la structure de *Zg*3515

Les valeurs entre parenthèse concernent les données pour la plus haute résolution.

Ce jeu de données de diffraction a été indexé en $P42_12_1$ (92). J'ai d'abord tenté le phasage par remplacement moléculaire avec une structure homologue : une arylsulfatase putative d'*E. coli* CFT073 (PDB : 3ED4). La sulfatase *Zg*3515 (sous-famille S1_20) et cette protéine (sous-famille S1_23) partage 27% d'identité sur 90% de la séquence couverte. Malheureusement, cette approche n'a pas été concluante.

J'ai donc produit la protéine sous sa forme séléniée et j'ai obtenu des cristaux dans la même condition que celle employée pour la protéine native. J'ai collecté un jeu de données avec une résolution de 1,7 Å à la longueur d'onde de résonance du sélénium sur la ligne Proxima 1 du synchrotron SOLEIL. J'ai pu ensuite tenter de phaser selon la

méthode SAD (*Single Anomalous Diffraction*) avec SHELX C/D/E. Toutes les sélénométhionines (16 au totale) que comporte le dimère cristallographique ont été identifiées. Le meilleur taux d'occupation était de 1, la plupart des autres atomes ayant un taux d'occupation d'environ 0,7, excepté deux atomes qui avaient un taux d'occupation de 0,47 et 0, 49. Cela m'a permis de définir sans ambiguïté les positions des atomes dans la maille cristallographique. Lors de l'étape de phasage, il est apparu que le groupe d'espace dans lequel notre jeu était indexé ($P42_12_1$ (92)) n'était pas correct et a dû être remplacé par le groupe d'espace $P22_12_1$ (19), la haute résolution de notre jeu nous avait fait penser à une symétrie plus haute. Cette erreur d'indexation est surement la cause de l'échec du phasage par remplacement moléculaire.

Un modèle, comportant deux monomères dans l'unité asymétrique, a été construit avec ArpWarp. Ce modèle a été corrigé avec Coot et la formylglycine a été positionnée manuellement. L'affinement a été réalisé avec Refmac (R : 0,16 et R_{free} : 0,18). Comme la protéine est monomérique en solution, la présence d'un dimère est un artefact de cristallisation. Cette structure arbore les caractéristiques typiques des sulfatases de la famille S1 avec un domaine N-terminal très important constitué de 9 brins β organisés en un feuillet parallèle (sauf les brins β5 et 8 qui sont antiparallèles), entouré de dix hélices α bien définies. Le domaine C-terminal est constitué d'un feuillet de quatre brins β antiparallèles adossés au domaine N-terminal, ce feuillet est entouré par deux hélices α (**Figure 68**)

Figure 68 : Structure tertiaire de la sulfatase *Zg*3515
Le domaine N-terminal est en violet et le domaine C-terminal est en bleu.

L'analyse de la surface électrostatique de la sulfatase *Zg*3515 révèle une grande zone chargée positivement autour d'une gorge au fond de laquelle se situe le site actif. Les résidus impliqués sont essentiellement des arginines (R353, R381, R525 et R533) et des lysines (K114, K115, K216 et K530). Ces résidus sont probablement impliqués dans la reconnaissance du substrat soit par des interactions électrostatiques directes, soit *via* un réseau de molécules d'eau (**Figure 69**).

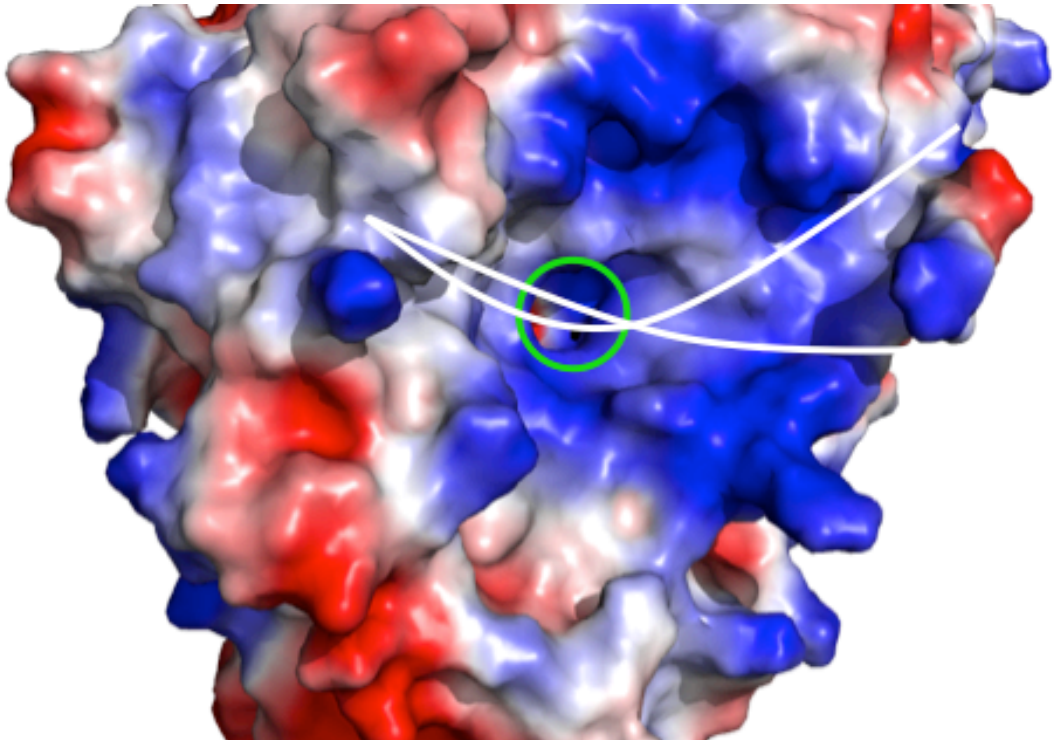


Figure 69 : Représentation de la gorge du site actif de la sulfatase *Zg*3515
Le site actif est présenté sous la forme de la surface électrostatique. En rouge sont représentées les zones électronégatives et en bleu les zones électropositives. La position du résidu formylglycine est signalée par le cercle vert, tandis que les positions où pourrait se fixer le substrat sont signalées par les traits blancs. De nombreux résidus basiques situés tout autour du site S sont représentés par une large zone électropositive.

A proximité du site actif se trouve également le site de fixation du cation divalent, qui est un élément conservé chez les sulfatases de la famille S1, chélaté par 3 résidus aspartate (D47, D48 et D328) et une asparagine (N329). Ce cation a été modélisé comme étant un calcium, car c'est l'ion le plus fréquent dans les structures de sulfatases. Cependant, cette hypothèse ne satisfait pas la densité électronique car, après affinement, un surplus d'électron est clairement signalé. La présence d'un ion magnésium ne satisfait pas non plus la densité car dans ce cas il manque des électrons.

Cela signifie que soit c'est bien un calcium et nous avons un taux d'occupation partiel du site par ce dernier, soit que nous sommes en présence d'un mélange calcium/ magnésium dans le site. Les dix résidus conservés au sein des sulfatases de la famille S1, et impliqués dans la reconnaissance du sulfate, sont retrouvés au sein de la structure (**Figure 69**).

La nouveauté se situe au niveau du résidu C86, contenu dans le motif consensus CAPAR, et qui est le résidu modifié en formylglycine. Dans la structure nous avons une densité non ambiguë au niveau de ce résidu qui indique clairement que nous sommes en présence de la FGly sous sa forme hydraté (gem-diol). Ce qui confirme la capacité d'*E. coli* à réaliser la modification co-traductionnelle de manière efficace pour *Zg*3515. En effet, si nous avions une grande proportion d'enzyme non modifiée, la densité serait moins nette. Mais nous ne pouvons pas conclure sur le fait que la forme gem-diol est présente en l'absence de substrat. En effet, le site actif présente une densité importante non expliquée en continuité avec l'Oγ_1 du résidu FGly comme si ce dernier était lié covalemment avec une espèce chimique. Il s'agit probablement d'un élément que la protéine a capté au sein d'*E. coli* ou lors de la purification (**Figure 71**). Ainsi la question de l'hydratation de la FGly pour former le gem-diol reste ouverte : est-ce la forme préférentielle ou la FGly était-elle hydratée consécutivement à l'entrée dans le site actif du substrat ?

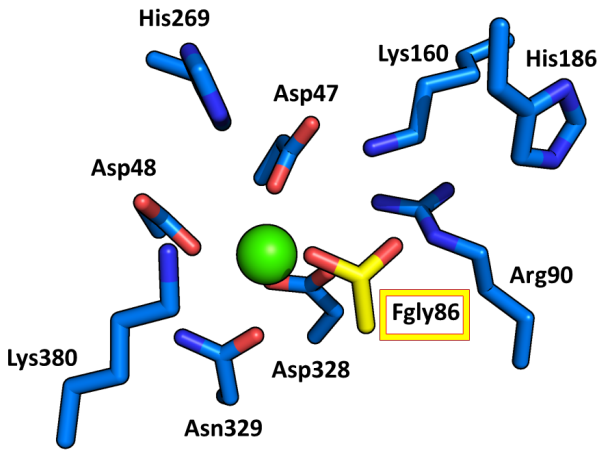


Figure 70 : Résidus impliqué dans la reconnaissance du sulfate

Dans le site actif de *Zg*3515 sont présents les résidus nécessaires à la reconnaissance du sulfate et à son hydrolyse. La sphère verte symbolise le calcium qui est chélaté par les trois acides aspartiques (47, 48 et 328) et par une asparagine (329).

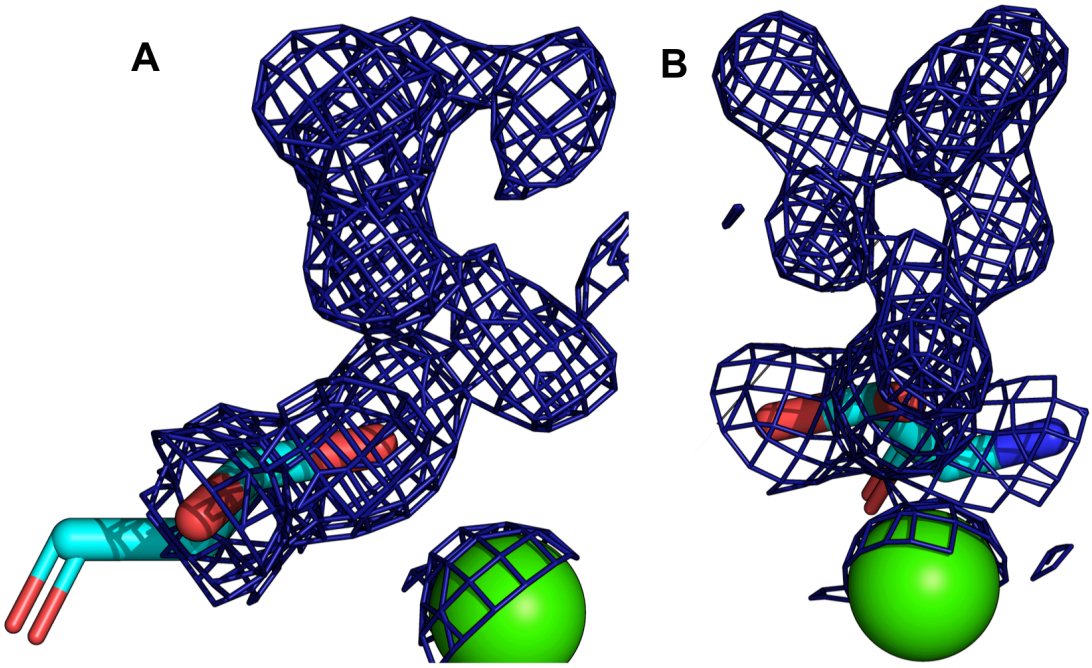


Figure 71 : Densité non modélisée au niveau de la formylglycine du site actif de la sulfatase *Zg*3515

La formylglycine est représentée ici en bâton de couleur cyan, et le calcium est représenté sous la forme d'une sphère verte. La densité en bleu marine, représentée ici, correspond à la carte 2Fo-Fc. La figure A montre la densité vue de profil qui présente deux protubérances marquées. La figure B montre la densité vue de face. Il est clairement visible qu'un cycle de type cinq carbone est présent.

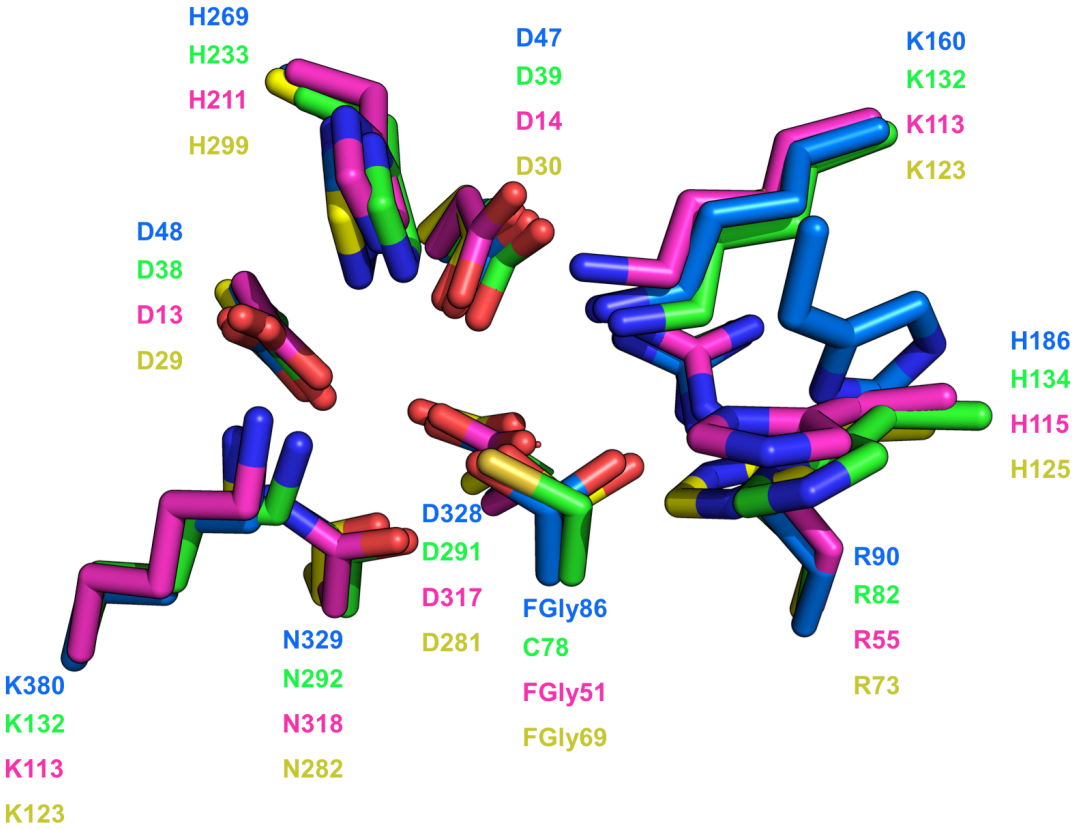


Figure 72 : Conservation structurale des résidus impliqués dans la reconnaissance des groupements sulfate

Comparaison de quatre sites actifs de sulfatase : PsS1_19A en vert (S1_19 ; PDB : 6BIA), AtsA en rose (S1_4 ; PDB : 1HDH), ARSA en jaune (S1_1 ; PDB : 1AUK) et *Zg*3515 en bleu. Si la position de neuf des résidus est conservée chez ces sulfatases, nous noterons que pour *Zg*3515 l'histidine 186 (milieu à droite) n'est pas à la même position que les histidines dans les trois autres structures.

Le site actif où se déroule le clivage de la liaison thioester ou sous-site S, selon la nomenclature de Hettles et collaborateurs (2018), se situe au niveau d'une vaste gorge (**Figure 69**). La difficulté pour l'interprétation, et une éventuelle modélisation, est que cette dernière est bifide. Ainsi, nous ne pouvons pas émettre d'hypothèse sur le positionnement de l'oligosaccharide au sein de la gorge. Néanmoins, la conformation de cette gorge rappelle la structure de la PsSI_19A qui est une *endo*-4S-iota-carrageenan sulfatase (Hettle et al., 2018).

Au sein des structures de sulfatases, il a été retrouvé une histidine (His115, H125, H134 dans la **Figure 72** ; HisA dans la **Figure 73**) dont le Nδ1 serait impliqué dans un transfert d'hydrogène qui permet ainsi l'élimination de l'adduit sulfate (HisA dans la **Figure73**) par une réaction d'élimination E2. Mais dans *Zg*3515, cette histidine n'est pas conservée au sein de la séquence. Cependant, l'histidine 186, positionnée différemment

par rapport aux autres sulfatases, est présente. Dans *Zg*3515, le Nδ1 de l'histidine H186 est à une distance de 4,6 Å de l'Oγ1 de la FGly. Or, l'étude de Hatcher et co-auteurs (Hatcher et al., 2004), montre que les transferts d'hydrogène ce déroulent préférentiellement quand la distance entre l'accepteur et le donneur est de 2,69 Å. Cette distance est en effet retrouvée dans les autres structures de sulfatase comme celle de l'aryl-sulfatase B humaine ARSB (S1_2 ; PDB : 1FSU) où la distance est de 2,7 Å. Cette différence de distance provient vraisemblablement d'une insertion au niveau de la séquence, qui est conservée chez toutes les sulfatases de la famille S1_20. Cette différence mène à la question cruciale, à savoir, comment la catalyse s'effectue-t-elle ?

Plusieurs hypothèses s'offrent à nous. Aurions-nous un mouvement dans le site actif permettant le positionnement correct de H186, ou bien au contraire de ce qui est communément admis, serions-nous face à une réaction de type SN_2 ?

Dans le cas d'un mécanisme de type SN_2, il est probable que l'histidine B (**Figure 73**) agit en tant que base sur une molécule d'eau pour permettre une attaque de l'adduit sulfate-protéine. Or, dans la structure de *Zg*3515, une molécule d'eau est très bien définie entre l'histidine 186 et l'Oγ1 de la FGly. D'autres possibilités existent également tel qu'un relai de proton impliquant H186 et une molécule d'eau.

Mais cette molécule d'eau pourrait également servir de relai de proton pour permettre une élimination de type E2.

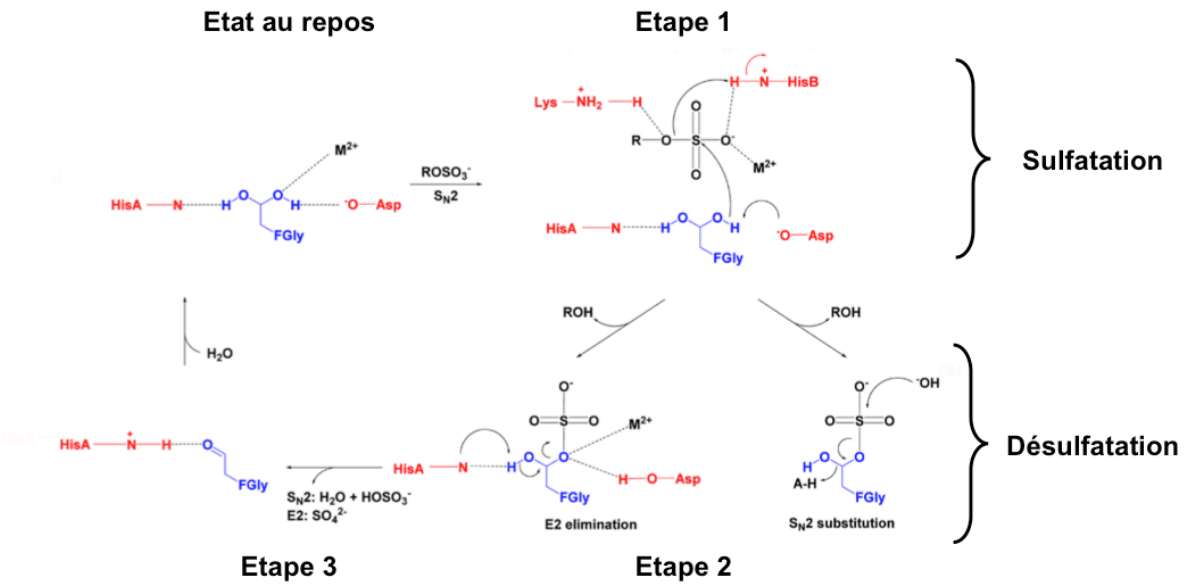


Figure 73 : Schéma réactionnelle des sulfatase de la famille 1 ou formylglycine dépendante. (Tiré de Appel et al., 2014)

L'état au repos (en l'absence de substrat) est symbolisé par une formylglycine sous forme gem-diol. L'étape 1 schématise l'attaque nucléophile réalisée par la FGly après son activation par un aspartate. L'étape 2, est la plus débattue même si le mécanisme d'élimination E2 est privilégié. L'étape 2 impliquerait une base catalitique qui serait l'histidine A, tandis que dans le cas du mécanisme de type S_N2 une molécule d'eau activée réaliserait l'attaque du groupement sulfate.

V. 2. B. g Etude cinétique préliminaires de différents mutants de Zg3515

Plusieurs mutants de la sulfatase *Zg*3515 ont été réalisés. Ces deniers concernent les résidus impliqués dans la catalyse et visait à expliciter le mécanisme des sulfatases. L'autre but de ces expériences était d'explorer la possibilité de transformer une sulfatases en trans-sulfatase. Pour rappel, dix résidus sont impliqués dans la catalyse (Hanson et al., 2004). Trois acides aspartiques et une glutamine permettent de fixer un cation divalent (généralement du calcium ou du magnésium), une histidine (HisA) permet de stabilisé la forme gem-diol de la formylglycine (FGly), qui est le résidu nucléophile, qu'un aspartate va activer par une catalyse base. L'attaque nucléophile conduit à la formation d'un adduit sulfate–enzyme qui sera par la suite éliminé. Cette élimination est réalisée selon deux mécanismes qui sont en discussion. Soit une élimination de type E2, impliquant une catalyse base par l'hisitidine A, soit une substitution nucléophile faisant intervenir une molécule d'eau activée (Appel et al., 2014). Dans ce dernier cas, concernant le résidu impliqué dans l'activation aucun candidat sérieux n'a été identifié.

Ainsi, le résidu dont le rôle est le plus à vérifier est l'histidine A (H186 chez *Zg*3515) qui joue un rôle cruciale dans le cas où nous aurions un mécanisme de type E2. De plus dans l'optique de faire de la sulfatase *Zg*3515 une *trans*-sulfatase (mais aussi de réaliser des co-cristallisations), il était nécessaire de muter la cystéine C86 qui est modifiée pour devenir par la suite la FGly. Celle-ci a été mutée notamment en sérine. Mais il est nécessaire de préciser que la sérine peut être modifiée par la FGE (Bojarova et al., 2008) et que chimiquement, le soufre de la fonction thiol peut permettre une attaque nucléophile.

Ainsi, le mutant C86S s'est révélé être actif avec la MUS utilisée comme substrat. Comme pour la sulfatase native, le pH optimum de la sulfatase mutée sur ce substrat a été déterminé à 7,33 comme pour la sulfatase native. La différence entre les deux est une diminution de l'activité de 4,57 fois par rapport à l'enzyme native dans les mêmes conditions (**Figure 60**).

Inactivation par oses simples et disaccharides

Pour analyser les propriétés de reconnaissance du site actif et réaliser les premières expériences de *trans*-sulfatation, des tests d'inhibition de l'enzyme en présence d'oses simples et de disaccharides, à concentration fixe (100 mM), ont été réalisés avec l'enzyme sauvage et le mutant C86S. Globalement les résultats d'inhibition sont très proches entre le sauvage et le mutant C86S pour le fructose, le glucose et le fucose (de la série D). Les différences notables se situent aux niveaux du D-galactose, du lactose, du maltose, du sucrose et du mélibiose qui sont activateur pour l'enzyme sauvage et inhibiteur pour le mutant C86S (**Figure 74**).

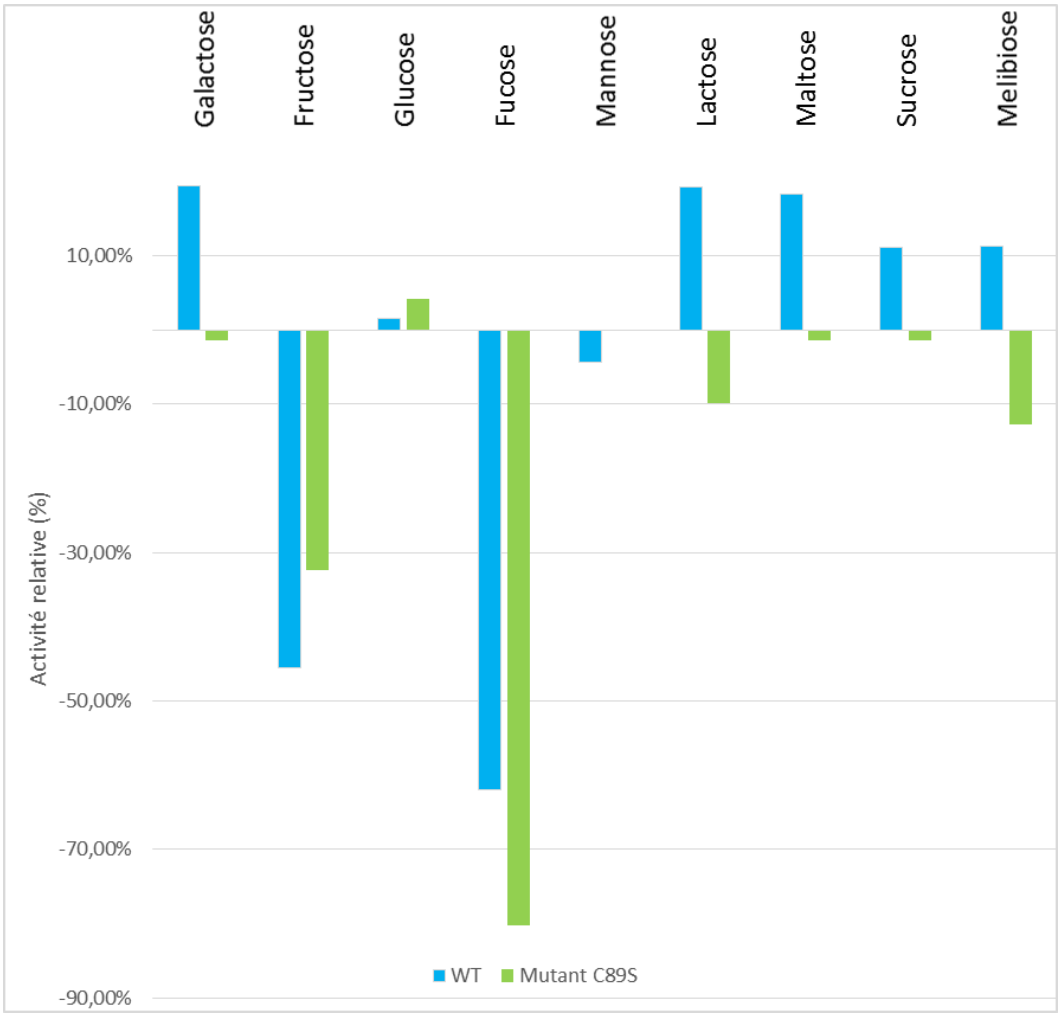


Figure 74 : Activité relative de *Zg*3515 avec le substrat MUS en présence de différents oses.

Les oses testés sont à une concentration de 100mM. Tous les oses simples sont de la série D. Le pourcentage d'activité est calculé à partir de l'enzyme seule en présence de MUS. De façon claire nous observons une augmentation de l'activité de l'enzyme sauvage (en bleu) en présence de D-galactose, lactose, maltose, sucrose et mélibiose. Pour le mutant C86S (en vert) et l'enzyme sauvage nous avons une inhibition marquée en présence de fructose et de fucose.

Impact de la stéréochimie sur la catalyse

Une étude a été réalisée sur la stéréosélectivité de *Zg*3515 en testant différent oses de série D ou L : L-galactose, D-galactose, D-fructose et L-fucose, la concentration de ceux-ci était de 50 mM. Les mutants testés ici étaient C86S et H186N. L'enzyme sauvage est fortement inhibé (60 et 80 % d'inhibition) en présence de L-galactose mais pas en présence de D-galactose. Mais à 50 mM contrairement à 100mM de D-galactose nous n'avons plus l'effet activateur de ce dernier. Curieusement, le D-fructose inhibe également l'enzyme de manière efficace. Les autres monosaccharides ou disaccharides ont peu ou pas d'effet sur l'activité.

Le fait que des oses inhibent l'activité tend à indiquer que le site actif le reconnaît et bloque l'entrée de la MUS. Ce résultat suggère que le site -1 ou +1 est susceptible de reconnaître un L-galactose, ce qui cohérent avec la présence de cet ose dans la structure des agars. Le fait que le D-Fructose soit aussi inhibiteur est sans doute lié à sa structure de type pyranose qui est majoritaire dans l'eau (Schneider et al., 1985). Cependant lorsque celui-ci est lié à un autre ose, comme dans le saccharose, le pouvoir inhibiteur est perdu, probablement en raison du caractère discriminant du glucose associé.

Le L-fucose est lui aussi largement inhibiteur, plus que le L-galactose dans le cas de l'enzyme native. Cela peut s'expliquer par la proximité structurale avec le L-galactose (la seule différence entre les deux est que le L-galactose possèdent un hydroxyle en position 6 alors que le L-fucose n'as qu'un méthyl en position 5). Cette explication est également valable pour le D-fructose où les positions des hydroxyles sur les C_2, C_3 et C_4 sont identiques à ce que l'on retrouve chez le L-galactose. Nous pouvons en déduire que la reconnaissance du L-galactose est très peu dépendante de l'hydroxyle 6, et que l'enzyme a une nette préférence pour la série L ce qui est cohérent avec le fait que *Zg*3515 agit sur des agars (**Figure 75**).

Néanmoins ces données sont parcellaires et pour conclure sur les motifs reconnus ou non par l'enzyme, il est nécessaire d'entreprendre une étude cinétique systématique. Un autre point à vérifier serait de déterminer l'impact, sur l'étape de désulfatation, de la mutation de l'histidine H186.

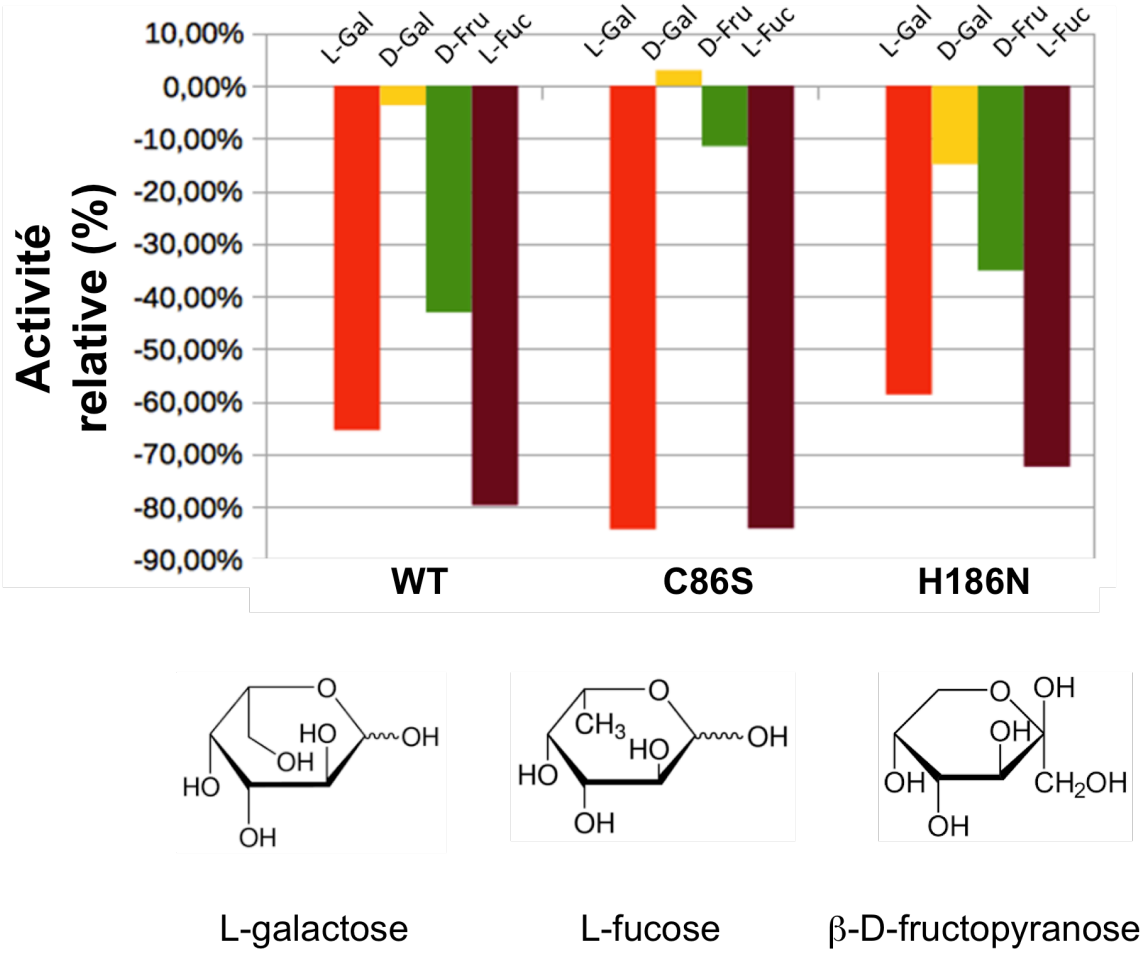


Figure 75 : Effet de la stéréospécificité sur l'acitivité sulfatase de *Zg*3515.
L'histogramme montre l'activité relative de *Zg*3515 et de ces mutants (C86S et H186N) en présence de 100 mM d'oses simples : L-galactose (L-Gal), D-galactose (D-Gal), D-fructose (D-Fru) et L-fucose (L-Fuc). En dessous se trouvent les représentations d'Haworth du L-galactose, L-fucose et du β-D-fructopyranose.

V. 2. B. h Premier test de *trans*-sulfatation avec *Zg*3515

Pour ces tests il était nécessaire que le résidu nucléophile soit éliminé donc nous avons utilisé le mutant C86G. De plus, pour éviter une élimination du sulfate du donneur activé l'histidine censée réalisée l'étape d'élimination (H186) a également été mutée en glutamine. Le test consistait à voir si le mutant C86G/H186Q pouvait transférer le sulfate d'un substrat activé (l'acide sulfamique) sur le L-galactose. Malgré les tests avec plusieurs concentrations d'acides sulfamiques et le test de deux conditions de tampon, aucune sulfatation du L-galactose n'a pu être obtenue.

Plusieurs raisons ont été avancées pour expliquer ces résultats négatifs. Tous d'abord que l'acide sulfamique n'est pas été reconnu par l'enzyme ou qu'il ne soit pas assez activé. En effet, quand ce dernier est utilisé pour sulfater des polysaccharides chimiquement, le taux de sulfatation est à son plus haut quand la réaction se déroule à 90°C, alors que nous travaillons à 20°C (Vasil'eva et al., 2013). De plus nous n'avons tenté la réaction que sur un L-galactose et ce dernier n'est peut-être pas suffisamment stabilisé au sein de l'enzyme. Il est également possible que la transulfatation se fasse de façon trop infime pour être détectée.

V. 3 Conclusions et perspectives

Pour l'étude du couple *Zg*3515/*Zg*3516, nous avons dû rechercher les substrats naturels. Des premiers éléments étaient disponibles pour nous permettre d'orienter cette recherche : le gène *zg*3516 était annoté comme étant une xylanase de la famille GH30 (sous-famille 4). De plus, nous disposions de données expérimentales. En effet, le Dr. F. Thomas a réalisé des expériences de transcriptomiques montrant une activation de la transcription de *Zg*3515 en présence de porphyrane et de κ-carraghénane et de *Zg*3516 en présence de κ-carraghénane (Thomas et al., 2017). Ainsi, les recherches étaient orientées vers trois polysaccharides : le porphyrane, le κ-carraghénanes et le xylane sulfaté.

Néanmoins, ces premières indications n'excluaient en rien d'autres polysaccharides. Ces derniers pouvant être très complexes, en effet plusieurs gènes codant pour des enzymes éliminant des modifications sont présents au sein de l'opéron 37 ce qui rend difficile la découverte du substrat. Face aux nombreuses possibilités de substrat potentielles, plusieurs expériences étaient possibles. Une première analyse, la plus rapide, consistait à tester l'action de la glycoside hydrolase sur différents substrats artificiels. La difficulté de cette approche est qu'il s'agit de petit substrat, sans modification, qui peuvent ne pas être clivé par une glycoside hydrolase qui clive seulement si l'oligosaccharide a une certaine taille, ce qui est caractéristique de certaines enzymes endolytiques.

La seconde approche est de réalisé des électrophorèse de type FACE ou C-PAGE qui permettent de visualiser les oligosaccharides générés à partir d'un polysaccharide. Les problèmes encourus lors de ces expériences sont de ne pas avoir à disposition le

polysaccharide substrat ou bien que l'enzyme ne puisse agir que sur des petits oligosaccharides non visualisables sur gel d'éléctrophorèse.

La dernière approche, la chromatographie sur couche mince, permet de visualiser les petits oligosaccharides produits par l'action de l'enzyme. Cette dernière approche nous a permis d'avoir la première piste sur le substrat de *Zg*3516. L'utilisation de la seconde méthode sur des polysaccharides déjà digérés par des glycosides hydrolases nous a permis de confirmer l'activité de *Zg*3516 sur des oligosaccharides générés par *Zg*PorB à partir de porphyrane. Mais les tests de dégradation enzymatique par *Zg*3516, sur des polysaccharides entiers, dont le porphyrane, ce sont révélés négatifs.

Des expériences avec des oligosaccharides purifiés, ont démontrées que *Zg*3516 peut dégrader des oligoagaroses, s'ils sont d'une taille supérieure à un DP4, mais pas des oligoporphyranes. La structure de *Zg*3516, présente un site actif de faible taille et très peu ouvert, *a contrario* de ce que l'on retrouve dans la majorité des glycosides hydrolases endolytique *Zg*AgaB, *Zg*PorB ou *Zg*AgaC.

La structure de *Zg*3516 a également révélé une organisation de ses structures secondaires non rencontré à ce jour chez les GH30. Ce qui, avec le fait qu'il n'y avait jusqu'à présent pas d'agarases dans la famille GH30, nous amène à dire que *Zg*3516 fait partie d'une nouvelle sous-famille structurale de GH30 encore non décrite.

Bien que tous les éléments concourent à démontrer que *Zg*3516 est une enzyme avec un mode d'action exolytique, il reste à réaliser un suivi cinétique pour démontrer que ce sont bien des oligosaccharides de type DP2 qui sont libérés par l'action de l'enzyme. Ce suivi permettra, par la même occasion, de définir le substrat préférentiel de l'enzyme.

Les résultats obtenus pour *Zg*3516, nous ont permis de réduire le champ des possibilités pour déterminer le substrat de *Zg*3515. Cette sulfatase fait partie de la sous-famille S1_20 pour laquelle aucune donnée n'a été publiée à ce jour. En effet, une activité sulfatase a pu être révélée sur des oligoporphyranes par électrophorèse. les mêmes oligoporphyranes pour lesquels nous avons obtenu une activité pour *Zg*3516. Ce premier test nous a permis d'établir que *Zg*3515 n'est pas active sur la sulfatation majoritaire du porphyrane, en effet dans le cas contraire de très nombreux oligoporphyranes auraient été impactés, ce qui n'est pas le cas ici. Mais les difficultés pour accumuler les oligosaccharides et caractériser ces derniers, nous ont conduits à rechercher de nouvelles pistes.

Pour ce faire, nous avons utilisé une technique de criblage à moyen débit. Ce criblage repose sur la comparaison simultanée par électrophorèse d'oligosaccharides, générés par l'action de plusieurs agarases et porphyranases sur différents polysaccharides, en présence et absence de sulfatase. Plusieurs polysaccharides se sont révélés intéressants. Nous avons néanmoins éliminé ceux issus d'*Osmundea pinnatifida* (trop difficile à utiliser à cause des nombreuses modifications) et de *Chondria dasyphylla* (trop peu de biomasse disponible pour une caractérisation).

En outre, une structure à haute résolution et de nombreuses données cinétiques sur différents mutants, ont pu être obtenues pour *Zg*3515. Ce sont les premières données expérimentales sur la sous-famille S1_20. De plus, l'étude de la structure a révélé une différence notable au niveau des résidus fortement conservés chez les sulfatases formylglycine dépendante. Parmi, ces derniers une des histidines est supposée permettre l'élimination de l'adduit sulfate sur le résidu catalytique : l'histidine 186. Mais pour cela, l'histidine devrait agir en tant que catalyseur base ce qui n'est possible que si elle se situe à une certaine distance de la FGly. Or, cette histidine (H186) n'est pas positionnée de façon à pouvoir tenir ce rôle. Ainsi, mon travail suggère de réaliser une étude plus poussée du mécanisme des sulfatases de la famille S1_20, car il est possible que cette sous-famille ait un mécanisme catalytique différent.

L'autre volet important de cette étude est de poursuivre la recherche du substrat de *Zg*3515 notamment en explorant la piste de *Deleisseria sanguinea*.

Mon travail apporte des éclairages sur la nature du substrat catabolisé par le PUL 37 auquel appartiennent *Zg*3515 et *Zg*3516. Ces deux enzymes ayant une activité sur les agars il est donc très probable que les autres enzymes agissent également sur le même polymère. Cela a d'autant plus de sens que la transcription des enzymes de ce PUL sont induites en présence de galactanes sulfatés tels que le porphyrane (l'agar extrait de *Porphyra* sp.) ou des carraghénanes (Thomas et al., 2017). Ainsi, le travail sur le couple *Zg*3515/*Zg*3516 ouvre la possibilité de caractériser de nouvelles enzymes agissant sur l'agar, notamment concernant les sulfatases : *Zg*3513 et Z*g*3517. En effet, *Zg*3513 est induite en présence de porphyrane mais pas *Zg*3517. Il est donc possible que *Zg*3513 soit impliquée dans la désulfatation du L-galactose 6-sulfate présent dans le porphyrane, même si nous ne pouvons exclure que ce soit en réalité *Zg*3517 ou bien aucune des deux enzymes. Une autre conséquence de ce travail est qu'il y a une forte probabilité que

*Zg*3514 soit une pyruvyl-estérase/galacturonyl-estérase agissant sur l'agar. Si cela était confirmé expérimentalement il s'agirait de la première activité de ce type décrite sur l'agar.

La découverte de l'activité de *Zg*3516 fait écho à une question restée en suspens lors de l'introduction générale : *quid* de l'équivalent d'Aga50D chez *Z. galactinovorans* ? En effet, dans l'introduction nous faisions état de l'existence d'une GH de la famille 50 qui permet de générer des agarobioses à partir d'oligosaccharides de grande taille (Pluvinage et al., 2013). Cette enzyme permet donc de passer des oligosaccharides de grande taille, générés par des enzymes comme *Zg*AgaB, vers des molécules de degré de polymérisation DP2/DP4, pouvant ensuite être dégradé en galactose par des enzymes comme *Zg*4663. Mais il n'y a pas d'enzyme de la famille GH50 chez *Z. galactinovorans*. Par conséquent est-ce que *Z. galactanivorans* a une enzyme alternative pour passer des grands oligoagars à des petits dégradés par *Zg*AhgA (famille GH117, Rebuffet et al., 2011) ou est-ce *Zg*AhgA suffit à la dégradation de ces longs oligosaccharides ?

La réponse à cette question semble être l'activité de *Zg*3516 (GH30). Cette dernière permet en effet de générer à partir de grands oligosaccharides, comme des DP10, des oligosaccharides plus petits, pouvant être par la suite hydrolysés en unité galactose par *Zg*AhgA et *Zg*4655 (famille GH117)(Ficko-Blean et al., 2015). Cette solution suppose que *Z. galactanivorans* puissent importer dans le périplasme, où se situe *Zg*3516, des oligosaccharides de type DP6 *a minima*, puisqu'elle semble incapable de cliver du DP4. Le modèle de dégradation du xyloglucane chez *Cellvibiro japonicus* implique qu'un DP4 ramifié avec un DP3 puisse être importé au sein du périplasme (Gardner, 2016). Ainsi, l'import d'un DP6 au sein du périplasme ne semble pas improbable.

L'autre avancée apportée par ce travail concerne la famille des GH30. Ici, nous avons deux faits concomitants : une nouvelle organisation du β-sandwich et nouvelle spécificité de substrat. Est-il possible qu'il y est un rapport de cause à effet ? Une hypothèse a été émise par St John et coauteur en 2010 dans ce sens. Ils proposaient que la conservation du β-sandwich est due au fait que ce dernier permettrait d'établir une plateforme proche du site actif et facilement modifiable, et permettant ainsi l'acquisition rapide de nouvelles spécificités de substrat.

VI Conclusions et perspectives générales.

VI. 1 Principaux résultats

Z. galactanivorans est une flavobactérie qui a été isolée sur *Delesseria sanguinea* (Potin et al., 1991), qui appartient à la famille des Delesseriaceae. Les organismes de cette famille sont connus pour contenir des agars très modifiés (G6S, G4S, LA2M (Craigie, 1990). Une question d'importance était donc de savoir comment *Z. galactanivorans* peut digérer un agar très modifié comme celui de *Delesseria sanguinea*.

Au laboratoire, *Zobellia galactanivorans* est capable de croitre avec comme seule source de carbone de l'agarose (Barbeyron et al., 2016). Pour ce faire, elle possède un large panel d'enzymes et de protéines dédiées à la reconnaissance, à la dégradation et à l'assimilation des polysaccharides. Concernant la dégradation des agars, elle dispose notamment de neuf gènes de la famille des GH16s codant pour des β-agarases ou des β-porphyranases. Cinq enzymes de la famille GH16 ont été caractérisées précédemment (*Zg*AgaA, *Zg*AgaB, *Zg*AgaD, *Zg*PorA et *Zg*PorB) et il s'est avéré qu'aucune de ces enzymes n'est capable de dégrader un agar très modifié (Hehemann et al., 2010 ; Hehemann et al., 2012). En effet, les modifications chimiques reconnues et acceptées par ces enzymes sont essentiellement des L6S et des G6M, par ailleurs fréquemment rencontrées sur l'agar industriel. Dans le cas des β-porphyranases le L6S est le motif reconnu et le substrat de l'enzyme. Les agars les plus modifiés présentent des branchements xylose ou galactose, des acides galacturoniques, des sulfatations et des méthylations à d'autres positions que celles précédemment mentionnées.

Par le passé il a été mis en évidence l'existence dans le génome de *Z. galactanivorans* d'une séquence divergente des autres β-agarases : *Zg*AgaC (Hehemann et al., 2010). Une précédente étude avait montré que *Zg*AgaC a un comportement peu classique vis-à-vis de l'agarose : elle peut dégrader de l'agarose mais plus lentement que les autres β-agarases, et par ailleurs elle n'est pas retenue sur une matrice de Sepharose (GE) (billes d'agarose) *a contrario* des autres β-agarases (Jam et al., 2005). Ces éléments nous ont fait supposé que *Zg*AgaC pouvait dégrader un agar présentant d'autres modifications que celles susmentionnées.

La caractérisation de *Zg*AgaC a confirmé notre hypothèse. En effet, pour la première fois, j'ai pu montrer qu'une GH16 pouvait cliver des oligosaccharides avec des motifs LA2M, LA2S et des D-galactoses présentant des branchements xylose. L'efficacité catalytique (k_{cat}/K_M) de *Zg*AgaC sur l'agarose est de 5839 $s^{-1}/\%^{-1}$ tandis que sur l'agar

extrait de *O. pinnatifida* elle est de 420 $s^{-1}/\%^{-1}$. *Zg*AgaC, malgré sa capacité à accepter des décorations, a donc une meilleure efficacité catalytique pour l'agarose, que pour l'agar modifié d'*Osmundea pinnatifida*. Cette différence pourrait s'expliquer en partie par la présence de nombreuses modifications autres que celles-prises en charge par l'enzyme, ou par un des sous-sites: G6M, G2S, L2M6S… (Bowker et Turvey 1968 a, b). Elle peut également s'expliquer par le fait que les paramètres cinétiques ont été déterminés dans les conditions optimales de l'enzyme pour le polysaccharide testé (pH 6,6 pour l'agarose et pH 9,1 pour l'agar d'*O. pinnatifida*), conditions très éloignées des conditions physiologiques. En effet, *Zg*AgaC est une enzyme extracellulaire et l'eau de mer à un pH de 8,3 environ, ce qui est plus proche de l'optimum de pH de *Zg*AgaC pour l'agar sulfaté que pour l'agarose.

Dans le contexte biologique, nous pouvons également imaginer que d'autres enzymes, telles que des sulfatases, pourraient agir avant *Zg*AgaC et éliminer les modifications non reconnues par l'enzyme avant son action, augmentant ainsi son efficacité. La caractérisation cinétique de *Zg*AgaC nous permet néanmoins de prouver pour la première fois l'impact important des microenvironnements sur l'activité des glycoside hydrolases.

Grâce à un concours de circonstances, nous avons découvert une nouvelle voie du catabolisme des agars chez *Z. galactanivorans*. En effet, la caractérisation de Zg3516, enzyme périplasmique et première GH30 active sur de l'agarose, permet de proposer l'hypothèse suivante : 1) *Z. galactanivorans* peut importer de longs oligosaccharides d'agarose au sein de son périplasme, 2) dans ce compartiment les sulfatations et peut-être les branchements finissent d'être éliminés et 3) Zg3516 transforme les oligosaccharides libérés en di et tétraagarobioses qui sont ensuite réduits en oses simples par l'action de *Zg*AghA et de *Zg*4655 (Rebuffet et al., 2011). Si cette capacité à importer des longs oligosaccharides était vérifiée, cela signifierait que *Z. galactanivorans* agirait selon une stratégie « égoïste », stratégie décrite pour d'autres bactéries tel que *Bifidobacterium thetaiotaomicron* sur le mannane (Cuskin et al., 2015). En effet, elle limiterait au maximum la dégradation des agars dans le milieu extracellulaire et les importeraient au sein de son périplasme pour éviter que les bactéries du milieu bénéficient de cette source de carbone.

Concernant les sulfatases actives sur les agars, des progrès significatifs ont été réalisés. Pour être actives, ces enzymes nécessitent d'être modifiées de façon co-traductionelle et cette modification est extrêmement difficile à générer chez *E. coli*. S'agissant de l'activation des sulfatases au sein d'*E. coli*, plusieurs protocoles ont été mis au point au cours de ma thèse (co-expression avec chaperonnes, traduction différentielle périplasmique). De plus, un protocole moyen débit a été établi pour permettre l'identification de leurs substrats naturels. Actuellement, nous disposons de plusieurs pistes qui pourraient nous permettre d'élucider la spécificité de substrat des sulfatases *Zg*3515 et *Zg*4660. La sulfatase *Zg*3515 semble active sur plusieurs fractions oligosaccharidiques générées à partir de différentes algues (*O. pinnatifida* et *D. sanguinea*) par *Zg*AgaC. La résolution de la structure et l'analyse du site actif de *Zg*3515 a soulevé de nombreuses questions sur le mécanisme catalytique des sulfatases de la famille S1-20. En effet, l'une des histidines conservées au sein des structures connues de sulfatases n'est pas à la même position chez *Zg*3515. Ainsi, la distance entre cette histidine et la formylglycine est trop importante pour permettre une réaction de β-elimination de type E_2. La question est donc de savoir comment la désulfatation de la FGly a lieu ? Le mécanisme de cette sulfatase repose - t'il sur un mécanisme de type SN_2 ? Ou serions- nous face à un relais de proton médié par une molécule d'eau ?

L'étude des sulfatases nous a amené à mettre au point un grand nombre de protocoles tels que l'extraction des polysaccharides sous pression ou la purification des oligosaccharides sulfatés *via* une matrice anionique. Toutes ces techniques pourront être utiles pour d'autres sujets portant sur des polysaccharides anioniques solubles natifs de nature et d'origine diverses. En effet, la méthode d'extraction est peu destructrice pour les modifications, elle permet d'avoir un polysaccharide très pur et avec un bon rendement et ceci rapidement. La technique de purification permet quant à elle de purifier des mélanges très complexes de façon efficace et avec un éluant dont les ions sont facilement éliminables par lyophilisation, évitant ainsi tous les problèmes liés à une trop haute concentration ionique.

VI. 2 Perspectives générales.

Pour finir la caractérisation de la sulfatase *Zg*4660, il est nécessaire d'obtenir l'enzyme sous forme active et donc de gérer la modification co-traductionnelle. Pour ce faire, plusieurs protocoles sont envisagés. Dans un premier temps la meilleure piste pour aboutir rapidement à un résultat consistera à produire l'enzyme dans son organisme d'origine. En effet, *Z. galactinovorans* comporte tous les mécanismes optimisés pour produire ses sulfatases sous leurs formes actives. Mais même si cette technique nous permet de définir les spécifictés de substrat de Zg4660 cela ne résoudra pas le problème de l'activation d'autres sulfatases venant d'organismes pour lesquels la transformation et la production de protéines *in situ* est impossible.

Pour pouvoir étudier toutes les sulfatases existantes, provenant de différents organismes, la mise au point d'un protocole universel de production et d'activation des sulfatases chez *E. coli* est indispensable. Pour ce faire, il est essentiel de poursuivre la mise au point du système d'expression différentielle des sulfatases et de la FGE au sein du périplasme. Dans ce contexte, il est nécessaire de tester différentes séquences d'adressage périplasmique, différentes concentrations d'inducteur de la traduction et également de déterminer le taux optimum d'induction de la FGE de l'organisme considéré, pour obtenir une grande quantité de sulfatases modifiées. De plus, les premières analyses de ma thèse indiquent que les sulfatases sont probablement obtenues sous forme active s'il y a une pause traductionnelle pendant la production de la protéine. Cette piste du contrôle cinétique de la traduction des sulfatases pour permettre la modification devra être poursuivie, car si cette hypothèse est confirmée, alors ce sera un élément essentiel à intégrer dans le système de production au sein du périplasme.

Concernant, l'étude du catabolisme des agars chez *Zobellia galactanivorans*, il serait intéressant de vérifier la taille des oligosaccharides importés par la bactérie, notamment pour vérifier quel est le substrat maximum disponible pour la GH16 *Zg*3516. Pour ce faire nous pourrions envisager de mettre la bactérie en présence d'oligosaccharides marqués de grandes tailles et vérifier si l'import périplasmique est possible, par validation en microscopie par exemple. De plus comparer la bactérie sauvage avec un

mutant de *Z. galactanivorans* où les β-agarases extracellulaires auraient été délétées pourrait également nous éclairer sur les mécanismes de dégradation et de transport des oligosaccharides en milieu physiologique.

*Zg*3516 est inclus dans le PUL37 dans lequel nous trouvons différentes sulfatases et une GDSL-lipase. Il serait nécessaire de caractériser ces enzymes pour confirmer les hypothèses que nous avons proposées concernant les différentes voies du catabolisme.

Par exemple nous pouvons citer le cas du porphyrane (PUL 42) est ce que ce sont les deux sulfatases de ce PUL qui permettent la désulfatation du porphyrane (L6S) ou est-ce que Zg3517 ou Zg3513, qui sont des sulfatases extracellulaires, qui assurent cette désulfatation afin de permettre l'import de l'oligosaccharide ? (**Figure 76**) Y aurait-il d'autres sulfatations qui sont impactées ? Plusieurs systèmes SusC/SusD sont proposés comme permettant l'import d'oligoagars. Cela pose la question des motifs reconnus/acceptés par ces protéines d'import périplasmique. En effet, nous avons identifié dans les trois PUL agissant sur les agars des gènes annotés comme codant des protéines SusC/SusD ce qui permet de s'interroger sur leur spécificité d'action ? Sont-elles différentes ? Sont-elles-complémentaires ? Comment ces PULs sont-ils activés : de façon synergique ou y a-t-il une spécificité en fonction d'un agar précis ?

Toutes ces questions portant sur la spécificité et les flux métaboliques au sein de la bactérie devront trouver une réponse si nous voulons être à même de comprendre le catabolisme des agars dans leur totalité.

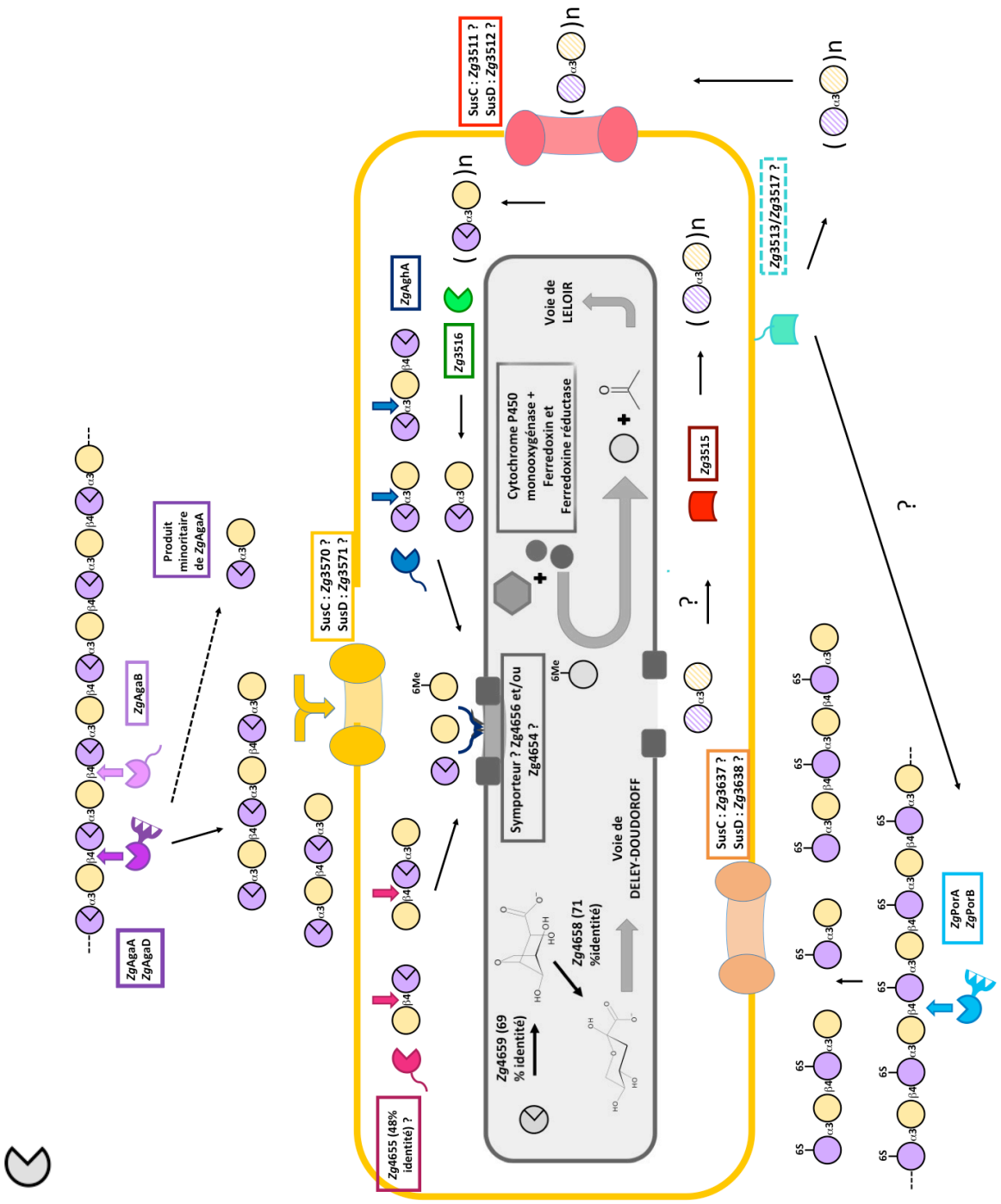


Figure 76 : Nouveau modèle de dégradation et relation entre les différents acteurs protéiques de la dégradation des agars par *Zobellia galactanivorans*

Les oligosaccharides générés par *Zg*AgaA, *Zg*AgaB et *Zg*AgaD sont représentés sous forme de DP4 et de DP6, mais nous ne connaissons pas la taille des oligosaccharides générés et internalisés par la bactérie dans le contexte physiologique. La partie grisée correspond à la partie non discutée ici mais détaillée dans la Figure 1 de l'introduction. Les sulfatases sont représentées sous forme de cylindre concave. Les oligosaccharides remplis avec des hachures signifient que nous ne connaissons pas leurs modifications. Est-ce que la voie de dégradation gérée par le PUL 37 et la voie impliquant les β-porphyranases (*Zg*PorA/*Zg*PorB) sont reliées ? Si oui, est-ce que les sulfatases extracellulaires du PUL 37 (*Zg*3513 et *Zg*3517) agissent avant les porphyranases ou après ? Ou bien est-ce que ce sont les enzymes périplasmiques *Zg*3515 et *Zg*3516 qui agissent après la voie de dégradation initiée par *Zg*PorA et *Zg*PorB ? Ou aucune de ces enzymes ?

Dans le PUL 42 nous retrouvons *Zg*PorE (β-porphyranase putative) et une sulfatase. Cette dernière permet t'elle de désulfater le L6S ? Dans ce cas quels rôles jouent les sulfatases du PUL37 ? De même, si ces PUL sont interconnectés, comment sont-ils induits ? Une autre question est : quelle relation existe-t'il

entre les β-agarases et le PUL37 ? En effet, le PUL37 n'est pas induit par l'agarose dans ce cas est-ce que *Zg*AghA et *Zg*4655 suffisent à dégrader les oligoagaroses ou bien y a t'il une autre enzyme impliquée ? Si cette enzyme existe, est-elle intracellulaire ou extracellulaire ? Est-ce que tous les SusC/SusD de *Z. galactanivorans* sont capables d'importer de longs oligosaccharides ?

VII. Bibliographie

Adl, S.M., Simpson, A.G.B., Farmer, M.A., Andersen, R.A., Anderson, O.R., Barta, J.R., Bowser, S.S., Brugerolle, G.U.Y., Fensome, R.A., Fredericq, S.*, et al.* (2005). The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists. Journal of Eukaryotic Microbiology *52*, 399-451.

Akoh, C.C., Lee, G.-C., Liaw, Y.-C., Huang, T.-H., and Shaw, J.-F. (2004). GDSL family of serine esterases/lipases. Progress in lipid research *43*, 534-552.

Allouch, J., Helbert, W., Henrissat, B., and 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.

Anderson, K.L., and Salyers, A.A. (1989a). Biochemical evidence that starch breakdown by Bacteroides thetaiotaomicron involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. Journal of bacteriology *171*, 3192-3198.

Anderson, K.L., and Salyers, A.A. (1989b). Genetic evidence that outer membrane binding of starch is required for starch utilization by Bacteroides thetaiotaomicron. Journal of bacteriology *171*, 3199-3204.

Araki, C. (1966). Some recent studies on the polysaccharides of agarophytes. Paper presented at: Proceedings of the Fifth International Seaweed Symposium, Halifax, August 25–28, 1965 (Elsevier).

Araki, S., Abe, S., Ando, S., Kon, K., Fujiwara, N., and Satake, M. (1989). Structure of phosphonoglycosphingolipid containing pyruvylated galactose in nerve fibers of Aplysia kurodai. Journal of Biological Chemistry *264*, 19922-19927.

Armisen, R., and Galatas, F. (1987). Production and utilization of products from commercial seaweeds, D.J. McHugh, ed. (Rome :: Food and Agriculture Organization of the United Nations).

Arnott, S., Fulmer, A., Scott, W.E., Dea, I.C.M., Moorhouse, R., and Rees, D.A. (1974). The agarose double helix and its function in agarose gel structure. Journal of Molecular Biology *90*, 269-284.

Atalla, R.H., and VanderHart, D.L. (1999). The role of solid state 13C NMR spectroscopy in studies of the nature of native celluloses. Solid State Nuclear Magnetic Resonance *15*, 1-19.

Barbeyron, T., Brillet-Gueguen, L., Carre, W., Carriere, C., Caron, C., Czjzek, M., Hoebeke, M., and Michel, G. (2016a). Matching the Diversity of Sulfated Biomolecules: Creation of a Classification Database for Sulfatases Reflecting Their Substrate Specificity. PloS one *11*, e0164846.

Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., and Potin, P. (2001). Zobellia galactanovorans gen. nov., sp. nov., a marine species of Flavobacteriaceae isolated from a red alga, and classification of [Cytophaga] uliginosa (ZoBell and Upham 1944)

Reichenbach 1989 as Zobellia uliginosa gen. nov., comb. nov. International journal of systematic and evolutionary microbiology *51*, 985-997.

Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., Goesmann, A., Leblanc, C., Oliver Glockner, F., Czjzek, M.*, et al.* (2016b). Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium Zobellia galactanivorans Dsij(T). Environmental microbiology *18*, 4610-4627.

Barbier, G., Oesterhelt, C., Larson, M.D., Halgren, R.G., Wilkerson, C., Garavito, R.M., Benning, C., and Weber, A.P.M. (2005). Comparative Genomics of Two Closely Related Unicellular Thermo-Acidophilic Red Algae, Galdieria sulphuraria and Cyanidioschyzon merolae, Reveals the Molecular Basis of the Metabolic Flexibility of Galdieria sulphuraria and Significant Differences in Carbohydrate Metabolism of Both Algae. Plant Physiology *137*, 460-474.

Bast, F. (2014). An illustrated review on cultivation and life history of agronomically important seaplants. Seaweed: mineral composition, nutritional and antioxidant benefits and agricultural uses Nova Publishers, New York, 39-70.

Baydoun, E.A.H., and Fry, S.C. (1988). [2-3H]Mannose Incorporation in Cultured Plant Cells: Investigation of L-Galactose Residues of the Primary Cell Wall. Journal of Plant Physiology *132*, 484-490.

Berteau, O., and 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*, 29R-40R.

Bixler, H.J., and Porse, H. (2011). A decade of change in the seaweed hydrocolloids industry. Journal of applied Phycology *23*, 321-335.

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 thet aiotaomicron, to the suckling period. Journal of biological chemistry.

Bojarova, P., and Williams, S.J. (2008). Sulfotransferases, sulfatases and formylglycine-generating enzymes: a sulfation fascination. Current opinion in chemical biology *12*, 573-581.

Boltes, I., Czapinska, H., Kahnert, A., von Bülow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M.A., and Usón, I. (2001). 1.3 Å structure of arylsulfatase from Pseudomonas aeruginosa establishes the catalytic mechanism of sulfate ester cleavage in the sulfatase family. Structure *9*, 483-491.

Bond, C.S., Clements, P.R., Ashby, S.J., Collyer, C.A., Harrop, S.J., Hopwood, J.J., and Guss, J.M. (1997). Structure of a human lysosomal sulfatase. Structure *5*, 277-289.

Bowker, D.M., and Turvey, J.R. (1968a). Water-soluble polysaccharides of the red alga Laurencia pinnatifida. Part I. Constituent units. Journal of the Chemical Society C: Organic, 983-988.

Bowker, D.M., and Turvey, J.R. (1968b). Water-soluble polysaccharides of the red alga Laurencia pinnatifida. Part II. Methylation analysis of the galactan sulphate. Journal of the Chemical Society C: Organic, 989-992.

Brawley, S.H., Blouin, N.A., Ficko-Blean, E., Wheeler, G.L., Lohr, M., Goodson, H.V., Jenkins, J.W., Blaby-Haas, C.E., Helliwell, K.E., Chan, C.X., *et al.* (2017). Insights into the red algae and eukaryotic evolution from the genome of Porphyra umbilicalis (Bangiophyceae, Rhodophyta). Proceedings of the National Academy of Sciences *114*, E6361-E6370.

Brown, E.D., Vivas, E.I., Walsh, C.T., and Kolter, R. (1995). MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in Escherichia coli. Journal of bacteriology *177*, 4194-4197.

Cardozo, K.H.M., Guaratini, T., Barros, M.P., Falcão, V.R., Tonon, A.P., Lopes, N.P., Campos, S., Torres, M.A., Souza, A.O., Colepicolo, P., *et al.* (2007). Metabolites from algae with economical impact. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology *146*, 60-78.

Carlucci, M.J., Mateu, C.G., Artuso, M.C., and Scolaro, L.A. (2012). Polysaccharides from red algae: Genesis of a renaissance. In The Complex World of Polysaccharides (InTech).

Cartmell, A., Lowe, E.C., Baslé, A., Firbank, S.J., Ndeh, D.A., Murray, H., Terrapon, N., Lombard, V., Henrissat, B., and Turnbull, J.E. (2017). How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. Proceedings of the National Academy of Sciences *114*, 7037-7042.

Cassano, V., Díaz-Larrea, J., Sentíes, A., Oliveira, M.C., Gil-Rodríguez, M.C., and Fujii, M.T. (2009). Evidence for The Conspecificity of Palisada Papillosa With P. Perforata (Ceramiales, Rhodophyta) from The Western and Eastern Atlantic Ocean on The Basis of Morphological and Molecular Analyses. Phycologia *48*, 86-100.

Chapman, E., Best, M.D., Hanson, S.R., and Wong, C.-H. (2004). Sulfotransferases: Structure, Mechanism, Biological Activity, Inhibition, and Synthetic Utility. Angewandte Chemie International Edition *43*, 3526-3548.

Cole, K.M., and Sheath, R.G. (1990). Biology of the Red Algae (Cambridge University Press).

Colin, S., Deniaud, E., Jam, M., Descamps, V., Chevolot, Y., Kervarec, N., Yvin, J.-C., Barbeyron, T., Michel, G., and 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.

Collén, J., Porcel, B., Carré, W., Ball, S.G., 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. Proceedings of the National Academy of Sciences *110*, 5247-5252.

Colley, K.J. (1997). Golgi localization of glycosyltransferases: more questions than answers. Glycobiology *7*, 1-13.

Correc, G., Hehemann, J.-H., Czjzek, M., and Helbert, W. (2011). Structural analysis of the degradation products of porphyran digested by Zobellia galactanivorans β-porphyranase A. Carbohydrate polymers *83*, 277-283.

Cosenza, V.A., Navarro, D.A., Ponce, N.M.A., and Stortz, C.A. (2017). Seaweed Polysaccharides: Structure and Applications. In Industrial Applications of Renewable Biomass Products (Springer), pp. 75-116.

Craigie, J.S. (1990). Cell walls. Biology of the red algae, 221-257.

Craigie, J.S., Wen, Z.C., and Van der Meer, J.P. (1984). Interspecific, Intraspecific and Nutritionally-Determined Variations in the Composition of. Agars from Gracilaria spp. Botanica Marina *27*, 55-62.

Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., Porter, N.T., Urs, K., Thompson, A.J., and Cartmell, A. (2015). Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. Nature *517*, 165.

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

de Graffenried, C.L., Laughlin, S.T., Kohler, J.J., and Bertozzi, C.R. (2004). A small-molecule switch for Golgi sulfotransferases. Proceedings of the National Academy of Sciences of the United States of America *101*, 16715-16720.

De Marco, A. (2007). Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in Escherichia coli. Nature protocols *2*, 2632.

de Reviers, B. (2002). Biologie et phylogénie des algues (Belin).

Deniaud Boüet, E. (2002). Structure et interactions des polymères dans les parois cellulaires de Palmaria palmata (Rhodophyta) (Paris 6).

Deniaud, E., Quemener, B., Fleurence, J., and Lahaye, M. (2003). Structural studies of the mix-linked β-(1→3)/ β-(1→4)-d-xylans from the cell wall of Palmaria palmata (Rhodophyta). International Journal of Biological Macromolecules *33*, 9-18.

Deniaud-Bouët, E., Hardouin, K., Potin, P., Kloareg, B., and Hervé, C. (2017). A review about brown algal cell walls and fucose-containing sulfated polysaccharides: Cell wall context, biomedical properties and key research challenges. Carbohydrate polymers *175*, 395-408.

Descamps, V., Colin, S., Lahaye, M., Jam, M., Richard, C., Potin, P., Barbeyron, T., Yvin, J.-C., and Kloareg, B. (2006). Isolation and culture of a marine bacterium degrading the sulfated fucans from marine brown algae. Marine biotechnology *8*, 27-39.

Diebold, M.-L., Fribourg, S., Koch, M., Metzger, T., and Romier, C. (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. Journal of structural biology *175*, 178-188.

Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M.A., and von Figura, K. (1998). Posttranslational formation of formylglycine in prokaryotic sulfatases by modification of either cysteine or serine. Journal of Biological Chemistry *273*, 25560-25564.

Dixon, P.S. (1973). Biology of the Rhodophyta, Vol 4 (Oliver & Boyd Edinburgh).

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallographica Section D: Biological Crystallography *60*, 2126-2132.

Ferrara, M.C., Cobucci-Ponzano, B., Carpentieri, A., Henrissat, B., Rossi, M., Amoresano, A., and Moracci, M. (2014). The identification and molecular characterization of the first archaeal bifunctional exo-β-glucosidase/N-acetyl-β-glucosaminidase demonstrate that family GH116 is made of three functionally distinct subfamilies. Biochimica et Biophysica Acta (BBA)-General Subjects *1840*, 367-377.

Ferreira, L.G., Noseda, M.D., Gonçalves, A.G., Ducatti, D.R.B., Fujii, M.T., and Duarte, M.E.R. (2012). Chemical structure of the complex pyruvylated and sulfated agaran from the red seaweed Palisada flagellifera (Ceramiales, Rhodophyta). Carbohydrate research *347*, 83-94.

Ficko-Blean, E., Duffieux, D., Rebuffet, É., Larocque, R., Groisillier, A., Michel, G., and Czjzek, M. (2015). Biochemical and structural investigation of two paralogous glycoside hydrolases from Zobellia galactanivorans: novel insights into the evolution, dimerization plasticity and catalytic mechanism of the GH117 family. Acta Crystallographica Section D: Biological Crystallography *71*, 209-223.

Ficko-Blean, E., Prechoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., Stam, M., Genicot, S., Jam, M., Calteau, A.*, et al.* (2017). Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. Nature communications *8*, 1685.

Flament, D., Barbeyron, T., Jam, M., Potin, P., Czjzek, M., Kloareg, B., and Michel, G. (2007). Alpha-agarases define a new family of glycoside hydrolases, distinct from beta-agarase families. Applied and environmental microbiology *73*, 4691-4694.

Frey, P.A. (1996). The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. The FASEB Journal *10*, 461-470.

Gabrielson, P.W., Garbary, D.J., and Scagel, R.F. (1985). The nature of the ancestral red alga: Inferences from a cladistic analysis. Biosystems *18*, 335-346.

Gardner, J.G. (2016). Polysaccharide degradation systems of the saprophytic bacterium Cellvibrio japonicus. World Journal of Microbiology and Biotechnology *32*, 121.

Genicot-Joncour, S., Poinas, A., Richard, O., Potin, P., Rudolph, B., Kloareg, B., and Helbert, W. (2009). The Cyclization of the 3,6-Anhydro-Galactose Ring of ι-Carrageenan Is Catalyzed by Two Galactose-2,6-Sulfurylases in the Red Alga Chondrus crispus. Plant Physiology *151*, 1609-1616.

Ghosh, D. (2005). Three‐dimensional structures of sulfatases. Methods in enzymology *400*, 273-293.

Ghosh, T., Pujol, C.A., Damonte, E.B., Sinha, S., and Ray, B. (2009). Sulfated Xylomannans from the Red Seaweed Sebdenia Polydactyla: Structural Features, Chemical Modification and Antiviral Activity. Antiviral Chemistry and Chemotherapy *19*, 235-242.

Gibeaut, D.M. (2000). Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides. Plant Physiology and Biochemistry *38*, 69-80.

Groisillier, A., Herve, C., Jeudy, A., Rebuffet, E., Pluchon, P.F., Chevolot, Y., Flament, D., Geslin, C., Morgado, I.M., Power, D.*, et al.* (2010). MARINE-EXPRESS: taking advantage of high throughput cloning and expression strategies for the post-genomic analysis of marine organisms. Microbial cell factories *9*, 45.

Grondin, J.M., Tamura, K., Déjean, G., Abbott, D.W., and Brumer, H. (2017). Polysaccharide utilization loci: fuelling microbial communities. Journal of bacteriology, JB. 00860-00816.

Grünewald, N., and Alban, S. (2009). Optimized and Standardized Isolation and Structural Characterization of Anti-inflammatory Sulfated Polysaccharides from the Red Alga Delesseria sanguinea (Hudson) Lamouroux (Ceramiales, Delesseriaceae). Biomacromolecules *10*, 2998-3008.

Grünewald, N., Groth, I., and Alban, S. (2009). Evaluation of Seasonal Variations of the Structure and Anti-inflammatory Activity of Sulfated Polysaccharides Extracted from the Red Alga Delesseria sanguinea (Hudson) Lamouroux (Ceramiales, Delesseriaceae). Biomacromolecules *10*, 1155-1162.

Hanson, S.R., Best, M.D., and Wong, C.H. (2004). Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. Angewandte Chemie *43*, 5736-5763.

Hatcher, E., Soudackov, A.V., and Hammes-Schiffer, S. (2004). Proton-coupled electron transfer in soybean lipoxygenase. Journal of the American Chemical Society *126*, 5763-5775.

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.

Hehemann, J.H., Correc, G., Thomas, F., Bernard, T., Barbeyron, T., Jam, M., Helbert, W., Michel, G., and Czjzek, M. (2012a). Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium Zobellia galactanivorans. The Journal of biological chemistry *287*, 30571-30584.

Hehemann, J.H., Kelly, A.G., Pudlo, N.A., Martens, E.C., and Boraston, A.B. (2012b). Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. Proceedings of the National Academy of Sciences of the United States of America *109*, 19786-19791.

Hehemann, J.H., Smyth, L., Yadav, A., Vocadlo, D.J., and Boraston, A.B. (2012c). Analysis of keystone enzyme in Agar hydrolysis provides insight into the degradation (of a polysaccharide from) red seaweeds. The Journal of biological chemistry *287*, 13985-13995.

Helbert, W. (2017). Marine polysaccharide sulfatases. Frontiers in Marine Science *4*, 6.

Hemmingson, J.A., Furneaux, R.H., and Murray-Brown, V.H. (1996a). Biosynthesis of agar polysaccharides in Gracilaria chilensis bird, McLachlan et Oliveira. Carbohydrate research *287*, 101-115.

Hemmingson, J.A., Furneaux, R.H., and Wong, H. (1996b). In vivo conversion of 6-O-sulfo-l-galactopyranosyl residues into 3,6-anhydro-l-galactopyranosyl residues in Gracilaria chilensis Bird, McLachlan et Oliveira. Carbohydrate research *296*, 285-292.

Hennet, T. (2002). The galactosyltransferase family. Cellular and Molecular Life Sciences CMLS *59*, 1081-1095.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. The Biochemical journal *280 (Pt 2)*, 309-316.

Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. Biochemical Journal *316*, 695.

Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.P., and Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proceedings of the National Academy of Sciences of the United States of America *92*, 7090-7094.

Henrissat, B., and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. Current opinion in structural biology *7*, 637-644.

Hernandez-Guzman, F.G., Higashiyama, T., Pangborn, W., Osawa, Y., and Ghosh, D. (2003). Structure of human estrone sulfatase suggests functional roles of membrane association. Journal of Biological Chemistry.

Hettle, A.G., Vickers, C., Robb, C.S., Liu, F., Withers, S.G., Hehemann, J.H., and Boraston, A.B. (2018). The Molecular Basis of Polysaccharide Sulfatase Activity and a Nomenclature for Catalytic Subsites in this Class of Enzyme. Structure *26*, 747-758 e744.

Hori, H., and Osawa, S. (1987). Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. Molecular Biology and Evolution *4*, 445-472.

Jam, M., Flament, D., Allouch, J., Potin, P., Thion, L., Kloareg, B., Czjzek, M., Helbert, W., Michel, G., and Barbeyron, T. (2005). The endo-beta-agarases AgaA and AgaB from the marine bacterium Zobellia galactanivorans: two paralogue enzymes with different molecular organizations and catalytic behaviours. The Biochemical journal *385*, 703-713.

Jimenez-Barbero, J., Bouffar-Roupe, C., Rochas, C., and Pérez, S. (1989). Modelling studies of solvent effects on the conformational stability of agarobiose and neoagarobiose and their relationship to agarose. International journal of biological macromolecules *11*, 265-272.

Jonas, S., van Loo, B., Hyvönen, M., and Hollfelder, F. (2008). A new member of the alkaline phosphatase superfamily with a formylglycine nucleophile: structural and kinetic characterisation of a phosphonate monoester hydrolase/phosphodiesterase from Rhizobium leguminosarum. Journal of molecular biology *384*, 120-136.

Kabsch, W. (2010). Xds. Acta Crystallographica Section D: Biological Crystallography *66*, 125-132.

Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N., Kimura, A., Tanaka, I., and Yao, M. (2008). Structural and functional analysis of a glycoside hydrolase family 97 enzyme from Bacteroides thetaiotaomicron. Journal of Biological Chemistry *283*, 36328-36337.

Kleywegt, G.J., Zou, J.-Y., Divne, C., Davies, G.J., Sinning, I., Ståhlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T.T., and Jones, T.A. (1997). The crystal structure of the catalytic core domain of endoglucanase I from Trichoderma reesei at 3.6 Å resolution, and a comparison with related enzymes1. Journal of molecular biology *272*, 383-397.

Kloareg, B., and Quatrano, R.S. (1988). Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. Oceanography and marine biology: an annual review *26*, 259-315.

Knutsen, S.H., Myslabodski, D.E., Larsen, B., and Usov, A.I. (1994). A modified system of nomenclature for red algal galactans. Botanica marina *37*, 163-170.

Koshland Jr, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. Biological reviews *28*, 416-436.

Kylin, J.H. (1956). Die gattungen der Rhodophyceen (CWK Gleerups Forlag).

Lahaye, M. (2001). Developments on gelling algal galactans, their structure and physico-chemistry. Journal of Applied Phycology *13*, 173-184.

Lahaye, M., and Robic, A. (2007). Structure and functional properties of ulvan, a polysaccharide from green seaweeds. Biomacromolecules *8*, 1765-1774.

Lahaye, M., Rondeau-Mouro, C., Deniaud, E., and Buléon, A. (2003). Solid-state 13C NMR spectroscopy studies of xylans in the cell wall of Palmaria palmata (L. Kuntze, Rhodophyta). Carbohydrate research *338*, 1559-1569.

Landráu, M.E., and Welschmeyer, N.A. (2000). THE UBIQUITOUS PRESENCE OF CHLOROPHYLL D IN RED ALGAE FROM THE MONTEREY PENINSULA. Journal of Phycology *36*, 41-41.

Lechat, H. (1998). Nature et organisation des polysaccharides parietaux de l'algue rouge eucheuma cottonii (Nantes).

Lee, C.H., Kim, H.T., Yun, E.J., Lee, A.R., Kim, S.R., Kim, J.-H., Choi, I.-G., and Kim, K.H. (2014). A novel agarolytic β-galactosidase acts on agarooligosaccharides for complete hydrolysis of agarose into monomers. Applied and environmental microbiology, AEM. 01577-01514.

Lee, R.E. (2008). Phycology. Edit (Cambridge University Press. New York, USA).

Lee, W.-K., Namasivayam, P., Ong Abdullah, J., and Ho, C.-L. (2017a). Transcriptome profiling of sulfate deprivation responses in two agarophytes Gracilaria changii and Gracilaria salicornia (Rhodophyta). Scientific Reports *7*, 46563.

Lee, W.K., Lim, Y.Y., Leow, A.T., Namasivayam, P., Ong Abdullah, J., and Ho, C.L. (2017b). Biosynthesis of agar in red seaweeds: A review. Carbohydrate polymers *164*, 23-30.

Leloir, L.F. (1951). The enzymatic transformation of uridine diphosphate glucose into a galactose derivative. Archives of biochemistry and biophysics *33*, 186-190.

Li, L., Saga, N., and Mikami, K. (2008). Phosphatidylinositol 3-kinase activity and asymmetrical accumulation of F-actin are necessary for establishment of cell polarity in the early development of monospores from the marine red alga Porphyra yezoensis. Journal of experimental botany *59*, 3575-3586.

Li, L., Saga, N., and Mikami, K. (2009). Ca2+ influx and phosphoinositide signalling are essential for the establishment and maintenance of cell polarity in monospores from the red alga Porphyra yezoensis. Journal of experimental botany *60*, 3477-3489.

Li, S.-Y., Lellouche, J.-P., Shabtai, Y., and Arad, S. (2001). FIXED CARBON PARTITIONING IN THE RED MICROALGA PORPHYRIDIUM SP. (RHODOPHYTA). Journal of Phycology *37*, 289-297.

Li, S.-Y., Shabtai, Y., and Arad, S. (2002). FLORIDOSIDE AS A CARBON PRECURSOR FOR THE SYNTHESIS OF CELL-WALL POLYSACCHARIDE IN THE RED MICROALGA PORPHYRIDIUM SP. (RHODOPHYTA)1. Journal of Phycology *38*, 931-938.

Liao, M.L., Kraft, G.T., Munro, S.L.A., Craik, D.J., and Bacic, A. (1993). BETA/KAPPA‐CARRAGEENANS AS EVIDENCE FOR CONTINUED SEPARATION OF THE FAMILIES DICRANEMATACEAE AND SARCODIACEAE (GIGARTINALES, RHODOPHYTA) 1. Journal of Phycology *29*, 833-844.

Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., Von Figura, K., and Saenger, W. (1998). Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. Biochemistry *37*, 3654-3664.

Lyko, F. (2017). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nature Reviews Genetics *19*, 81.

Macler, B.A. (1986). Regulation of Carbon Flow by Nitrogen and Light in the Red Alga, Gelidium coulteri. Plant Physiology *82*, 136-141.

Major, L.L., Wolucka, B.A., and Naismith, J.H. (2005a). Structure and function of GDP-mannose-3',5'-epimerase: an enzyme which performs three chemical reactions at the same active site. Journal of the American Chemical Society *127*, 18309-18320.

Major, L.L., Wolucka, B.A., and Naismith, J.H. (2005b). Structure and Function of GDP-Mannose-3‘,5‘-Epimerase: An Enzyme which Performs Three Chemical Reactions at the Same Active Site. Journal of the American Chemical Society *127*, 18309-18320.

Martens, E.C., Koropatkin, N.M., Smith, T.J., and Gordon, J.I. (2009). Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. The Journal of biological chemistry *284*, 24673-24677.

McBride, M.J., and Kempf, M.J. (1996). Development of techniques for the genetic manipulation of the gliding bacterium Cytophaga johnsonae. Journal of bacteriology *178*, 583-590.

McCarter, J.D., and Withers, G.S. (1994). Mechanisms of enzymatic glycoside hydrolysis. Current opinion in structural biology *4*, 885-892.

McHugh, D.J. (2003). A guide to the seaweed industry (Food and Agriculture Organization of the United Nations Rome).

McIntosh, L.P., Hand, G., Johnson, P.E., Joshi, M.D., Körner, M., Plesniak, L.A., Ziser, L., Wakarchuk, W.W., and Withers, S.G. (1996). The p K a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a 13C-NMR study of Bacillus circulans xylanase. Biochemistry *35*, 9958-9966.

McLean, M.W., and Williamson, F.B. (1979). Glycosulphatase from Pseudomonas carrageenovora: purification and some properties. European journal of biochemistry *101*, 497-505.

Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001). The kappa-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.

Michel, G., Nyval-Collen, P., Barbeyron, T., Czjzek, M., and Helbert, W. (2006). Bioconversion of red seaweed galactans: a focus on bacterial agarases and carrageenases. Applied microbiology and biotechnology *71*, 23-33.

Michel, G., Tonon, T., Scornet, D., Cock, J.M., and Kloareg, B. (2010). Central and storage carbon metabolism of the brown alga Ectocarpus siliculosus: insights into the origin and evolution of storage carbohydrates in Eukaryotes. The New phytologist *188*, 67-81.

Miller, G., Furneaux, C.W., Davis, P., Love, P., and O'Donnell, A. (2009). Built environment procurement practice: Impediments to innovation and opportunities for changes.

Miller, I.J. (2003). The chemical structure of galactans from Sarcodia montagneana and from Sarcodia flabellata. Botanica marina *46*, 392-399.

Miller, I.J. (2005). Evaluation of the structures of polysaccharides from three taxa in the genus Hymenena and from Acrosorium decumbens (Rhodophyta, Delesseriaceae). Botanica Marina *48*, 148-156.

Miller, I.J., and Blunt, J.W. (2002). Evaluation of the structure of the polysaccharides from Chondria macrocarpa and Ceramium rubrum as determined by 13C NMR spectroscopy. Botanica marina *45*, 1-8.

Mukai, L.S., Craigie, J.S., and Brown, R.G. (1981). CHEMICAL COMPOSITION AND STRUCTURE OF THE CELL WALLS OF THE CONCHOCELIS AND THALLUS PHASES OF PORPHYRA TENERA (RHODOPHYCEAE)1,2. Journal of Phycology *17*, 192-198.

Müller, K.M., Cannone, J.J., Gutell, R.R., and Sheath, R.G. (2001). A Structural and Phylogenetic Analysis of the Group IC1 Introns in the Order Bangiales (Rhodophyta). Molecular Biology and Evolution *18*, 1654-1667.

Nishiyama, Y., Sugiyama, J., Chanzy, H., and Langan, P. (2003). Crystal Structure and Hydrogen Bonding System in Cellulose Iα from Synchrotron X-ray and Neutron Fiber Diffraction. Journal of the American Chemical Society *125*, 14300-14306.

Ohta, Y., Hatada, Y., Miyazaki, M., Nogi, Y., Ito, S., and Horikoshi, K. (2005). Purification and characterization of a novel α-agarase from a Thalassomonas sp. Current microbiology *50*, 212-216.

Otwinowski, Z., and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. In Methods in enzymology (Elsevier), pp. 307-326.

Parsons, M.J., Pickmere, S.E., and Bailey, R.W. (1977). Carrageenan composition in New Zealand species of Gigartina (Rhodophyta): Geographic variation and interspecific differences. New Zealand journal of botany *15*, 589-595.

Patterson, J.H., McConville, M.J., Haites, R.E., Coppel, R.L., and Billman-Jacobe, H. (2000). Identification of a methyltransferase from Mycobacterium smegmatis involved in glycopeptidolipid synthesis. Journal of Biological Chemistry.

Percival, E., and McDowell, R.H. (1990). 15 - Algal Polysaccharides. In Methods in Plant Biochemistry, P.M. Dey, ed. (Academic Press), pp. 523-547.

Perez, S., and Kouwijzer, M. (1999). Shapes and interactions of polysaccharide chains. In Carbohydrates (Springer), pp. 258-293.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. Acta Crystallographica Section D: Biological Crystallography *57*, 1445-1450.

Petersen, B.L., Egelund, J., Damager, I., Faber, K., Krüger Jensen, J., Yang, Z., Bennett, E.P., Scheller, H.V., and Ulvskov, P. (2009). Assay and heterologous expression in Pichia pastoris of plant cell wall type-II membrane anchored glycosyltransferases. Glycoconjugate Journal *26*, 1235.

Pierre, G., Delattre, C., Laroche, C., and Michaud, P. (2015). Galactans and its applications. Polysaccharides: Bioactivity and Biotechnology, 753-794.

Pleschberger, M., Hildner, F., Rünzler, D., Gelbmann, N., Mayer, H.F., Sleytr, U.B., and Egelseer, E.M. (2013). Identification of a novel gene cluster in the upstream region of the S-layer gene sbpA involved in cell wall metabolism of Lysinibacillus sphaericus CCM 2177 and characterization of the recombinantly produced autolysin and pyruvyl transferase. Archives of Microbiology *195*, 323-337.

Pluvinage, B., Hehemann, J.-H., and Boraston, A.B. (2013). Substrate recognition and hydrolysis by a family 50 exo-β-agarase, Aga50D, from the marine bacterium Saccharophagus degradans. Journal of Biological Chemistry *288*, 28078-28088.

Popper, Z.A., Michel, G., Herve, C., Domozych, D.S., Willats, W.G., Tuohy, M.G., Kloareg, B., and Stengel, D.B. (2011). Evolution and diversity of plant cell walls: from algae to flowering plants. Annual review of plant biology *62*, 567-590.

Potin, P., Richard, C., Rochas, C., and Kloareg, B. (1993). Purification and characterization of the α - agarase from Alteromonas agarlyticus (Cataldi) comb. nov., strain GJ1B. European journal of biochemistry *214*, 599-607.

Potin, P., Sanseau, A., Le Gall, Y., Rochas, C., and Kloareg, B. (1991). Purification and characterization of a new k‐carrageenase from a marine Cytophaga‐like bacterium. European journal of biochemistry *201*, 241-247.

Powell, A.K., Ahmed, Y.A., Yates, E.A., and Turnbull, J.E. (2010). Generating heparan sulfate saccharide libraries for glycomics applications. nature protocols *5*, 821.

Prado, H.J., Ciancia, M., and Matulewicz, M.C. (2008). Agarans from the red seaweed Polysiphonia nigrescens (Rhodomelaceae, Ceramiales). Carbohydrate research *343*, 711-718.

Préchoux, A., Genicot, S., Rogniaux, H., and Helbert, W. (2013). Controlling carrageenan structure using a novel formylglycine-dependent sulfatase, an endo-4S-iota-carrageenan sulfatase. Marine biotechnology *15*, 265-274.

Préchoux, A., Genicot, S., Rogniaux, H., and Helbert, W. (2016). Enzyme-assisted preparation of furcellaran-like κ-/β-carrageenan. Marine biotechnology *18*, 133-143.

Prosselkov, P.V., Gross, W., Igamberdiev, A.U., and Schnarrenberger, C. (1996). Purification and characterization of UDP-D-galactose 4-epimerase from the red alga Galdieria sulphuraria. Physiologia Plantarum *98*, 753-758.

Ragan, M.A., and Gutell, R.R. (1995). Are red algae plants? Botanical Journal of the Linnean Society *118*, 81-105.

Rajan, S.S., Yang, X., Collart, F., Yip, V.L.Y., Withers, S.G., Varrot, A., Thompson, J., Davies, G.J., and Anderson, W.F. (2004). Novel catalytic mechanism of glycoside hydrolysis based on the structure of an NAD+/Mn2+-dependent phospho-α-glucosidase from Bacillus subtilis. Structure *12*, 1619-1629.

Ramawat, K.G., and Mérillon, J.-M. (2015). Polysaccharides: bioactivity and biotechnology (Springer).

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. Environmental microbiology *13*, 1253-1270.

Rees, D.A. (1961). Enzymic synthesis of 3:6-anhydro-l-galactose within porphyran from l-galactose 6-sulphate units. Biochemical Journal *81*, 347-352.

Reintjes, G., Arnosti, C., Fuchs, B.M., and Amann, R. (2017). An alternative polysaccharide uptake mechanism of marine bacteria. The ISME journal *11*, 1640.

Reisky, L., Büchsenschütz, H.C., Engel, J., Song, T., Schweder, T., Hehemann, J.-H., and Bornscheuer, U.T. (2018). Oxidative demethylation of algal carbohydrates by cytochrome P450 monooxygenases. Nature chemical biology *14*, 342.

Reyes, F., and Orellana, A. (2008). Golgi transporters: opening the gate to cell wall polysaccharide biosynthesis. Current Opinion in Plant Biology *11*, 244-251.

Roeser, D., Preusser-Kunze, A., Schmidt, B., Gasow, K., Wittmann, J.G., Dierks, T., von Figura, K., and Rudolph, M.G. (2006). A general binding mechanism for all human sulfatases by the formylglycine-generating enzyme. Proceedings of the National Academy of Sciences *103*, 81-86.

Ruiz, O.A., and Ugalde, R.A. (1998). Partial characterization and photolabeling of a Rhizobium meliloti polysaccharide methyltransferase with S-adenosylmethionine. Int Microbiol *1*, 225-230.

Saunders, G.W., and Hommersand, M.H. (2004). Assessing red algal supraordinal diversity and taxonomy in the context of contemporary systematic data. American Journal of Botany *91*, 1494-1507.

Saunders, G.W., and Kraft, G.T. (1997). A molecular perspective on red algal evolution: focus on the Florideophycidae. In Origins of Algae and their Plastids, D. Bhattacharya, ed. (Vienna: Springer Vienna), pp. 115-138.

Schmidt, B., Selmer, T., Ingendoh, A., and von Figurat, K. (1995). A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. Cell *82*, 271-278.

Schmidt-Nielsen, K. (1997). Animal Physiology: Adaptation and Environment (Cambridge University Press).

Schmitz, F. (1892). Klasse Rhodophyceae 2. Unterklasse Florideae. Syllabus der Vorlesungen Uber Specielle und Medicinisch-Pharmaceutische Botanik Grosse Ausgabe.

Sharp, P.M., Cowe, E., Higgins, D.G., Shields, D.C., Wolfe, K.H., and Wright, F. (1988). Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity. Nucleic acids research *16*, 8207-8211.

Sheldrick, G.M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallographica Section D *66*, 479-485.

Siow, R.-S., Teo, S.-S., Ho, W.-Y., Shukor, M.Y.A., Phang, S.-M., and Ho, C.-L. (2012). MOLECULAR CLONING AND BIOCHEMICAL CHARACTERIZATION OF GALACTOSE-1-PHOSPHATE URIDYLYLTRANSFERASE FROM GRACILARIA CHANGII (RHODOPHYTA)1. Journal of Phycology *48*, 155-162.

Siow, R.-S., Teoh, S., Teo, S.-S., bin Abd. Shukor, M.Y., Phang, S.-M., and Ho, C.-L. (2013). Molecular cloning and characterization of GDP-mannose-3′,5′-epimerase from Gracilaria changii. Journal of Applied Phycology *25*, 1309-1318.

Smirnoff, N., and Wheeler, G.L. (2000). Ascorbic acid in plants: biosynthesis and function. Critical reviews in biochemistry and molecular biology *35*, 291-314.

Spillmann, D., Hård, K., Thomas-Oates, J., Vliegenthart, J.F., Misevic, G., Burger, M.M., and Finne, J. (1993). Characterization of a novel pyruvylated carbohydrate unit implicated in the cell aggregation of the marine sponge Microciona prolifera. Journal of Biological Chemistry *268*, 13378-13387.

St John, F.J., González, J.M., and Pozharski, E. (2010). Consolidation of glycosyl hydrolase family 30: a dual domain 4/7 hydrolase family consisting of two structurally distinct groups. FEBS letters *584*, 4435-4441.

Steven, G.B., and Matthew, K.M. (2003). From Bacterial Glycogen to Starch: Understanding the Biogenesis of the Plant Starch Granule. Annual review of plant biology *54*, 207-233.

Stiger-Pouvreau, V., Bourgougnon, N., and Deslandes, E. (2016). Chapter 8 - Carbohydrates From Seaweeds. In Seaweed in Health and Disease Prevention, J. Fleurence, and I. Levine, eds. (San Diego: Academic Press), pp. 223-274.

Sutherland, J.E., Lindstrom, S.C., Nelson, W.A., Brodie, J., Lynch, M.D.J., Hwang, M.S., Choi, H.-G., Miyata, M., Kikuchi, N., Oliveira, M.C.*, et al.* (2011). A NEW LOOK AT AN ANCIENT ORDER: GENERIC REVISION OF THE BANGIALES (RHODOPHYTA)1. Journal of Phycology *47*, 1131-1151.

Szumiło, T. (1981). Pathway for D-galactonate catabolism in nonpathogenic mycobacteria. Journal of bacteriology *148*, 368-370.

Terrapon, N., and Henrissat, B. (2014). How do gut microbes break down dietary fiber? Trends in biochemical sciences *39*, 156-158.

Thomas, F., Bordron, P., Eveillard, D., and Michel, G. (2017). Gene Expression Analysis of Zobellia galactanivorans during the Degradation of Algal Polysaccharides Reveals both Substrate-Specific and Shared Transcriptome-Wide Responses. Frontiers in microbiology *8*, 1808.

Thomas, F., Hehemann, J.H., Rebuffet, E., Czjzek, M., and Michel, G. (2011). Environmental and gut bacteroidetes: the food connection. Frontiers in microbiology *2*, 93.

Torode, T. A., Marcus, S. E., Jam, M., Tonon, T., Blackburn, R. S., Hervé, C., & Knox, J. P. (2015). Monoclonal antibodies directed to fucoidan preparations from brown algae. *PLoS One*, *10*(2), e0118366.

Tsekos, I. (1981). Growth and differentiation of the Golgi apparatus and wall formation during carposporogenesis in the red alga, Gigartina teedii (Roth) Lamour. Journal of cell science *52*, 71-84.

Tveter-Gallagher, E., Cogburn, J.N., Mathieson, A.C., and Schiff, J.A. (1984). Location of carrageenan and incorporation of 35SO2–4 into sulfur amino acids in the red alga Chondrus crispus (Dordrecht: Springer Netherlands).

Reyes-Prieto, A., Weber, A. P., & Bhattacharya, D. (2007). The origin and establishment of the plastid in algae and plants. *Annu. Rev. Genet.*, *41*, 147-168.

Ullah, H., Scappini, E.L., Moon, A.F., Williams, L.V., Armstrong, D.L., and Pedersen, L.C. (2008). Structure of a signal transduction regulator, RACK1, from Arabidopsis thaliana. Protein Science *17*, 1771-1780.

Urbániková, Ľ., Vršanská, M., Krogh, K.B.R.M., Hoff, T., and Biely, P. (2011). Structural basis for substrate recognition by Erwinia chrysanthemi GH30 glucuronoxylanase. The FEBS journal *278*, 2105-2116.

Usov, A.I. (2011). Polysaccharides of the red algae. Advances in carbohydrate chemistry and biochemistry *65*, 115-217.

Usov, A.I., and Ivanova, E.G. (1987). Polysaccharides of Algae. XXXVII. Characterization of Hybrid Structure of Substituted Agarose from Polysiphonia morrowii (Rhodophyta, Rhodomelaceae) Using β-Agarase and 13C-NMR Spectroscopy (Walter de Gruyter, Berlin/New York).

Vagin, A., and Lebedev, A. (2015). MoRDa, an automatic molecular replacement pipeline. Acta Crystallographica Section A: Foundations and Advances *71*, s19-s19.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. Journal of applied crystallography *30*, 1022-1025.

Van de Velde, F., Knutsen, S.H., Usov, A.I., Rollema, H.S., and Cerezo, A.S. (2002). 1H and 13C high resolution NMR spectroscopy of carrageenans: application in research and industry. Trends in Food Science & Technology *13*, 73-92.

Vries, J., and Archibald, J.M. (2018). Plant evolution: Landmarks on the path to terrestrial life. New Phytologist *217*, 1428-1434.

Weigl, J., and Yaphe, W. (1966). The enzymic hydrolysis of carrageenan by Pseudomonas carrageenovora: purification of a κ-carrageenase. Canadian journal of microbiology *12*, 939-947.

Wheeler, G.L., Jones, M.A., and Smirnoff, N. (1998). The biosynthetic pathway of vitamin C in higher plants. Nature *393*, 365-369.

Wiederschain, G.Y. (2009). Essentials of glycobiology. Biochemistry (Moscow) *74*, 1056-1056.

Winn, M.D., Murshudov, G.N., and Papiz, M.Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. In Methods in enzymology (Elsevier), pp. 300-321.

Wong, T.Y., and Yao, X.T. (1994). The DeLey-Doudoroff pathway of galactose metabolism in Azotobacter vinelandii. Applied and environmental microbiology *60*, 2065-2068.

Yang, E.C., Boo, S.M., Bhattacharya, D., Saunders, G.W., Knoll, A.H., Fredericq, S., Graf, L., and Yoon, H.S. (2016). Divergence time estimates and the evolution of major lineages in the florideophyte red algae. Scientific Reports *6*, 21361.

Yip, V.L.Y., Varrot, A., Davies, G.J., Rajan, S.S., Yang, X., Thompson, J., Anderson, W.F., and Withers, S.G. (2004). An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a family 4 β-glycosidase from Thermotoga maritima. Journal of the American Chemical Society *126*, 8354-8355.

Yu, S., and Pedersén, M. (1990). The α-galactosidase of Gracilaria tenuistipitata and G. sordida (Gracilariales, Rhodophyta). Phycologia *29*, 454-460.

Yun, E.J., Lee, S., Kim, H.T., Pelton, J.G., Kim, S., Ko, H.J., Choi, I.G., and Kim, K.H. (2015). The novel catabolic pathway of 3, 6‐anhydro‐L‐galactose, the main component of red macroalgae, in a marine bacterium. Environmental microbiology *17*, 1677-1688.

Zhang, B., Liu, X., Qian, Q., Liu, L., Dong, G., Xiong, G., Zeng, D., and Zhou, Y. (2011). Golgi nucleotide sugar transporter modulates cell wall biosynthesis and plant growth in rice. Proceedings of the National Academy of Sciences.

Zhu, Y., Thomas, F., Larocque, R., Li, N., Duffieux, D., Cladiere, L., Souchaud, F., Michel, G., and McBride, M.J. (2017). Genetic analyses unravel the crucial role of a horizontally acquired alginate lyase for brown algal biomass degradation by Zobellia galactanivorans. Environmental microbiology *19*, 2164-2181.

Les agars sont des polysaccharides majeurs issus de la paroi des algues rouges. Ces polysaccharides sont composés par des résidus D- et L-galactose alternativement liés par des liaisons glycosidiques β-1,4 et α-1,3. Ces galactanes présentent de nombreuses modifications: sulfatations, méthylations, pyruvilation, branchement, etc.. Toutes ces modifications rendent complexe la dégradation ces polysaccharides pour des bactéries marines hétérotrophes.
Zobellia galactanivorans est une flavobactérie marine capable de dégrader de nombreux polysaccharides d'algues dont les agars. L'objectif de cette thèse était d'identifier et de caractériser les outils enzymatiques dont dispose cette bactérie pour dégrader les agars complexes.
Les deux sujets d'étude de cette thèse sont une β-agarase divergente (*Zg*AgaC) et des sulfatases actives sur des agars. Pour pouvoir réaliser cette étude portant sur des enzymes spécifique des agars, nous avons mis au point des cribles d'activités sur une collection d'agars naturels qui a été préparée au cours de cette thèse. Ces cribles ont permis de mettre en évidence le comportement divergent de *Zg*AgaC sur des agars complexes, par rapport aux autres β-agarases et β-porphyranases de la famille 16 des glycoside hydrolases (GH16). Ils ont aussi permis d'identifier les agars complexes sur lesquels deux sulfatases différentes sont actives.
Pour pouvoir caractériser l'action de ces enzymes il était nécessaire de pouvoir purifier les oligosaccharides générés par *Zg*AgaC et ceux sur lesquels les sulfatases sont actives. Une méthode de purification des oligosaccharides sulfatés a donc été mise au point pour caractériser les substrats de ces enzymes. Ainsi les oligosaccharides produits par *Zg*AgaC ont pu être caractérisés par spectrométrie de masse. Combinés à la résolution de la structure cristalline de *Zg*AgaC, un modèle de reconnaissance des oligosaccharides par ce nouveau type d'agarase a pu être proposé.
Concernant les sulfatases, un protocole de production recombinante chez la bactérie *Escherichia coli* a été mis au point pour les produire sous formes actives, et plusieurs avancées ont été réalisées dans la détermination de leurs substrats. Ces recherches ont également permis la découverte et la caractérisation d'une β-agarase appartenant à une famille de glycoside hydrolase qui n'était pas connue à ce jour pour contenir une telle activité enzymatique. Dans ce contexte, la structure cristalline de cette nouvelle enzyme ainsi que d'une sulfatase associée ont pu être déterminée.
En conclusion, ce travail de thèse a permis de mettre au point différents outils permettant de mettre en évidence de nouvelles activités enzymatiques et aussi de mieux comprendre le catabolisme des agars chez *Z. galactanivorans*.

Agars are red algal polysaccharides. These are composed of D-galactose with L-galactose alterned by glycosidic bond β-1,4 and α-1,3. These galactans harbor several modifications : sulfatations, methylations, pyruvylations… All these modifications hinder the agar degradation by marine bacteria. *Zobellia galactanivorans* is a marine flavobacterium able to degrade marine polysaccharides, including agars. The aims of this thesis are to identify and characterize the enzymatic tools of *Z. galactanivorans* to degrade the complex agar.
The two subjects are a divergent β-agarase (*Zg*AgaC) and sulfatases that act on agars. To perform this study on the agar specific enzymes, we have developed activity screens on a complex agar collection of substrates produced during the thesis. These screens have been used to show the divergent behaviour of *Zg*AgaC on complex agar compared to the other β-agarases and β-porphyranases of the family 16 (GH16). These screens were further used to identify the substrate of two sulfatases active on agar.
To characterize these enzymes, it was necessary to purify the oligosaccharides that are produced by *Zg*AgaC and the oligosaccharides substrate of the sulfatases. Thus, a new method of sulfated oligosaccharides purification has been developed. With this method, we have purified oligosaccharides to be characterized by mass spectrometry. Combined with the structure resolution of *Zg*AgaC, an oligosaccharide pattern of recognition by this new subfamily of β-agarase has been proposed.
Concerning the sulfatases, a protocol of recombinant expression on *E.coli* has been developed to produce the sulfatase on active form. Several advances have been obtained to determine their substrate.
Finally, this research has allowed the discovery and characterization of another β-agarase, belonging to a GH family to date devoid of this specificity. In this context, the 3D structures of this enzyme and of a sulfatase have been resolved.
To conclude, this work has allowed to develop different tools to identify new enzymatic activities and to have a better view of the agar catabolism of *Z. galactanivorans*.