



HAL
open science

Développement d'enzymes recombinants issus des bactéries marines *P. carrageenovora* et SW5 pour la production d'oligo-fucoïdanes et d'oligo- λ -carraghenane

Sébastien Colin

► **To cite this version:**

Sébastien Colin. Développement d'enzymes recombinants issus des bactéries marines *P. carrageenovora* et SW5 pour la production d'oligo-fucoïdanes et d'oligo- λ -carraghenane. Génomique, Transcriptionnelle et Protéomique [q-bio.GN]. Université de Technologie de Compiègne, 2005. Français. NNT : 2005COMP1558 . tel-01115060

HAL Id: tel-01115060

<https://hal.sorbonne-universite.fr/tel-01115060v1>

Submitted on 11 Feb 2015

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

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

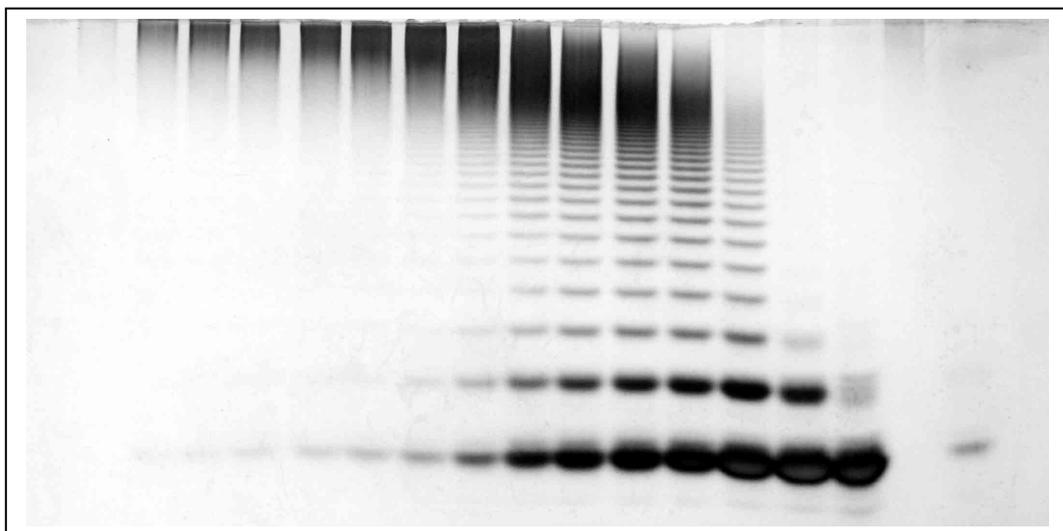


utc
Université de Technologie
Compiègne

Par **COLIN Sébastien**

***DEVELOPPEMENT D'ENZYMES RECOMBINANTS
ISSUS DES BACTERIES MARINES P. carrageenovora ET SW5
POUR LA PRODUCTION
D'OLIGO-FUCOÏDANES ET D'OLIGO- λ -CARRAGHENANES.***

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC.



Soutenue le : 3 mai 2005

Spécialité : **Génie Enzymatique, Bioconversion et
Microbiologie**

D1558

**DEVELOPPEMENT D'ENZYMES RECOMBINANTS ISSUS
DES BACTERIES MARINES *P. carrageenovora* ET SW5
POUR LA PRODUCTION
D'OLIGO-FUCOÏDANES ET D'OLIGO- λ -CARRAGHENANES.**

Soutenu le 3 mai 2005 devant le jury composé de :

**M. D. THOMAS (Président)
M. B. KLOAREG (Directeur de thèse)
M. J.C. YVIN (Co-Directeur)
M. G. BARBIER (Rapporteur)
M. P. BEGUIN (Rapporteur)
Mme B. THOMASSET**

A ma femme Susana, mon fils Vasco, et sa petite sœur Anna

A mes parents, mes frères,

*et à Pierrot Lastennet, mon grand-père, qui, embarqué comme mousse à douze ans, a
abrégé un début de scolarité brillant, à ses valeurs humaines, à sa droiture et son
amour.*



"Démocrite avait trouvé la manière la plus philosophique de jouir de la nature et des hommes; savoir, d'étudier l'une et de rire des autres."

(Jean le Rond d'Alembert, 1717-1783)

REMERCIEMENTS

La thèse concrétise l'achèvement d'un parcours éducatif. En premier lieu, je tiens donc à remercier tout particulièrement l'ensemble des enseignants qui ont jalonné ce parcours, les pédagogues au regard affûté comme les plus ennuyeux. A travers eux, c'est l'ensemble d'un système éducatif qu'il faut reconnaître. Sans eux, sans leurs efforts, pas d'éveil des curiosités, pas de construction de l'esprit critique, bref, le spectre de l'obscurantisme.

Tapis dans l'ombre de ces années d'études, mes parents évidemment. En cette occasion, je souhaiterais leur témoigner toute ma reconnaissance pour l'ensemble de leur œuvre. Moi même père depuis quelques temps, je sens combien la construction d'un enfant n'est pas chose facile. Vue de l'intérieur, j'ai le sentiment que la mienne a été plutôt bien menée. Et je repense alors au temps passé sur mon épaule, ici sur une grammaire rébarbative et là sur un exo de math récalcitrant, aux sacrifices silencieux également. Mon penchant pour l'observation du vivant n'est certainement pas étranger à mes déambulations dans les pas de mon père à la recherche du cèpe furtif, du nid de passereau caché dans les ronces, de cette branche de noisetier pour mon premier arc, ou encore de ces étrilles revêches ; à l'art culinaire de ma mère aussi, qui n'est pas si improvisé qu'elle aime à le prétendre. Je vous remercie donc de m'avoir transmis ce rapport au monde, et de ce soutien discret, mais sans faille, dont vous avez fait preuve. Mais abrégeons cet épanchement forcément réducteur et trop impudique à mon goût. Je remercie mes deux frères Yann et Pierre pour leur présence affective, latente et rassurante, les bons vins de l'un et l'ingéniosité de l'autre.

Il faut maintenant que je vous présente ma femme Susana, portugaise à tempérament, qui, le temps d'une thèse, m'aura donné deux enfants, Vasco deux ans et demi déjà, et sa petite sœur Anna dont la naissance est venue ensoleiller le moment où j'écrivais ces lignes. Elle a su faire preuve de solidité pour deux à des moments où j'étais proche du renoncement. Elle a supporté stoïquement, ou presque, mes doutes, mes sautes d'humeurs, mes adultères chroniques avec ma canne à pêche, et mes belottes nocturnes. Susana, laisse moi t'embrasser publiquement. Si ce travail a finalement abouti c'est aussi grâce à mon fils Vasco, haut comme trois pommes,

espiègle à souhait, sourire dévastateur (attention les filles). Merci mon bonhomme pour tes « xi-coração ». Mes beaux parents, Margarida et Manuel, je leur suis extrêmement reconnaissant de leur gentillesse et de leur délicatesse. Je pense avoir un peu plus de temps pour travailler mon portugais désormais.

Quand j'ai tapé à la porte du bureau de Bernard Kloareg, il y a près de cinq ans maintenant, je souhaitais quitter l'écologie bactérienne où je m'ennuyais passablement, pour retrouver mes premiers amours, à savoir la biochimie des protéines. Quelques mois plus tard, sans trop me connaître, il m'apportait sur un plateau un sujet et une bourse. Je te remercie donc Bernard de m'avoir accordé ta confiance et accueilli dans ton équipe dans de bonnes conditions matérielles. Je suis extrêmement reconnaissant également à Jean-Claude Yvin et Simon Berteau d'avoir accepté de me cofinancer avec l'ANRT. Et à travers eux, c'est la société GOËMAR toute entière que je remercie. J'espère sincèrement que mes travaux leur seront utiles dans leurs futurs développements.

C'est Tristan Barbeyron, en vieux routier (ce n'est pas péjoratif), qui m'a initié, contre quelques clopes, à la biologie moléculaire, à l'analyse de séquence, aux joies et aux déboires de la surexpression (son secret ? c'est LB + 15% de rillettes pur porc). Ce sont des mises à l'étrier qui comptent, merci Barboun. J'ai pris grand plaisir à échanger (surtout recevoir) et bavarder avec Gurvan Michel, expert ès carraghénases, qui a assuré la direction de l'équipe glycobiochimie marine entre le départ de Bernard et l'arrivée de Myriam. Gurvan, je te remercie de m'avoir, bon gré mal gré, co-encadré et épaulé durant le dernier tiers de ce travail. Je suis très reconnaissant à Sabine Génicot de m'avoir aidé à purifier la λ -carraghénase dans une situation pas simple pour elle, et à Myriam Czjzek qui arrive de l'AFMB (Marseille) pour sa chaleur humaine et les quelques discussions que nous avons pu avoir. Je voudrais maintenant remercier tous les complices de paille, Mumu, Tristan, Delphine, Sylvie, Sabine, Estelle, Cécile, Carole et tous les autres que-je-ne-nommerai-pas pour leurs bonnes dispositions au quotidien, leurs boîtes de cônes toujours pleines (maintenant je peux bien avouer que c'est Tristan qui les vide), etc...et le personnel de la station. Je finirai en remerciant Catherine Boyen qui dirige désormais l'UMR 7139, et qui a hérité de ma fin de thèse épineuse. J'ai été sensible à sa diplomatie et à sa patience durant ces longs mois de

rédaction. Et, je tiens enfin à souligner la très large liberté de manip et de rédaction dont j'ai bénéficié, à de rares exceptions près, de la part de l'équipe dirigeante. Je lui suis gré de cette marque de confiance et de cette démarche définitivement formatrice à défaut d'être efficace.

Et puis il y a tous les pôtes, les vieux, les neufs, les filles, les gars , ceux qui ont un bateau et pis ceux qu'en ont pas, ceux qui aiment manger, et ceux qui aiment boire, ceux du boulot, ceux du foot, du winch, du club du point de croix léonard, ceux qui laissent gagner à la pétanque, les loins, les proches, tout ceux-là qui n'ont pas besoin d'être nommés, ceux qui font que l'après boulot est au moins aussi intéressant que le boulot, qui sont là pour faire marrer au bon moment, merci à tous.

Pour conclure, loin de vouloir transformer cette page en une tribune vindicative mais souhaitant franchement me démarquer d'un orgueil intellectuel, ou pire, d'un carriérisme insidieux, je voudrais inviter mes lecteurs à partager l'émerveillement quotidien que me procure l'observation du vivant. Il constitue à mon sens le retour sur investissement le plus valorisant, et une invite à la modestie, à une naïveté de bon aloi, gage d'une certaine fraîcheur d'esprit. Merci à vous lecteurs.

SOMMAIRE

REMERCIEMENTS.....	5
SOMMAIRE	9
CHAPITRE 1	13
1. LA PAROI CELLULAIRE DES MACRO-ALGUES MARINES	16
1.1. LA VARIETE DU COMPARTIMENT PARIETAL A TRAVERS DIFFERENTS REGNES	19
1.1.1. <i>Structure et composition de la paroi procaryotique</i>	20
1.1.2. <i>Structure et composition de la paroi des champignons</i>	23
1.1.3. <i>Structure et composition de la paroi des plantes vasculaires</i>	23
1.2. LA PAROI CELLULAIRE DES MACROALGUES MARINES	24
1.2.1. <i>Particularités des parois d'algues</i>	24
1.2.2. <i>Les polysaccharides squelettiques</i>	28
1.2.3. <i>Les polysaccharides matriciels</i>	28
1.2.3.1. <i>Les algues vertes</i>	31
1.2.3.2. <i>Les algues brunes</i>	31
1.2.3.3. <i>Les algues rouges</i>	32
1.2.4. <i>Importance industrielle et commerciale</i>	35
2. λ-CARRAGHENANES ET FUCOÏDANES, DEUX POLYSACCHARIDES TRES SULFATES.....	39
2.1. PHYSICO-CHIMIE ELEMENTAIRE DES POLYSACCHARIDES	39
2.2. LES POLYSACCHARIDES SULFATES ANIMAUX	45
2.3. LES CARRAGHENANES	47
2.3.1. <i>Classification et nomenclature</i>	48
2.3.2. <i>La famille des λ-carraghénanes</i>	51
2.3.3. <i>Composition des λ-carraghénanes</i>	52
2.3.4. <i>Structures secondaires, tertiaire et quaternaire des carraghénanes : leur impact sur la solubilité et la gélification</i>	54
2.4. LES FUCOÏDANES	59
2.4.1. <i>Nomenclature</i>	60
2.4.2. <i>Structure et composition des fucoïdanes</i>	61
2.5. ACTIVITES BIOLOGIQUES DES POLYSACCHARIDES SULFATES	67
2.5.1. <i>Activités sur les systèmes animaux</i>	67
2.5.1.1. <i>Propriétés anticoagulantes et antithrombiques</i>	68
2.5.1.2. <i>Propriétés diverses</i>	68
2.5.1. <i>Activités sur les systèmes végétaux</i>	70
3. DEGRADATION ENZYMATIQUE DES CARRAGHENANES ET DES FUCOÏDANES.....	73
3.1. CATABOLISME ENZYMATIQUE DES CARRAGHENANES	73
3.1.1. <i>Recherches d'activités carraghénases</i>	73
3.1.2. <i>Les scénarios de dégradation des carraghénanes</i>	76
3.1.2. <i>Les rapports structure/fonction des carraghénases</i>	77
3.1.3. <i>La λ-carraghénase de <i>P. carrageenovora</i></i>	84
3.2. CATABOLISME ENZYMATIQUE DES FUCOÏDANES	84
3.2.1. <i>Recherches d'activités fucoïdanes</i>	86
3.2.2. <i>Les α-L-fucosidases</i>	91
3.2.3. <i>Les sulfatases</i>	92
3.2.4. <i>La α-L-fucoïdanase de SW5, FcnA</i>	93

3.3. LES GLYCOSIDES HYDROLASES.....	97
3.3.1. Deux mécanismes enzymatiques	100
3.3.1.1. La rétention de configuration	100
3.3.1.2. L'inversion de configuration	103
3.3.2. Mode d'action et topologie du site actif.....	103
3.3.2.1. Les modes d'action	103
3.3.2.2. Les topologies de sites actifs.....	104
3.4. INTERET DES ENZYMES EN GLYCOBIOLOGIE	107
3.4.1. Analyse structurale	107
3.4.2. Synthèse d'oligosaccharides.....	111
4. PRESENTATION DU TRAVAIL	115
CHAPITRE 2	117
1. LA PURIFICATION A HOMOGENEITE DE LA λ-CARRAGHENASE DE <i>P. CARRAGEENOVORA</i> 120	
2. UNE METHODE DE DOSAGE POUR LA λ-CARRAGHENASE	125
3. ETUDE DE LA SPECIFICITE DE SUBSTRAT	129
4. CLONAGE ET SEQUENÇAGE DU GENE DE LA λ-CARRAGHENASE.....	130
5. ANALYSE DE LA SEQUENCE	134
5.1. ANALYSE DU CONTIG	134
5.2. ANALYSE DE LA SEQUENCE PROTEIQUE DE LA λ -CARRAGHENASE.....	139
5.2.1. <i>Le peptide signal</i>	139
5.2.2. <i>Recherche des domaines de la λ-carraghénase</i>	140
5.2.3. <i>Conclusions de l'analyse de séquence</i>	148
5. LA SUREXPRESSION DE LA λ-CARRAGHENASE	151
6. CONCLUSIONS GENERALES ET PERSPECTIVES.....	152
6.1. LA λ -CARRAGHENASE, UN OUTIL EN DEVENIR.....	152
6.2. LA λ -CARRAGHENASE, UNE NOUVELLE FAMILLE DE CARRAGHENASES	154
<u>ARTICLE 1: λ-CARRAGEENASE, A NEW GLYCOSIDASE DISTINCT FROM K- AND I-CARRAGEENASES.....</u>	157
CHAPITRE 3	185
1. CLONAGE ET SEQUENÇAGE DU GENE DE LA FUCOÏDANASE (FcnA) DE SW5	189
2. ANALYSE DE SEQUENCE	196
2.1. ANALYSE DU CONTIG	196
2.2. ANALYSE DE LA SEQUENCE PROTEIQUE DE LA FUCOÏDANASE FcnA	201
2.2.1. <i>Trois répétitions contiguës</i>	203
2.2.2. <i>Le domaine C-terminal</i>	208
2.2.3. <i>Le domaine catalytique N-terminal</i>	212
2.3. ANALYSE DES AUTRES SEQUENCES PROTEIQUES DU CONTIG	222
2.3.1. <i>La sulfatase partielle SulB</i>	222
2.3.2. <i>La carboxylestérase EstA</i>	222
2.3.3. <i>La sulfatase entière Sula</i>	226
2.3.4. <i>La fucosidase FcsA</i>	232
2.4. L'HYPOTHESE DU CLUSTER FONCTIONNEL	233
2.4.1. <i>Clusters, opérons et transferts horizontaux</i>	234
2.4.2. <i>Hypothèse des transferts horizontaux</i>	238
2.4.2.1. <i>Des homologues forts entre espèces distantes phylogénétiquement</i>	238

2.4.2.2. <i>Des homologues forts entre espèces proches écologiquement (ou proximité écologique et TLG)</i>	239
2.4.2.3. <i>Répartitions en clusters des sulfatases de R. baltica</i>	240
2.4.3. <i>Conclusions</i>	241
3. SUREXPRESSION DE FcnA ET SulA	242
3.1. SUREXPRESSION DE FcnA1, FcnA2, ET SulA1	242
3.2. OPTIMISATION DE LA PRODUCTION DE FUCOÏDANASE RECOMBINANTE	245
3.3. PURIFICATION DE LA FUCOÏDANASE RECOMBINANTE	251
3.4. ANALYSE DE L'ACTIVITE DE FcnA1 ET 2.....	252
4. CONCLUSIONS GENERALES ET PERSPECTIVES	257
<u>ARTICLE 2: ISOLATION AND CULTURE OF A MARINE BACTERIUM DEGRADING THE SULFATED FUCANS</u> FROM BROWN ALGAE	261
<u>ARTICLE 3: PURIFICATION AND EXPRESSION OF FUCOÏDANASE, A NOVEL GLYCOSIDE HYDROLASE</u>	281
BIBLIOGRAPHIE	309
ANNEXES	329

CHAPITRE 1

INTRODUCTION GÉNÉRALE

Ce travail a pour cadre l'utilisation d'oligosaccharides à des fins phytosanitaires. En effet, dans les relations plante-pathogène la reconnaissance et la virulence des partenaires sont très largement inféodées aux propriétés pariétales des partenaires (Baker et al., 1997 ; Nimchuk et al., 2003 ; Vorweck et al., 2004). Plus précisément, les agressions pathogéniques entraînent généralement la dégradation des parois cellulaires de la plante infestée et éventuellement de l'agresseur. Les oligosaccharides produits lors de cette dégradation permettent alors de déclencher des mécanismes de défense chez l'hôte (Potin et al., 1999 ; Bouarab et al., 1999). Selon le principe du vaccin (Baker et al., 1997), la plante élicitée par un mélange d'oligosaccharides bio-actifs est alors plus résistante à la maladie. Il s'avère que certains oligosaccharides sulfatés issus de la dégradation de polysaccharides extraits de macroalgues marines, sont très actifs de ce point de vue (Klarzynski et al., 2003 ; Menard et al., 2004). Ce travail consistait à produire en vue de leur caractérisation, deux enzymes bactériens hydrolysant d'une part un polysaccharide pariétal d'algue rouge (le λ -carraghénane), et d'autre part un polysaccharide d'algue brune (le fucoïdane).

Dans ce chapitre introductif, j'ai délibérément abordé le développement des notions essentielles à ce travail par les substrats pour finir par les enzymes impliqués dans leur dégradation. Par ailleurs, pour replacer ce travail dans son contexte et détailler les deux systèmes modèles, il m'a paru pertinent de procéder par zooms successifs: règnes, algues, paroi, polysaccharides, polysaccharides sulfatés, les carraghénanes et les fucoïdanes. Ainsi, dans le premier sous-chapitre, je suis parti de l'arbre des différents règnes du vivant pour mettre en exergue les particularités du compartiment pariétal des macro-algues marines. Ces dernières reposent essentiellement sur la forte teneur en polysaccharides sulfatés des macro-algues marines notamment rouges et brunes. Dans le second sous-chapitre, je me suis attaché à décrire en détail la composition, les structures et les propriétés biologiques de deux de ces polysaccharides sulfatés, les carraghénanes et les fucoïdanes. Le troisième sous-chapitre traite de l'enzymologie associée à la dégradation de ces deux polymères ce qui m'amènera naturellement à finir par une description des glycosides hydrolases, groupe enzymatique auquel appartiennent la λ -carraghénase et la fucoïdanase étudiées.

1. La paroi cellulaire des macro-algues marines

Il serait illusoire de vouloir réaliser ici un descriptif exhaustif du compartiment pariétal à travers la diversité végétale tant elle s'est complexifiée au cours des cinquante dernières années. Cette complexité repousse la science du découpage et de la classification du vivant dans ces derniers retranchements, les différences entre protophytes et protozoaires pouvant être par exemple très diffuses. La taxonomie des algues en particulier s'enrichit régulièrement de nouveaux clades en réponse aux techniques d'investigations plus performantes. A tel point que d'un point de vue systématique, le terme « algue » a perdu toute signification puisqu'il désigne des organismes appartenant à différents règnes du vivant (De Reviers, 2003 ; Simpson et Roger, 2004) comme l'indiquent les figures 1 et 2 : Eubacteria, Plantae, Hétérokonta (contenant entre autres, les diatomées, les oomycètes, les algues dorées et jaunes), Alveolata (contenant les Ciliés, Dinoflagellés)...

En guise de préambule, il s'avère donc nécessaire de préciser la terminologie qui sera employée plus avant. Sera considéré comme végétal tout organisme eucaryote photosynthétique possédant des chloroplastes. Les plantes correspondent alors aux Embryophyta (mousses, fougères, conifères, angiospermes...). A l'exception des Cyanobactéries (procaryotes), les différents groupes d'algues, au nombre de onze, seront donc des végétaux (bien que certains groupes aient perdu secondairement la photosynthèse) mais pas des plantes (Fig. 1 et 2). Ils se répartissent en six ou sept grandes lignées évolutives (De Reviers, 2003).

Laissant de côté le phytoplancton, mon propos ne concernera que les macroalgues marines (les « seaweeds » anglosaxonnes). Ces dernières se répartissent en trois grands groupes (Fig. 1 et 2). Les macroalgues vertes appartiennent à l'infra-règne des Chlorophyta (classe des Ulvophyceae). Ce dernier forme avec l'infra-règne des Streptophyta (contenant les Angiospermes) le sous-règne des Chlorobionta (« la lignée verte »). Les algues rouges appartiennent à un sous-règne frère, les Rhodobionta, et plus précisément au phylum des Rhodophyta. Ces deux sous-règnes rassemblent la majeure partie des organismes du règne des Plantae. De ce point de vue, les algues rouges et, plus proches encore, les algues vertes, partagent une origine commune avec

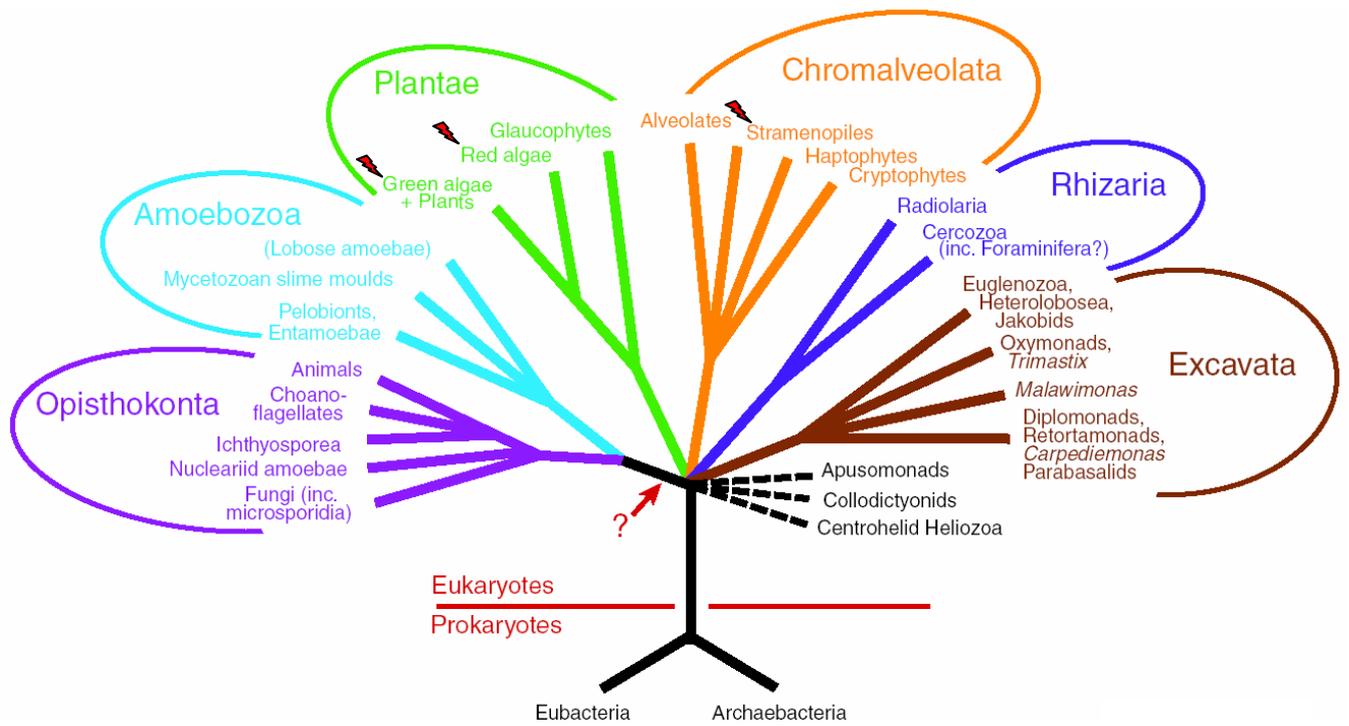
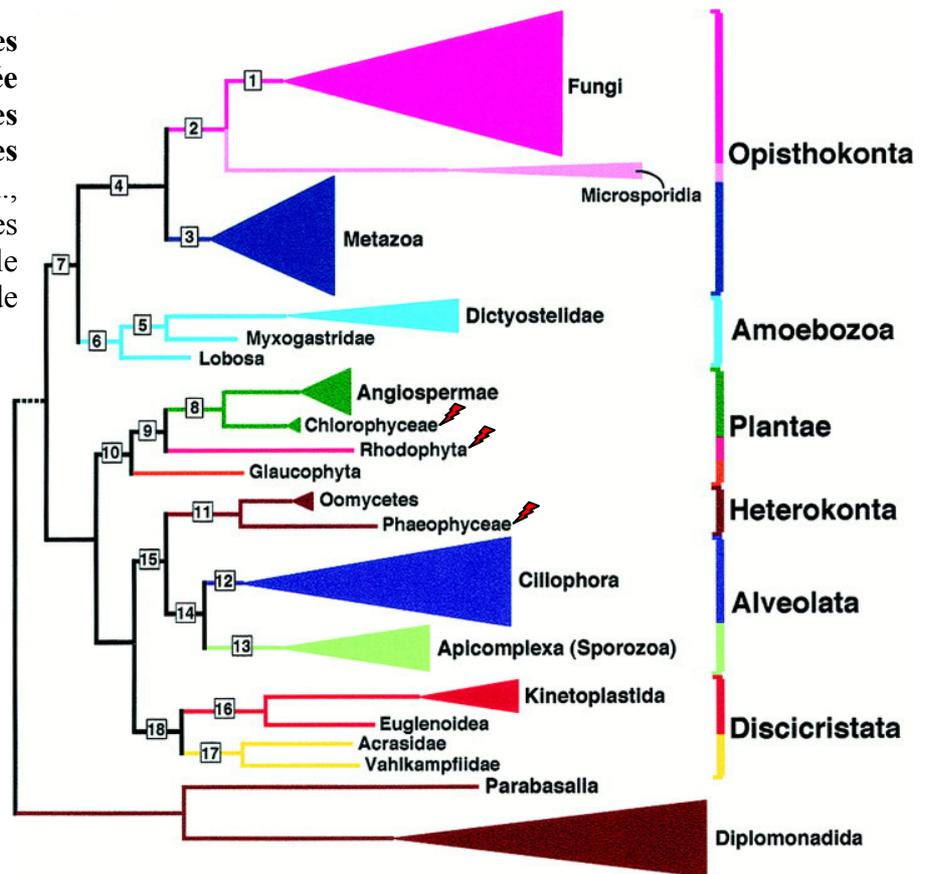


Figure 1 : Arbre schématique décrivant l'organisation des Eucaryotes en 6 règnes (d'après Simpson et Roger, 2004). On notera la divergence précoce des Rhodophyceae et des Stramenopiles.

Figure 2 : Phylogénie des règnes Eucaryotes basée sur la comparaison des séquences de 5 protéines (d'après Baldauf et al., 2000). La largeur des triangles représente le nombre de taxons de l'embranchement.



les plantes terrestres (Baldauf et al., 2000). Il est admis qu'elles résultent toutes d'un même événement d'endosymbiose avec une cyanobactérie, dite endosymbiose primaire (Simpson et Roger, 2004). Les algues brunes (classe des Phaeophyceae) n'appartiennent pas à ce règne et leur lignée est donc très distante phylogénétiquement de celles des algues vertes et rouges (Fig. 1 et 2). Les Phaeophyceae forment une classe faisant partie du vaste sous-règne des Hétérokonta (ou Straménopiles) où figurent également les diatomées. Ces organismes appartiennent au règne des Chromalveolata qui comprend quatre groupes majeurs d'algues eucaryotes dont les dinoflagellés (Simpson et Roger, 2004). Ils résultent d'une endosymbiose dite secondaire puisque impliquant une algue rouge.

1.1. La variété du compartiment pariétal à travers différents règnes

Force est de constater que dans les différents règnes du vivant (Archae, Eubacteria, Plantae, Metazoa, Eumycota, Heterokonta...) si l'isolement du compartiment intracellulaire vis à vis de l'environnement extérieur par le plasmalemme est une constante, le compartiment pariétal a cependant largement divergé, répondant étroitement à la variation des contraintes physiologiques et mécaniques.

Chez les Eucaryotes, les cellules animales se distinguent par l'absence de paroi malgré la présence d'un manteau cellulaire, le glycocalix, qui tapisse la face externe du plasmalemme. Ce revêtement de nature glucidique et glycoprotéique (protéoglycane) joue un rôle actif dans la perception et la transmission des informations externes. Néanmoins cette structure ne partage en aucun cas les propriétés mécaniques des parois végétales, fongiques ou procaryotiques. Dans ces trois derniers groupes la paroi assume un rôle similaire de tampon vis à vis de l'espace extracellulaire. Indépendamment de sa structure, la paroi est ici le premier (et le seul chez les organismes unicellulaires) rempart avec l'environnement extérieur (Braam, 1999). Elle se doit d'être semi-perméable et d'assurer tout le transit nécessaire aux métabolismes de la cellule. Elle fournit une protection osmotique, chimique et mécanique, mais aussi une résistance biologique. Elle participe de surcroît à la

morphogénèse, à la croissance des cellules et à l'agencement des biofilms et elle est enfin le siège des systèmes de perception des variations extérieures. Chez les végétaux et champignons pluricellulaires, elle contribue également aux propriétés tissulaires. Elle y est ainsi sujette à une dynamique de développement. Au cours de la croissance, la différenciation cellulaire entraîne au sein de la paroi un certain nombre de modifications structurales et chimiques qui s'accompagne d'une stratification de cette enveloppe. Enfin, ce compartiment cellulaire joue évidemment un rôle clé dans les relations symbiotiques ou hôtes-pathogènes puisque le contact physique entre deux organismes, et donc leur reconnaissance mutuelle, se fait au niveau de leur surface (Fig. 3).

Il est admis que chez les eucaryotes à paroi, celle-ci se compose d'une phase squelettique très organisée charpentant l'ensemble, et d'une phase amorphe dite matricielle qui enveloppe le squelette et cimente l'ensemble. Quoiqu'il en soit, malgré leur diversité, le point commun de tous les organismes à paroi est la présence, en proportion plus ou moins importantes de polysaccharides structurant ce compartiment cellulaire. Le caractère polyol des monosaccharides permet de vastes possibilités de connexion et de substitutions ce qui constitue une palette extrêmement riche pour la construction de polymères saccharidiques. L'observation faite de la grande diversité des polysaccharides de paroi montre que la nature a su apprivoiser cette richesse au cours de l'évolution, afin d'accommoder au mieux les contraintes environnementales et les propriétés physico-chimiques des polymères pariétaux.

1.1.1. Structure et composition de la paroi procaryotique

La paroi des procaryotes peut se résumer à travers trois grandes catégories. Chez les Eubactéries, d'une part, coexistent deux types de parois permettant de distinguer les souches Gram⁺ où la paroi est de type monocouche, et des Gram⁻, où la structure pariétale est plus complexe (Fig. 4). Elles se différencient par l'importance de la couche de peptidoglycane (beaucoup plus épaisse chez les Gram⁺, 20 à 80 nm, que chez les Gram⁻, 2 à 3 nm) entourant immédiatement la membrane plasmique, et par l'existence d'une seconde membrane (dite externe) chez les Gram⁺ qui enferme la

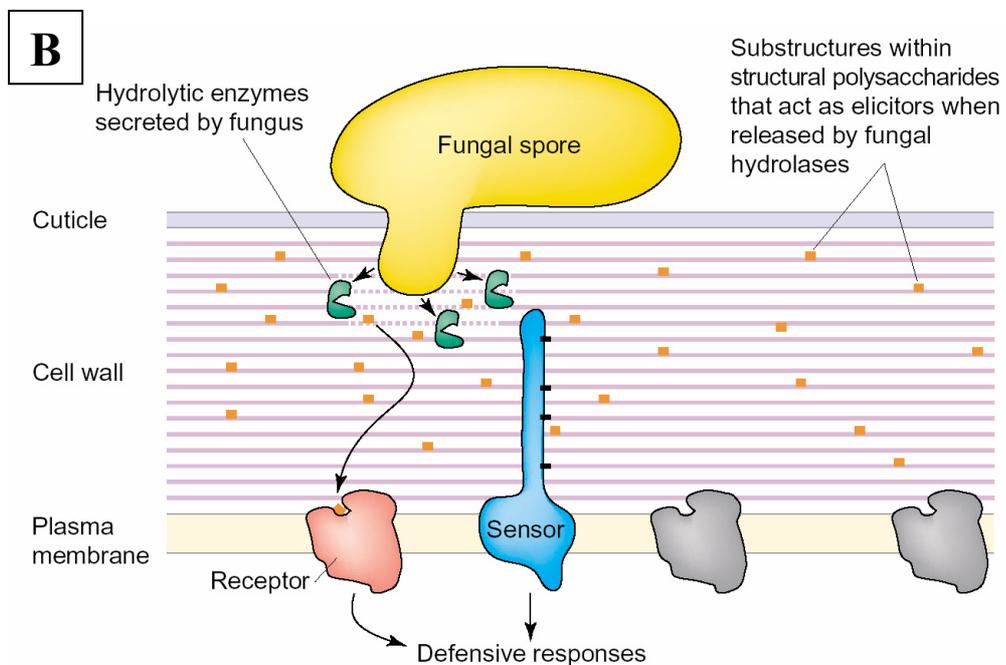
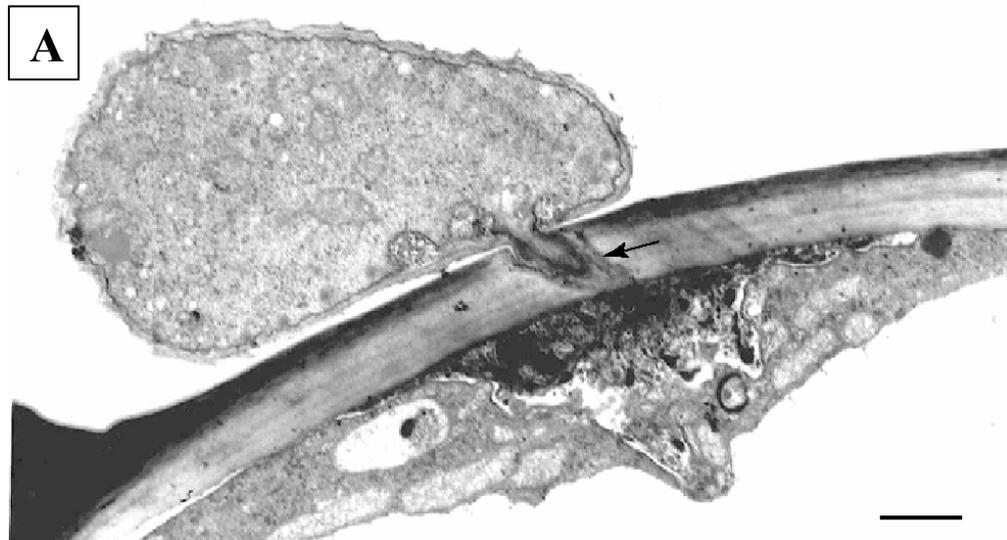


Figure 3 : Modélisation de l'infection de l'orge par un champignon (d'après Vorwerk et al., 2004).

A) Micrographie électronique de *Blumeria graminis* pénétrant la paroi cellulaire de l'orge. On remarquera (flèche) que l'hyphe du champignon est bordé d'une zone où la paroi a été dégradée (échelle 250 nm).

B) Evènements associés à la dégradation de la paroi par le pathogène. La pénétration de ce dernier engendre la libération enzymatique d'oligosaccharides qui diffusent dans la phase liquide de la paroi pour finalement déclencher une réaction de défense via leur reconnaissance par des récepteurs membranaires spécifiques. La présence de senseurs des propriétés mécaniques de la paroi est également suspectée.

couche de peptidoglycanes dans un espace dit périplasmique. Cette seconde membrane en bicouche comporte des phospholipides sur sa face interne et des glycolipides sur la face externe. La troisième et dernière catégorie de parois concerne celle des Archaea qui présente une organisation et une composition différentes. Elle ne contient pas de peptidoglycane mais un ensemble peu étudié de protéines, glycoprotéines et de polysaccharides. Ces trois familles de parois peuvent enfin être enveloppées d'une structure protéique en monocouche cristalline nommée S-layer. Le peptidoglycane est un ensemble de chaînes polysaccharidiques alternant acide N-acétyl-murannique et N-acétyl-glycosamine, pontées par des peptides. Le maillage ainsi formé contribue de façon prépondérante à la rigidité et donc à la résistance mécanique de la paroi.

1.1.2. Structure et composition de la paroi des champignons

Les polysaccharides représentent 80 à 90% de la matière sèche de la paroi des champignons. Les polymères de la phase squelettique sont principalement composés de glucose et de N-acétyl-D-glucosamine (Fig. 5). Il s'agit surtout de chitine cristalline mais aussi de cellulose agencées en microfibrilles, de β -glucanes non cellulosiques, et de α -glucanes. Chitosan, mannanes et divers hétéropolysaccharides amorphes constituent la phase matricielle en association avec des protéines (Farkas, 1979 ; Bartnicki-Garcia, 1968).

1.1.3. Structure et composition de la paroi des plantes vasculaires

La paroi des Trachéophytes (plantes vasculaires) montre une remarquable diversité de composition. Elle est essentiellement constituée de polysaccharides de hauts poids moléculaires dont la nature varie d'un groupe taxonomique à l'autre, ou encore selon les tissus, l'organe considéré, l'âge de l'organisme, illustrant les rapports structures-fonctions de ces polymères. Une excellente revue concernant les parois de dicotylédones vient de paraître (Somerville et al., 2004) qui illustre l'énorme fossé qui existe entre la connaissance de la paroi de ces végétaux dit supérieurs et les autres.

La microscopie photonique a permis d'identifier chez les plantes supérieures (Angiospermes) plusieurs régions distinctes. La lamelle moyenne est la formation la

plus périphérique partagée par les cellules contiguës (Fig. 6). C'est un ciment constitué de polysaccharides acides, les pectines (homogalacturonane, rhamnogalacturonane I et rhamnogalacturonane II). Cette enveloppe hydrophile anionique participe à la cohésion intercellulaire. La région de la paroi plus interne, qui enveloppe individuellement chaque cellule, est biphasique (Preston, 1974). Mise en évidence par microscopie électronique, la phase squelettique cellulosique, cristalline et discontinue, est engoncée dans une phase matriciel, amorphe et continue de composition variable (Fig. 6 ; les hémicelluloses : xyloglucane et glucuronoarabinoxylane). Ce polysaccharide cristallin et insoluble forme chez les végétaux supérieurs des microfibrilles élémentaires de longueur indéfinie dont le diamètre varie de 3 à 4 nm. La cellulose assure ainsi un rôle de charpente au sein de la paroi. Ces dernières sont réticulées par des chaînes polysaccharidiques simples telles que des xyloglucanes (Fig. 6). Il est ainsi suggéré que la cohésion de la paroi repose en large partie sur les liaisons hydrogène que l'hémicellulose partage avec les microfibrilles de celluloses. Chez les cellules jeunes une paroi dite primaire se caractérise par sa forte hydrophilie et sa richesse en hémicellulose. La pauvreté en cellulose procure souplesse et plasticité à cette paroi. En fin de croissance, la paroi se fige par la mise en place de cellulose hautement cristalline associée en fibres très compactes et son taux d'hydratation baisse singulièrement (Preston, 1979). Une étude récente de la paroi cellulaire de la feuille d'*Arabidopsis* constitue un exemple commun de composition (Zabackis et al., 1995) et d'organisation pariétale chez les Angiospermes.

1.2. La paroi cellulaire des macroalgues marines

1.2.1. Particularités des parois d'algues

Les propriétés générales de la paroi sont relativement conservées entre trachéophytes et algues marines. D'un point de vue fonctionnel, ce compartiment cellulaire assume le même rôle de milieu tampon et de charpente. Il participe également au processus de polarisation lors du développement précoce et à l'adhésion au substrat chez les algues (Kloareg et Quatrano, 1988). Comme chez les plantes supé-

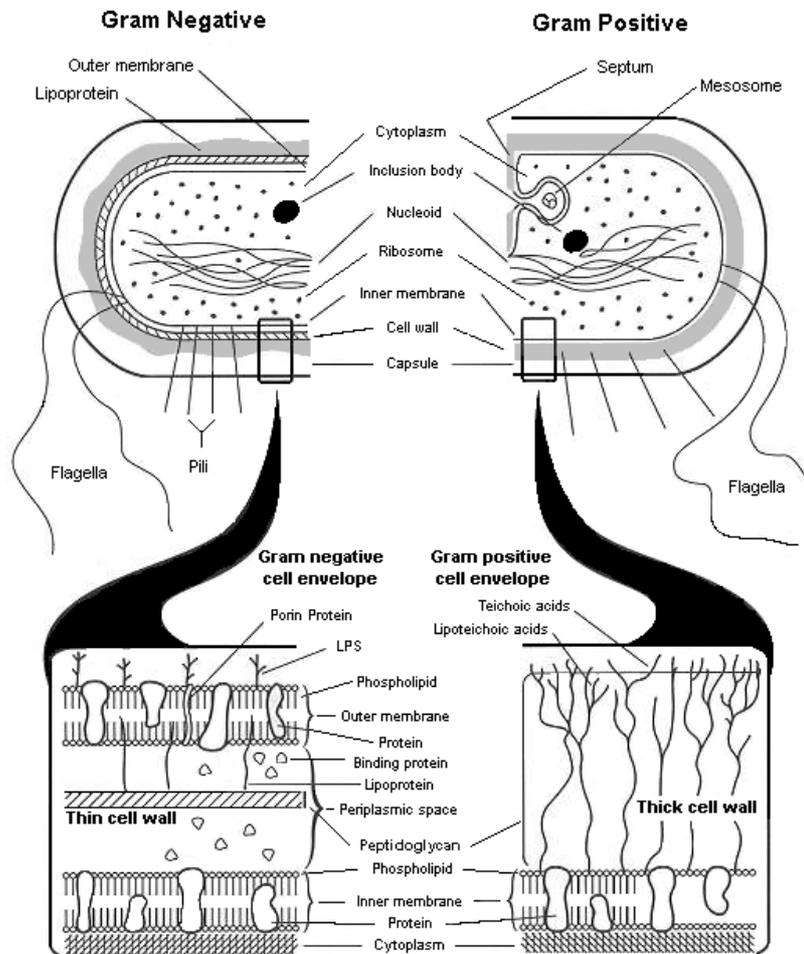


Figure 4 : Schématisation de l'organisation de la paroi des eubactéries Gram- et Gram+

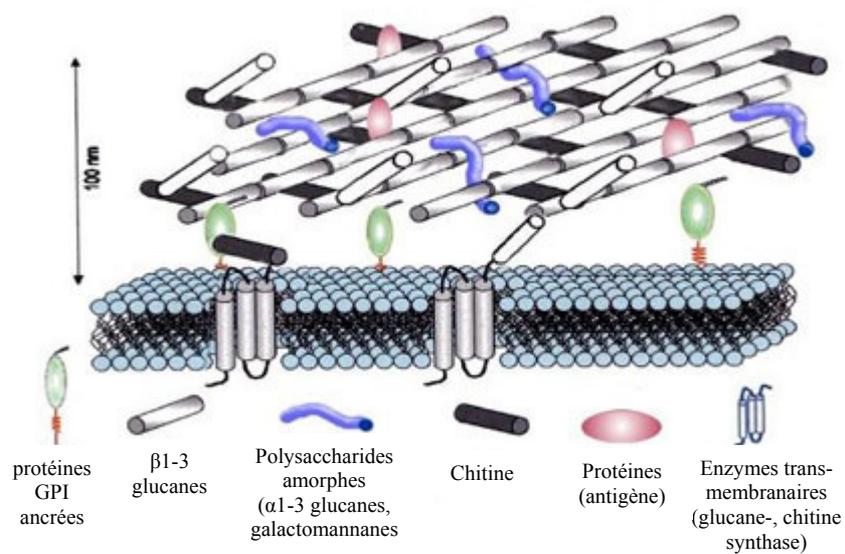


Figure 5 : Schématisation de l'organisation de la paroi cellulaire des champignons

érieures les parois de cellules d'algues sont considérées comme un système biphasique (Rees et al., 1982). Toutefois le squelette cristallin y est beaucoup plus réduit. En contrepartie, la phase matricielle amorphe et polysaccharidique est bien plus importante ce qui confère aux tissus et au végétal des propriétés adaptées à la vie des macroalgues en milieu marin (souplesse, élasticité, résistance à la dissémination). Un panorama détaillé des différents polysaccharides isolés des parois algales a été réalisé par Painter (1983) venant enrichir celui de Percival (1979). Bien que n'étant plus très récent, il présente l'avantage d'être relativement complet, donnant ainsi un bon aperçu de la diversité et des principales caractéristiques des constituants pariétaux des différents groupes d'algues, qui plus est, dans une perspective évolutive.

L'abondance des polysaccharides anioniques (acides ou sulfatés) constitue une particularité des parois algales (Kloareg et Quatrano, 1988), comparée à la paroi des plantes terrestres où les polysaccharides neutres sont prépondérants (exception faite des pectines). Les corrélations observées entre la composition de la paroi et la distribution écologique des espèces suggèrent que les polysaccharides matriciels anioniques participent chez les algues aux régulations mécaniques, hydriques et ioniques (Kloareg et Quatrano, 1988).

La paroi des algues multicellulaire est classiquement divisée en deux régions distinctes : une succession de couches organisées entourant le plasmalemme de chaque cellule (la paroi immédiate), et un matériel moins organisé qui remplit l'espace intercellulaire (la matrice intercellulaire). Cette distinction repose essentiellement sur une différence dans la répartition des polysaccharides squelettiques et matriciels. Ainsi chez *Gelidium pacificum*, les couches internes de la paroi sont essentiellement composées de celluloses, alors que les couches externes contiennent de l'agarose, et la matrice intercellulaire de l'agaropectine (Akatsuka et Iwamoto, 1979). Chez les Rhodophyceae carraghénophytes, une stratification similaire a été observée, les carraghénanes remplaçant l'agar (La Claire et Dawes, 1976 ; Gordon-Mills et McCandless, 1975). En outre ces études ont également montré l'existence d'un gradient croissant de sulfate dans les galactanes des couches internes vers les couches externes, mais également des cellules profondes vers les cellules périphériques.

1.2.2. Les polysaccharides squelettiques

Chez la plupart des algues multicellulaires, les polymères squelettiques sont des polysaccharides neutres et linéaires. Comme chez les plantes supérieures, le plus commun est la cellulose (Kreger, 1962). Cependant la cellulose extraite des parois algales est rarement un polymère pur de β -1,4-glucose (Turvey et Williams, 1970). Les thalles de Pheophyceae (algues brunes) et des Rhodophytes (algues rouges) contiennent 1 à 8 % de cellulose (Ross, 1953) alors que ce taux est d'environ 30 % chez les végétaux supérieurs (Preston, 1974). Les Siphonocladales, les Cladophorales, et les Ulvales, trois ordres appartenant à la classe des Chlorophytes (algues vertes), représentent néanmoins une exception remarquable. Leur paroi contient en effet 70 % de cellulose extrêmement cristalline (Frei et Preston, 1961). A l'opposé la cellulose est absente de la paroi de certaines algues alors qu'elle est systématiquement observée chez les végétaux supérieurs. Chez les Caulerpales et les Dichomosiphales (Chlorophytes) elle est substituée par du β -1,3-xylane hautement cristallin (Frei et Preston, 1961). La paroi des Codiales et des Dasycladales (Chlorophytes) est, de façon surprenante, composée quasi exclusivement de mannose avec une phase squelettique de β -1,4-mannane cristallin (Mackie et Preston, 1974) associé à du mannane amorphe. Les mannane et xylane sont également décrits chez les Rhodophytes comme polysaccharides structurants (Painter, 1983).

1.2.3. Les polysaccharides matriciels

Les polysaccharides matriciels sont isolés plus aisément que les composants de la phase squelettique. En effet, dans la majorité des cas, ils sont solubles dans l'eau en présence de sels, d'acides ou de bases, et peuvent donc être extraits par trempage d'extraits bruts d'algues à haute ou basse température. Quelle que soit la classe considérée, la matrice résulte d'un mélange de plusieurs types de polysaccharides. Ces derniers présentent par ailleurs une large variabilité qualitative et quantitative fonction de l'espèce, du tissu, de l'âge de l'organisme, de l'habitat et de l'époque de récolte (Craigie, 1990). Des différences quantitatives notoires ont également été observées au sein d'une même espèce entre le sporophyte et le gamétophyte.

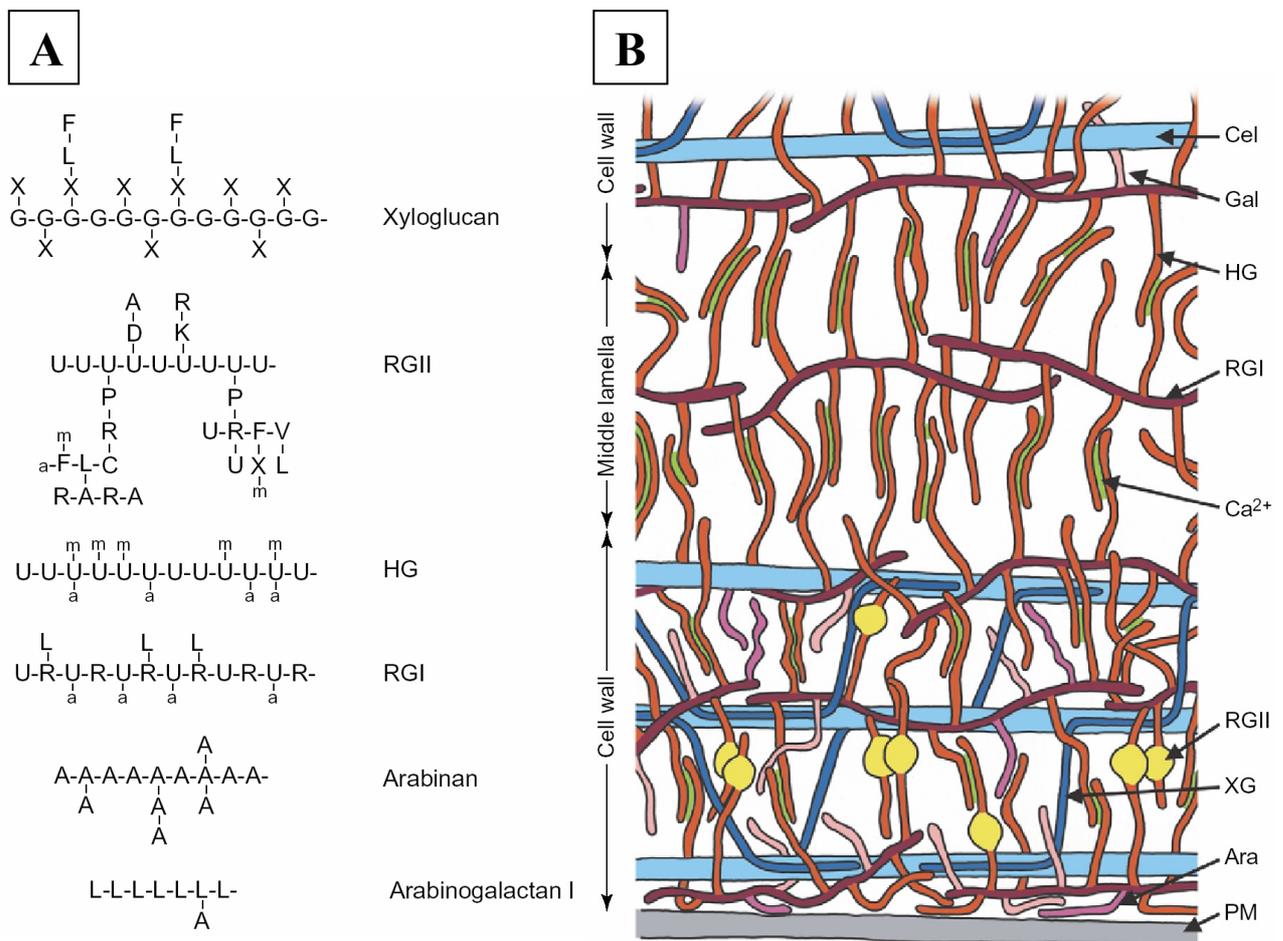
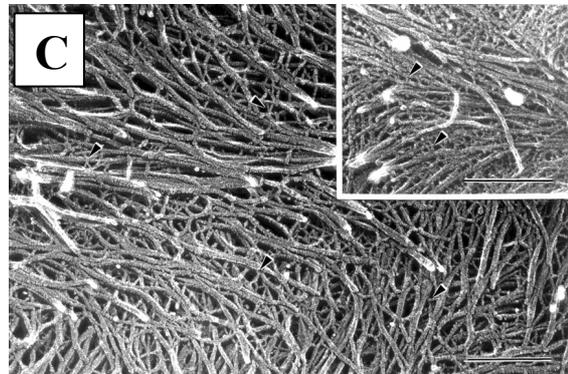


Figure 6 : La paroi cellulaire des Angiospermes (d'après Vorwerk et al., 2004 et Somerville et al., 2004).

A) Composants polysaccharidiques principaux de la paroi. Symboles : a, acétate; A, L-arabinose; C, acide L-acérique; D, acide D-3-déoxy-D-lyxo-2-heptulopyranosylarique; F, L-fucose; G, D-glucose; K, acide D-3-deoxy-D-manno-octulosonique; L, D-galactose; m, méthyle; P, D-apiose; R, L-rhamnose; U, acide D-galacturonique; V, acide D-glucuronique; X, D-xylose).

B) Modèle d'organisation des polysaccharides dans la paroi. Les microfibrilles de cellulose (Cel) sont pontées par l'hémicellulose, principalement le xylo-glucane (XG), lequel est lié par liaisons hydrogène à la cellulose. Le rhamnogalacturonane I (RGI) est également aligné avec les microfibrilles de cellulose. L'homogalacturonane (HG), les arabinanes (Ara), les galactanes (Gal) et le rhamnogalacturonane II (RGII) apparaissent comme étant des branches latérales attachées à de longues chaînes de RGI. Les molécules de RGII pontent différentes molécules de pectines via des esters de borate. En outre, le calcium (Ca^{2+}) permet l'association des régions acides de molécules de HG non-estérifiées. PM indique la membrane plasmique.

C) Réseau de cellulose de la paroi cellulaire. Micrographie électronique d'une cellule épidermique de *Pisum sativum* traitées à l'EDTA. Les microfibrilles de cellulose et leurs pontages sont indiqués par les flèches (échelle 200 nm). L'insert montre la même paroi non traitée.



1.2.3.1. Les algues vertes

Les chlorophytes présentent un mucilage polysaccharidique hybride rappelant les hémicelluloses et pectines des plantes terrestres. Ce sont majoritairement des hétéropolysaccharides polydisperses principalement composés de xylose, arabinose, et galactose (parfois rhamnose et mannose) probablement liés à des protéines en proportion variable. Ils sont sulfatés et hautement ramifiés (Percival, 1979). Ils présentent des propriétés gélifiantes. La phase matricielle des Cladophorales, et des Siphonales est constituée de arabinoxylolactanes sulfatés, majoritairement formés de blocs de L-arabinoses liés en α (1, 4), séparés par des résidus D-galactose, et branchés par des D-xyloses (Percival and Young, 1971). Par contre les mucilages des Ulvales et des Acrosiphoniales sont des glucuronoxylorhamnanes aussi bien sulfatés que substitués par des acides carboxyliques. Ils sont le plus fréquemment formés par la répétition de L-rhamnose-2-sulfate et d'acide D-glucuronique liés alternativement en α (1, 4) et β (1, 3), ramifiés par des D-xyloses liés en β (1, 4) (Medcalf et al., 1972 ; Percival, 1979).

1.2.3.2. Les algues brunes

Le polymère prépondérant de la matrice des Phaeophyceae est l'alginate (ou acide algique) qui peut représenter jusqu'à 40% de la matière sèche (Painter, 1983 ; Larsen et al., 2003). C'est également le plus connu du fait de sa très grande importance commerciale (Draget et al., 1997, 2003). Cet hétéropolysaccharide est constitué de deux acides uroniques, le β (1, 4)-mannuronate et son C₅ épimère le α (1, 4)-guluronate (Painter, 1983). Par hydrolyse acide ménagée, ainsi que par dégradation enzymatique (alginate lyases), il a été démontré que l'alginate est un polymère composé de blocs linéaires de poly- β -D-mannuronate (blocs MM), de poly- α -D-guluronate (bloc GG) et de mélange des deux acides (blocs MG) quasiment en alternance (Haug et al., 1974). Ainsi la structure primaire des alginates définit non seulement la proportion relative en acides mannuronique et guluronique, mais aussi la répartition et les rapports dans la chaîne des trois types de blocs. Le ratio M/G varie selon la nature et l'âge des tissus (Haug et al., 1974). A l'état natif, cette famille de polymère est un sel insoluble des principaux cations présents dans l'eau de mer (à savoir Na⁺, Mg²⁺ et Ca²⁺). Le calcium

forme avec les blocs G des structures type « egg-box » où il est coordonné avec les oxygènes du polysaccharide. De telles structures impliquent plusieurs chaînes s'empilant les unes sur les autres de manière à réaliser de véritables jonctions (Fig. 7). L'importance en taille et en nombre de ces dernières définit la stabilité du réseau connectant les chaînes entre elles, et donc la force du gel ainsi formé (Draget et al., 1997). La solubilisation de l'alginate passe par l'échange du Ca^{2+} par un cation monovalent.

Les algues brunes synthétisent en quantité moindre une autre famille de d'hétéropolysaccharides très complexes, rassemblés sous l'appellation fucoïdane. Leur particularité repose sur des teneurs généralement élevées en fucoses (plus ou moins sulfatés) mêlées à des résidus galactose, mannose, xylose et d'acides uroniques. Ils feront l'objet d'une description détaillée dans la partie § 2.4. Plusieurs polysaccharides cohabitent sous cette appellation (Kloareg et al., 1986). Les sargassanes (ou glycuronofucogalactanes) tout d'abord où les chaînes de $\alpha(1,4)$ -D-galactose sont prépondérantes, substituées par des résidus L-fucose sulfatés et des acides uroniques (principalement du D-glucuronate). Les ascophyllanes (ou xylofucoglycuronanes) très hétérogènes, sont eux constitués d'un squelette d'acide D-mannuronique lié en $\beta(1,4)$ et ramifiés par des fucose et des xyloses. Les homofucanes, enfin, comportent majoritairement du L-fucose sulfaté, différemment liés et plus ou moins branchés. Les connaissances encore parcellaires de ces composés pariétaux ont néanmoins donné lieu à un modèle d'organisation (Fig. 7).

1.2.3.3. Les algues rouges

L'excellente revue de Craigie (1990) bien que commençant à vieillir, analyse de façon exhaustive les différentes données à disposition concernant la paroi cellulaire des algues rouges. Les Rhodophyceae synthétisent principalement des galactanes linéaires et sulfatés de haut poids moléculaire ($> 10^6$ Da) (Tong et al., 1980). Ce sont les composants majoritaires dans la paroi de la grande majorité des algues rouges (Percival, 1979 ; McCandless et Craigie, 1979 ; Painter, 1983 ; Falshaw et al., 1999). Ce sont des polymères de galactoses liés alternativement en $\beta(1,4)$ et $\alpha(1,3)$. Cette structure régulière est souvent masquée par la présence d'ester-sulfates, de groupements

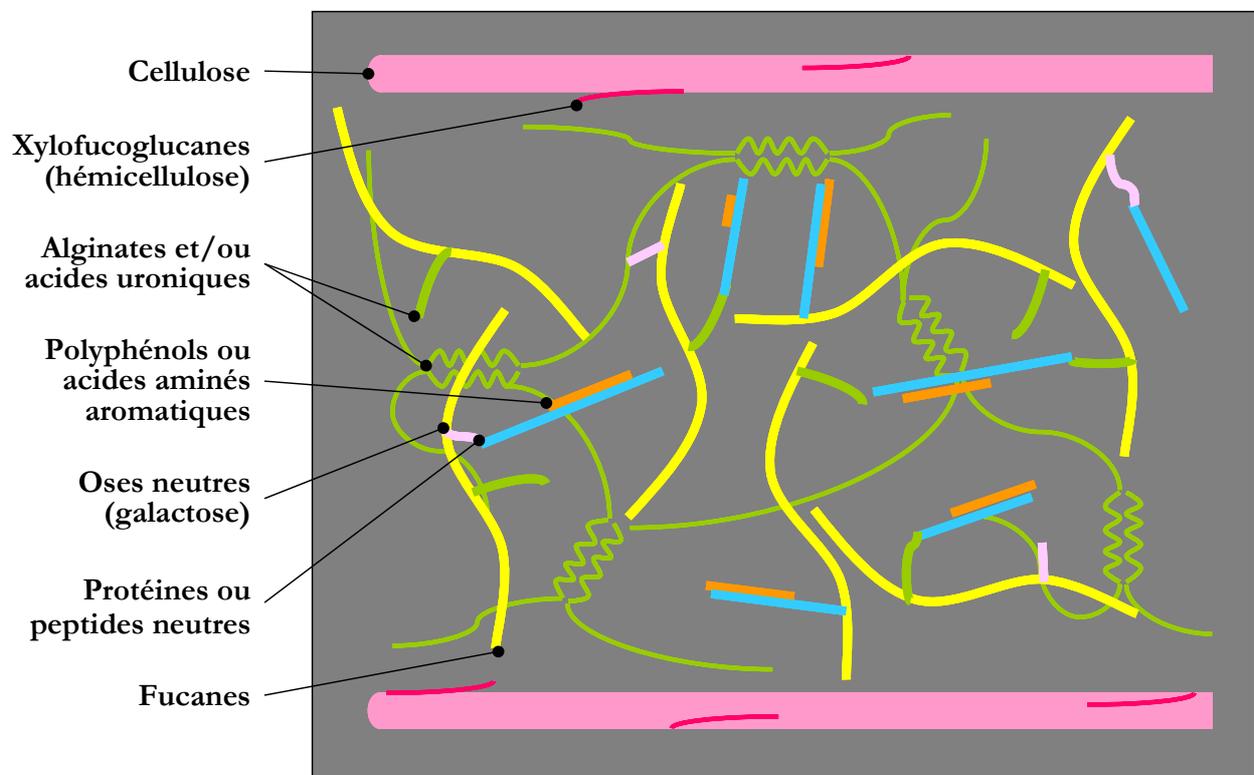


Figure 7 : Modélisation de l'organisation de la paroi des Fuciales (d'après des études d'Estelle Deniaud sur *Pelvetia canaliculata*, *Fucus sp.* et *Ascophyllum nodosum*, communication personnelle). Les alginates s'auto-associent via des ponts calcium dans des structures type « egg-box ». Les fucoïdanes, des oses neutres, des chaînes d'acides uroniques et des chaînes peptidiques pourraient être liés covalamment. Des polyphénols seraient associés avec la fraction peptidique via des interactions électrostatiques.

méthyles, ou de pyruvate. On distingue deux grandes familles de galactanes : les carraghénanes et les agars. Alors que le β -galactopyranosyl lié en 3 est en configuration D dans les deux familles, le α -galactopyranosyl lié en 4 est en configuration D dans les carraghénanes et en configuration L dans les agars. Je détaillerai plus avant ces polysaccharides sulfatés dans la section § 2.3. Néanmoins il faut rappeler que le α -galactopyranosyl lié en 4 est très souvent un 3,6-anhydrogalactose en conformation 1C_4 . C'est notamment le cas des κ - et ι -carraghénanes ainsi que de l'agarose. Ce dernier est considéré comme un polysaccharide neutre (bien que pouvant être légèrement sulfaté), par opposition aux carraghénanes canoniques qui possèdent de 1 à 3 groupement sulfates par disaccharides, respectivement pour le κ - et ι - et λ -carraghénane.

Chez les Rhodophyceae, il est rapporté que, outre les facteurs saisonniers et environnementaux, la ploïdie influe la nature des galactanes répartis au sein de la paroi (Craigie, 1990). La plupart des algues rouges adoptent en effet un cycle de vie diplobiontique alternant une génération diploïde, le sporophyte, et une génération haploïde, le gamétophyte. Le gamétophyte (mâle ou femelle) contient un mélange de κ/ι -carraghénanes (famille κ), alors que les carraghénanes de la famille λ caractérisent le tétrasporophyte (McCandless et al., 1983). Des observations du même ordre sont disponibles pour les agarophytes (Craigie, 1990). Toutefois dans l'ordre des Gigartinales, les familles des Solieriaceae, des Furcellariaceae et des Hypneaceae semblent synthétiser les mêmes carraghénanes indépendamment de l'alternance des générations (Matulewicz et al., 1989).

1.2.4. Importance industrielle et commerciale

Les parois cellulaires d'algues constituent une ressource fondamentale pour la production d'hydrocolloïdes à usage industriel. Leur usage remonte à 1658 quand les propriétés gélifiantes de l'agar extrait des algues rouges furent découvertes au Japon. Toutefois, les carraghénanes doivent leur nom à un petit village irlandais « Carraghen » où l'« Irish Moss » (*Chondrus crispus*, algue rouge) était utilisée depuis plus de 500 ans pour la confection de flans à base de lait. Et les extraits

d' « Irish Moss » contenant du carraghénane furent populairement employés en Europe au 19^{ème} siècle comme épaississant alimentaire. Mais il fallu attendre les années 1930 pour voir émerger la commercialisation d'alginate, extrait d'algues brunes, comme épaississant et gélifiant. Les applications industrielles des colloïdes algues se sont particulièrement enrichies après la seconde guerre mondiale, bien qu'étant parfois limitée par les problèmes d'approvisionnement. A l'heure actuelle, une forte proportion du matériel brut d'où sont extraits les polysaccharides pariétaux provient de l'aquaculture.

Cet aspect de l'exploitation des macroalgues est particulièrement bien développé d'un point de vue industriel et commercial dans une récente revue de la Food and Agriculture Organisation (FAO) des Nations Unies (McHugh, 2003). Aujourd'hui, approximativement 1 million de t/an d'algues sont récoltées afin de produire trois familles principales d'hydrocolloïdes. Ces derniers représentent 55 000 t/an pour une valeur commerciale de 585 millions US\$. La production d'alginate (213 millions US\$) provient exclusivement d'algues brunes alginophytes issues des ressources naturelles, principalement *Ascophyllum* (16%), *Durvillaea* (3,5%), *Ecklonia* (2,5%), *Laminaria* (34%), *Lessonia* (16%) et *Macrocystis* (28%), leur culture engendrant un coût de revient trop important. A l'inverse la production d'agar (132 millions US\$) est principalement issue de la culture de deux algues rouges agarophytes (*Gelidium* et *Gracilaria*). Cultivée depuis les années 60-70, l'exploitation de *Gracilaria* s'est intensifiée au cours des années 90 bien que fournissant un agarose de qualité moindre. Cette indépendance vis à vis de la ressource naturelle a permis une nette expansion de l'industrie de l'agar. Toutefois cette dernière exploite encore largement les ressources naturelles dans une proportion variant au gré des fluctuations du marché. Provenant également d'algues rouges, la production de carraghénanes (240 millions US\$) était à l'origine dépendante d'algues rouges carraghénophytes sauvages (spécialement *Chondrus crispus*) dont la ressource était limitée. Depuis les années 70, l'essor de son industrie est dû à la disponibilité de ressources cultivées en eau chaude dans des pays à bas coûts salariaux. C'est au Philippines que fut initiée la culture des deux principaux carraghénophytes actuels (89% des carraghénophytes récoltés) *Euचेuma cottonii* (ou *Kappaphycus alvarezii*) et *Euचेuma spinosum* (ou *Euचेuma*

denticulatum) qui présentent l'avantage de produire chacune un type prédominant de carraghénane, respectivement le kappa- et le iota-carraghénane. Le genre *Gigartina* représente 9% des récoltes, alors que *Chondrus* représente seulement 2%.

Que ce soit dans l'industrie agroalimentaire, cosmétique ou autres, ce sont les propriétés texturantes (gélifiant, épaississant, stabilisant) de ces hydrocolloïdes qui sont exploitées. L'agar se compose d'agarose (le polysaccharide responsable des propriétés gélifiantes) et d'agaropectine. Il est dédié à hauteur de 90% à l'agroalimentaire. Les 10% restant concernent son emploi en microbiologie et en biotechnologie (agarose purifié). Une utilisation pharmacologique mineure comme laxatif est à rappeler néanmoins. L'intérêt que l'agroalimentaire voue à l'agar repose sur sa résistance à la cuisson, sa gélification simple par refroidissement, l'augmentation de ses pouvoirs gélifiants en présence de sucrose, son absence de couleur et de goût, et son apport calorifique quasi nul (contrairement à l'amidon).

L'utilisation de l'alginate se base sur trois grandes propriétés. La première concerne sa capacité à augmenter la viscosité d'une solution aqueuse, lui attribuant une fonction d'épaississant. La seconde est sa capacité à former des gels en présence d'ion calcium à température ambiante. Contrairement à ceux obtenus avec l'agar, les gels d'alginate sont thermorésistants. La troisième propriété de ce polysaccharide concerne sa capacité à faire des films (en présence de sodium ou de calcium), et des fibres (en présence de calcium). Ces propriétés sont étroitement liées au ratio M/G du polymère et donc à la source de ce dernier. L'impression des textiles représente 50% du marché global de l'alginate, où il sert d'épaississant aux pâtes d'impression contenant les colorants. Ses propriétés épaississantes sont également exploitées dans l'agroalimentaire où il permet notamment de stabiliser des émulsions ou de retenir l'eau de certaine préparation. Il est aussi utilisé comme gélifiant dans la réalisation de substituts de fruits, la restructuration de produits carné, les desserts type « jelly ». Les films d'alginate sont utilisés pour éviter l'adhésion des aliments à leur emballage et pour préserver des produits d'effets néfastes de la congélation, de l'oxydation par l'air ou de risques microbiologiques. D'autres applications de l'alginate ont été développées dans l'industrie chimique de synthèse/conversion (immobilisation d'enzymes voire de cellules en gel), l'industrie pharmaceutique (utilisation sous forme de fibres dans les

pansements cicatrisants, dans l'alimentation diététique comme agent « remplissant », comme gel de moulage), dans l'industrie du papier (agent de surface). La pharmacie et l'agroalimentaire (20%) représentent 33% des débouchés de l'alginate.

Trois type de carraghénanes sont exploités : le κ -carraghénane qui permet l'obtention de gels forts et rigides en présence d'ions potassium ou très cassant avec des ions calcium ; le ι -carraghénane qui, avec des ions calcium (dans une moindre mesure avec les ions potassium), donne des gels élastiques résistant aux congélation/décongélation ; et enfin le λ -carraghénane qui bien que non gélifiant, forme des solutions très visqueuses. Une des particularités de cette famille de polysaccharides est leur fort caractère polyanionique. Ce dernier autorise des interactions avec les protéines chargées positivement en particuliers la caséine. Il en découle un usage important des carraghénanes dans les produits laitiers manufacturés où leur effet est perceptible à de très basse concentration (de l'ordre de 0,1% p/v, voire 0,01%). Les carraghénanes sont utilisés pures ou en mélange afin d'obtenir un ajustement des propriétés rhéologiques. Ils trouvent leurs principales applications dans l'industrie agroalimentaire des produits laitiers (33% du marché), des aliments liquides (15%), des produits allégés. Dans les produits carnés (15%) ils servent de stabilisateurs/rétenteurs de fluides. Ces polysaccharides sont en outre très employés dans la nourriture en conserve pour animaux (25%). En dehors de l'agroalimentaire, on les retrouve dans les dentifrices (6%) ou d'autres applications plus confidentielles (6%, immobilisation de biocatalyseurs, texturants cosmétiques).

Les fucoïdanes rentrent pour une part infime dans l'industrie des hydrocolloïdes. Leur principale exploitation se fait en Asie où ils entrent dans la composition de produits « nutraceutiques » en tant que composés bioactifs et non plus texturants.

2. λ -carraghénanes et fucoïdanes, deux polysaccharides très sulfatés.

Les polysaccharides sulfatés forment un groupe complexe de macromolécules affichant un large spectre de propriétés biologiques importantes. Ces polymères anioniques sont largement répandus dans la nature et apparaissent dans une grande variété d'organismes. Dans le règne animal, les polysaccharides sulfatés principaux sont les glycosaminoglycans (GAGs) sulfatés qui sont des composants ubiquistes des tissus connectifs (Falshaw et al., 1999) des vertébrés. Des fucanes et des galactanes ont néanmoins été décrits chez des échinodermes (oursins et holothuries, cf. ci-dessous) et des tunicés (ascidies ; Santos et al., 1992). Chez les végétaux, comme décrit précédemment, ils ont été détectés de façon exclusive dans la paroi des algues (principalement marines), mais jamais identifiés dans les plantes vasculaires. Toutefois une étude récente (Aquino et al., 2005) vient d'identifier du D-galactane sulfaté ($[\rightarrow 3\text{-}\beta\text{-D-Gal-2-sulfate-1}\rightarrow 4\text{-}\alpha\text{-D-Gal-1}\rightarrow 4\text{-}\alpha\text{-D-Gal-1}\rightarrow 3\text{-}\beta\text{-D-Gal-4-sulfate-1}\rightarrow]_n$) dans les racines et les rhizomes de l'angiosperme marine (« seagrass ») *Ruppia maritima* (Potamogetonaceae). Ces auteurs avancent l'hypothèse d'une évolution/adaptation convergente pour expliquer la présence de galactanes sulfatés dans des organismes partageant le même environnement marin, mais très éloignés phylogénétiquement (Fig 8).

2.1. Physico-chimie élémentaire des polysaccharides

Les polysaccharides forment la famille de biopolymères la plus abondante et la plus diverse. Avec plusieurs centaines d'exemples connus, ils offrent une riche diversité de structures chimiques, allant du simple homopolymère linéaire aux hétéropolymères branchés. Ce sont les macromolécules pouvant former le plus grand nombre de combinaisons structurales. En effet, elles peuvent être constituées de différents types d'oses lesquels adoptent deux formes énantiomériques possibles D et L. Les oses les plus fréquemment rencontrés sont les pentoses (trois centres asymétriques, soit 2^3 stéréoisomères) et les hexoses (quatre centres asymétriques soit 2^4 stéréoisomères). Ces résidus monosaccharidiques composant ces polymères ont des

structures cycliques de type pyranose ou furanose. Les résidus composant les fucoïdanes et les carraghénanes sont des pyranoses (des furanoses sont faiblement observés dans le fucoïdane) qui adoptent une conformation de type chaise conformément à la majorité des polysaccharides composés de pyranoses (Perez et Kouwijzer, 1999). Ils peuvent de surcroît être associés par différents type de liaisons glycosidiques (1→2, 1→3, 1→4, et 1→6) selon les deux anoméries possible α et β (amidon, $\alpha(1,4)$ - D-glucose, et cellulose, $\beta(1,4)$ - D-glucose). Ainsi on considère qu'un hexasaccharide peut en théorie former $1,05 \cdot 10^{12}$ structures différentes (Imberty et Perez, 2000). Les plus grandes unités répétitives caractérisées sont des octasaccharides. Les polysaccharides les plus simples, dont la structure répétée est composé d'un monosaccharide unique, sont généralement dédiés au stockage de l'énergie (amidon, glycogène). Les polysaccharides pariétaux affichent communément une complexité structurale plus forte qui les rends moins accessibles à l'investigation. C'est le cas des carraghénanes, et des fucoïdanes de manière encore plus accentuée.

Contrairement aux protéines globulaires la longueur d'une chaîne polysaccharidique n'est pas fixe, mais peut au contraire largement varier d'une molécule à l'autre. De surcroît, l'enchaînement peut être linéaire mais également ramifié. Une population de polysaccharide sera donc décrite par une distribution des poids moléculaires, mono- ou poly-disperse permettant de calculer un poids moléculaire moyen. Une extrémité de la chaîne polysaccharidique possède le caractère réducteur du monosaccharide, alors que l'autre extrémité est dite non réductrice, ce qui introduit la notion de polarité chez ces molécules.

En première approximation un polysaccharide peut être considéré comme une succession d'unités rigides reliées par une liaison glycosidique flexible (Fig. 9). Lorsque cette liaison intervient entre deux carbones de cycles pyranoses voisins (tous les cas sauf lors d'une liaison C1'→C6), la liaison O-glycosidique (ex. C1'→C4) définit deux angles dièdres φ (H1'-C1'-O1'-C4) et ψ (ex. C1'-O1'-C4-H4). Contrairement à la liaison peptidique des protéines, la rotation au tour de ces deux angles est relativement libre, occasionnant de faibles variations de l'énergie interne de la molécule. Des calculs d'énergie conformationnelle ont néanmoins montré que pour un disaccharide non substitué, 90% des couples (φ , ψ) sont énergétiquement défavora-

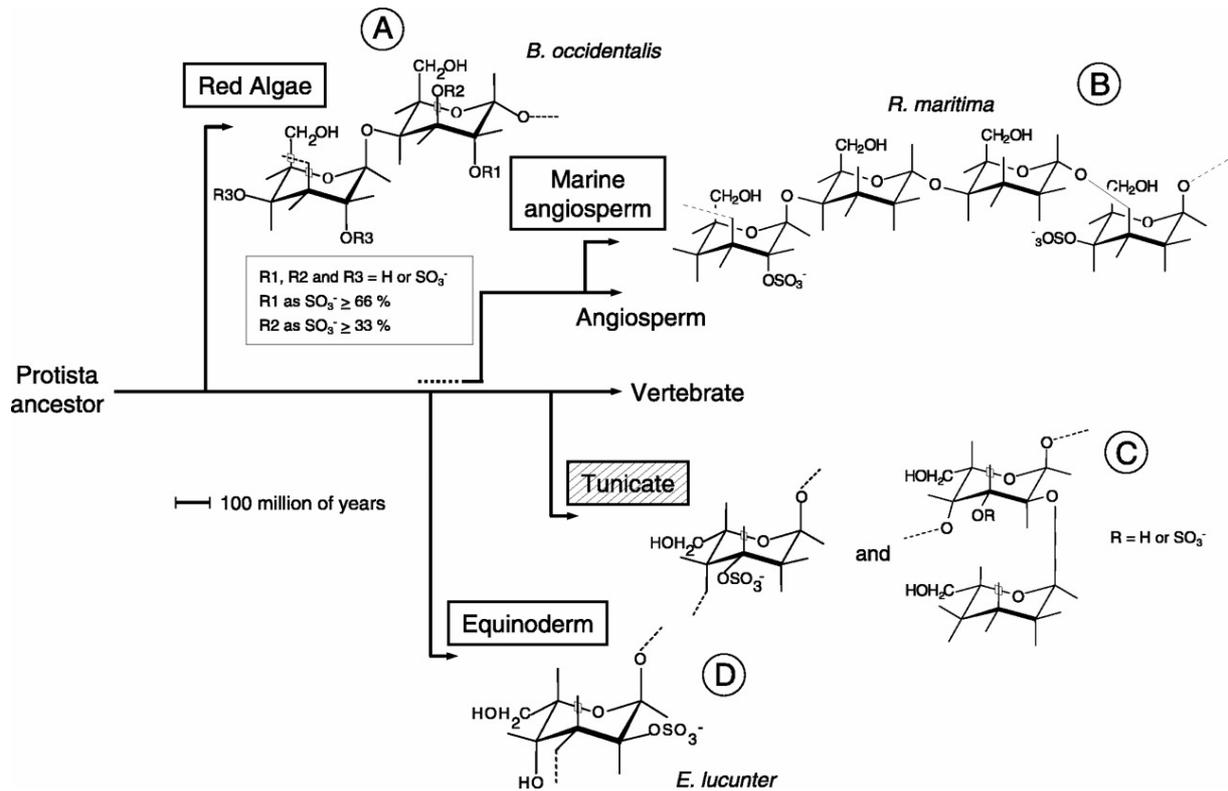


Figure 8 : Les relations phylogénétiques et les temps de divergences des rhodophycées marines, des angiospermes et des invertébrés couplés aux structures de leur galactanes sulfatés (extrait de Aquino et al. 2005). **A)** Les galactanes sulfatés des algues rouges marines ont une structure répétée $[-\rightarrow 4-\alpha\text{-D-Gal-1}\rightarrow 3-\beta\text{-D-Gal-1}\rightarrow]_n$ avec un profil de sulfatation variable. **B)** Les galactanes sulfatés de l'angiosperme marin *R. maritima* a la structure répétée $[-\rightarrow 3-\beta\text{-D-Gal-2}(\text{OSO}_3)\text{-1}\rightarrow 4-\alpha\text{-D-Gal-1}\rightarrow 4-\alpha\text{-D-Gal-1}\rightarrow 3-\beta\text{-D-Gal-4}(\text{OSO}_3)\text{-1}\rightarrow]_n$. **C)** Dans le cas des tunicés, les galactanes sulfatés ont des unités 3-(OSO₃)-4-linked- α -L-Gal, et dans plusieurs espèces des résidus de L-Gal non sulfatés forment des branchements. **D)** Les galactanes sulfatés d'oursins ont une structure répétée régulière composée d'un unique résidu monosaccharidique sulfaté 2-(OSO₃)-3-linked- α -L-Gal. Les angiospermes terrestres et les vertébrés ne contiennent pas de galactanes sulfatés.

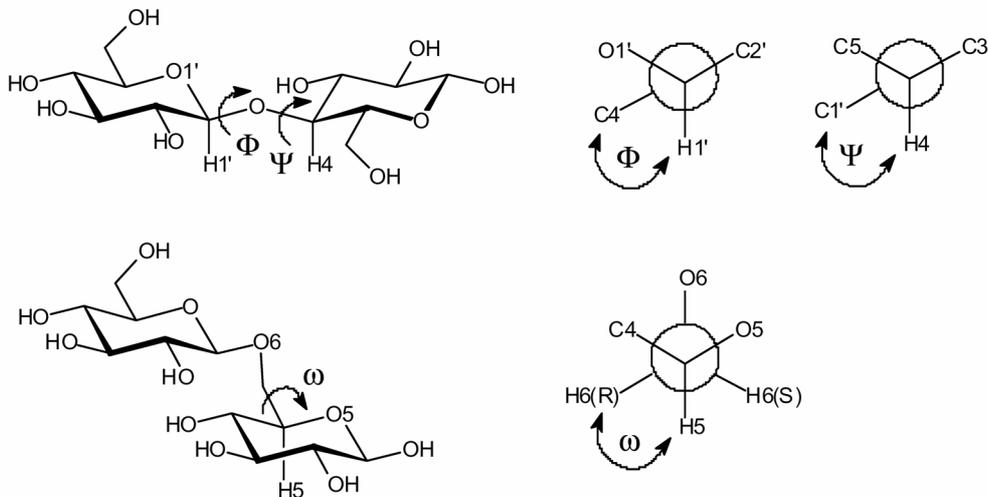


Figure 9 : Définitions des angles dièdres ϕ , ψ , et ω , permettant de décrire la conformation de la liaison glycosidique.

bles suite à des contraintes stériques (Rees et al., 1982). Mais certains couples (φ , ψ) seront également interdits par des répulsions électrostatiques dans le cas de substitutions (ex. sulfates), alors que d'autres pourront être privilégiés par des interactions électrostatiques (ex. liaisons hydrogènes) qui traduiront l'existence de conformations préférentielles (ex. structures secondaires). Par ailleurs, il y a une indépendance presque totale entre un couple (φ_i , ψ_i) et le couple (φ_{i+1} , ψ_{i+1}) (Perez et Kouwijzer, 1999). En outre la flexibilité interne aux résidus contribue de manière prépondérante à l'abaissement des barrières énergétiques des transitions conformationnelles autour de la liaison glycosidique. Ainsi les degrés de liberté permis par celle-ci engendrent malgré tout une forte entropie conformationnelle conduisant un polysaccharide en solution à adopter généralement une structure de type pelote statistique (Perez et Kouwijzer, 1999). D'une manière générale, en solution, les forces d'attraction (liaisons hydrogènes molécule/molécule et molécule/solvant, interactions dipolaires et ioniques, la contribution des interactions hydrophobes est négligeable) qui tendent à favoriser une conformation unique, ne sont pas suffisantes pour figer la flexibilité de la chaîne (Rees, 1975). Compte tenu de la faiblesse de ces interactions, l'entropie conformationnelle ne peut être contrebalancée que par l'établissement d'un grand nombre d'entre elles de manière coopérative (Rees et Welsh, 1977). De plus, en l'absence d'autres contributions enthalpiques, les seules interactions polymère/solvant fournissent seulement un petit incrément à l'entropie de mélange, ce qui rend la dissolution d'un polysaccharide non favorable thermodynamiquement parlant (Perez et Kouwijzer, 1999). Autrement dit, le comportement d'un polysaccharide en solution sera étroitement dépendant des interactions polymère/polymère.

Afin de comprendre les bases moléculaires des arrangements natifs des polysaccharides matriciels, ainsi que de relier leur propriétés et fonctions à leur structure, il est nécessaire de déterminer leurs différents niveaux d'organisation structurale. Comme pour les autres macromolécules, l'élucidation de leur structure primaire est un prérequis. Cette dernière correspond à la séquence de monosaccharides enchaînés covalamment par des liaisons glycosidiques. A la différence des autres biopolymères (protéines et des acides nucléiques) les polysaccharides peuvent avoir une structure primaire ramifiée. La structure secondaire correspond à un arrangement

géométriquement régulier (le couple $[\varphi_i, \psi_i]$ est constant) d'une séquence contenant des répétitions régulières, tel que les rubans, les hélices étirées ou creuses. Des interactions favorisées énergétiquement entre des chaînes de structures secondaires bien définies résultent des structures ordonnées dites tertiaires (i.e. doubles hélices d'agarose ou de carraghénane, « egg-box » d'alginate ou de pectate) qui stabilisent ces structures secondaires. L'association des ces dernières est homophile, et elle se fait donc entre des molécules identiques mais généralement distinctes. Mais cette observation n'exclue pas que dans des mélanges aussi complexes que le sont les parois cellulaires, des polysaccharides différents puissent interagir à travers des structures mixtes. Enfin, l'empilement ou l'agrégation de structures tertiaires constitue un niveau d'organisation nommé structure quaternaire (ex. l'empilement cation-dépendant des doubles hélices de ι -carraghénane et κ -carraghénane)

Les interactions chaîne/chaîne comme celles observées dans les arrangements cristallins, ont également lieu en solution. Lorsqu'elles se développent sur des longueurs et des durées significatives elles peuvent conduire à une réticulation des molécules engendrant des réseaux tridimensionnels communément appelé gels. Lorsque ces interactions sont plus transitoires ou quand elles perturbent significativement le solvant environnant, elles augmentent la viscosité de la solution (Perez et Kouwijzer, 1999). Les structures secondaires et tertiaires impliquent des segments de séquence de longueur variable. Cette dernière est directement liée à l'homogénéité du polysaccharide. En effet, un changement dans la séquence empêche la propagation de la conformation ordonnée et généralement un changement de direction de la chaîne polysaccharidique. Ces interruptions, appelées « kinks » en anglais, peuvent être dues à l'occurrence d'une ou plusieurs unités glucidiques différentes, à une ramification du polysaccharide ou à une substitution irrégulière. Ces points de rupture jouent un rôle crucial dans le contrôle de la densité et la distribution des pontages inter-chaînes et donc dans la porosité et le taux d'hydratation du gel (Rees et Welsh, 1977). Dans le cas de l'agarose, il est même proposé que ces kinks préviennent un arrangement trop parfait qui empêcherait la formation de gel et ferait précipiter ce polymère (Clark et Ross-Murphy, 1987). Cette hypothèse peut être étendue à l'ensemble des galactanes gélifiants (Perez et Kouwijzer, 1999). Néanmoins

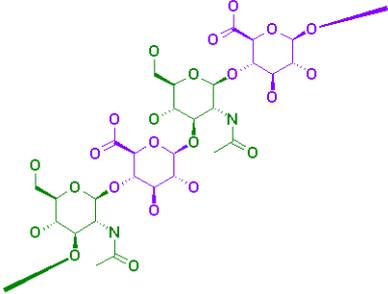
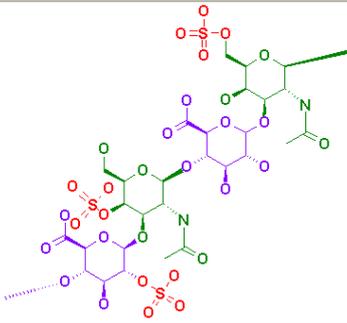
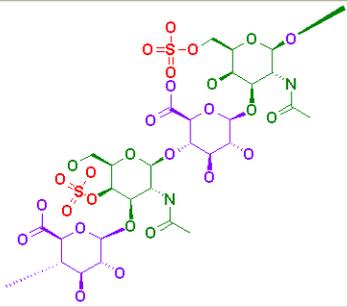
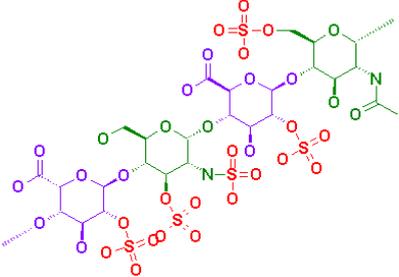
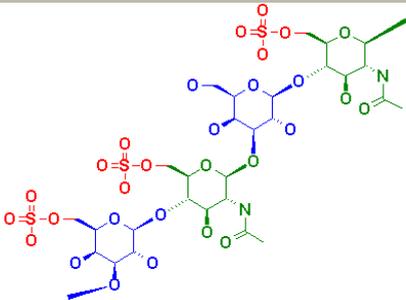
les jonctions inter-chaînes coopératives requièrent une longueur spécifique minimale pour assurer leur stabilité. De ce fait la quantité et la distribution des « kinks » revêtent une importance capitale dans les propriétés d'un gel.

2.2. Les polysaccharides sulfatés animaux

Dans le règne animal, les polysaccharides sulfatés principaux sont les glycosaminoglycanes (GAGs) sulfatés qui sont des composants ubiquistes des tissus connectifs (Falshaw et al., 1999). Ils sont classés en quatre familles structurales : les glucosaminoglycanes d'une part, héparane sulfate et keratan sulfate ; et les galactosaminoglycanes d'autres part, dermatan sulfate, et les chondroïtine 4- et 6-sulfates (tableau 1). Ce sont tous des polymères linéaires répétant des unités disaccharidiques composées d'hexosamine liée en 3 ou 4, et d'un acide uronique lié en 4 (à l'exception du kératane sulfate où ce dernier est un hexose). Ces polysaccharides se différencient par la stéréochimie ainsi que le pattern de sulfatation des disaccharides. Leur structure et leur fonction sont diverses. Les longueurs de chaînes sont de l'ordre de 40 unités disaccharidiques lié covalamment à un peptide, le rapport de masse étant largement en faveur du glycane. Les GAGs ont deux fonctions majeures : un rôle de composant structural de la matrice extracellulaire ; et un rôle de ligand destiné à la reconnaissance de diverses molécules à la surface de la cellule. En tant qu'agent structurant ils interviennent dans les cartilages, l'humeur vitreuse de l'œil, le liquide synovial, etc., se comportant en général comme des absorbeurs de chocs. Ils sont d'ores et déjà incorporés en tant que composés bioactifs, à de multiples produits pharmaceutiques ou ayant trait à la santé (pansements, additifs alimentaires, cosmétiques, matériel chirurgical) mais un de leur principal emploi repose sur les propriétés anticoagulantes de l'héparine (Falshaw et al., 1999).

Les mucines sont des oligosaccharides sulfatés liés par liaison O-glycosidiques à des sérine ou thréonine des protéines. Elles doivent leur nom au fait que ces glycoprotéines constituent les mucus excrétés par les cellules des membranes muqueuses. On les retrouve dans les mucus pulmonaire, buccaux et gastro-intestinaux où ils participent à la lubrification et la protection des organes concernés vis à vis de la

Tableau 1 : Les glycosaminoglycane (GAGs) sont les principaux polysaccharides sulfatés animaux

	<p>Hyaluronates:</p> <p>D-glucuronate + N-acétyl-glucosamine(GlcNAc)</p>
	<p>Dermatane sulfate:</p> <p>D-glucuronate (GlcA) ou L-iduronate (IdoA) + GalNAc-sulfate</p> <p>L'hétérogénéité du dermatane sulfate résulte du degré variable de sulfatation et de la présence des deux acides uroniques.</p>
	<p>Chondroïtine sulfate:</p> <p>D-glucuronate + GalNAc-6 (ou 4)-sulfate</p>
	<p>Héparine et Héparane sulfate:</p> <p>D-glucuronate-2-sulfate + N-sulfo-D-glucosamine-6 (ou 2)-sulfate</p> <p>(les héparanes ont moins de sulfate que les héparines)</p>
	<p>Kératane sulfate:</p> <p>galactose + galactose-6-sulfate + GlcNAc-6-sulfate</p>

flore microbienne.

Chez les animaux marins, enfin, plusieurs types de polysaccharides sulfatés ont été observés. La tunique des ascidies contient ainsi des polymères de L-galactane (Pavão et al., 1989). Santos et col. (1992) ont ainsi isolé un homopolymère de $\alpha(1,4)$ -L-galactopyranosyl-3-sulfate de *Herdmania monus*. Par ailleurs du $\alpha(1,3)$ -L-galactopyranosyl-2-sulfate a été identifiés dans l'œuf de l'oursin *Echinometra lucunter* (Alves et al., 1997). Cependant les principaux polysaccharides sulfatés d'Echinodermes demeurent les homofucanes qui sont décrits chez les oursins et holothuries. Ils feront l'objet d'une description détaillée dans la section concernant les fucoïdanes.

2.3. Les carraghénanes

La présence de D-galactose dans les carraghénanes est connue depuis plus d'un siècle. Cependant l'existence d'ester-sulfate substituant les hydroxyles et celle du 3,6-anhydro-D-galactose ont été révélées plus récemment. Les polysaccharides de la famille des carraghénanes sont des polymères de $[\rightarrow 3)\text{-}\beta\text{-D-galactopyranosyl-(1}\rightarrow 4)\text{-}\alpha\text{-D-galactopyranosyl-(1}\rightarrow]_n$ où le $\alpha\text{-D-galactopyranosyl}$ peut être modifié par un pont 3,6-anhydre (Fig 10). A ce jour, ce résidu n'a jamais été observé vierge ; il porte systématiquement une modification chimique le différenciant du simple galactose (Knutsen, 1994). Sur cette chaîne polymérique, les groupes hydroxyles peuvent être sulfatés, et plus rarement méthylés ou glycosylés par des monosaccharides. En outre il arrive que le $\beta\text{-D-galactopyranosyl}$ soit substitué sur les positions 4 et 6 par un ketal cyclique dérivant du pyruvate. L'ensemble de ses substitutions génère des variants disaccharidiques dont la répétition au sein d'une même molécule est en général prononcée. Toutefois il ne faut pas envisager les molécules de carraghénane comme de purs homopolymères. Il s'agit en effet de molécules hybrides qui comportent des zones irrégulières qui masquent partiellement la régularité du squelette. Par ailleurs, les variants disaccharidiques observés dérivent de la biosynthèse des carraghénanes. Ils sont donc non aléatoires et en nombre restreint. Dès lors, des classifications de ces

différents motifs dits canoniques ont été mises en place afin de faciliter les nomenclatures.

2.3.1. Classification et nomenclature

Les premières notions de classification sont apparues très tôt en se basant sur le critère de solubilité en présence de KCl (Smith et Cook, 1953) couplé aux premières structures chimiques élucidées (O'Neill, 1955 ; Dolan et Rees, 1965 ; Anderson et al., 1965). Les investigations se poursuivant, ce critère de classification dichotomique trouva rapidement ces limites, alors que, de façon concomitante, la notion de motifs précurseurs (du point de vue de la biosynthèse) apparaissait (Anderson et al., 1968 ; Craigie et Wong, 1978) permettant de regrouper les différents motifs structuraux en famille respectant l'histoire de leur biosynthèse (Craigie et Wong, 1978 ; McCandless et Craigie, 1979). Cette classification utilise des lettres grecques pour désigner les motifs disaccharidiques canoniques ou idéaux, c'est à dire les plus abondants (Fig. 10, Tab. 2). La limite de ce système réside effectivement dans le fait

Tableau 2 : Nomenclature des carraghénanes préconisée par Knutsen et col. (1994).

Code lettre	type de carraghénanes	nomenclature IUPAC* (recommandée en 1996)
G	$\alpha, \beta, \gamma, \delta$	3-linked β -D-galactopyranose
D	inexistant à l'état naturel	4-linked α -D-galactopyranose
DA	κ, β, ω	4-linked 3,6-anhydro- α -D-galactopyranose
S	$\kappa, \iota, \lambda, \mu, \nu, \theta$	ester sulfate ($O-SO_3^-$)
P	π	4,6-O-(1-carboxyéthylidène)
G2S	λ, ξ, θ	3-linked β -D-galactopyranose-2-sulfate
G4S	$\kappa, \iota, \mu, \nu, \omega$	3-linked β -D-galactopyranose-4-sulfate
G6S	ψ, ω	3-linked β -D-galactopyranose-6-sulfate
GP,2S	π	3-linked β -D-galactopyranose-2-sulfate-4,6-O-(1-carboxyéthylidène)
D2S	ξ, ω	4-linked α -D-galactopyranose-2-sulfate
D6S	γ, μ, ψ	4-linked α -D-galactopyranose-6-sulfate
D2S,6S	δ, λ, ν	4-linked α -D-galactopyranose-2,6-disulfate
DA2S	α, ι, θ	4-linked 3,6-anhydro- α -D-galactopyranose-2-sulfate

*International Union of Pure and Applied Chemistry

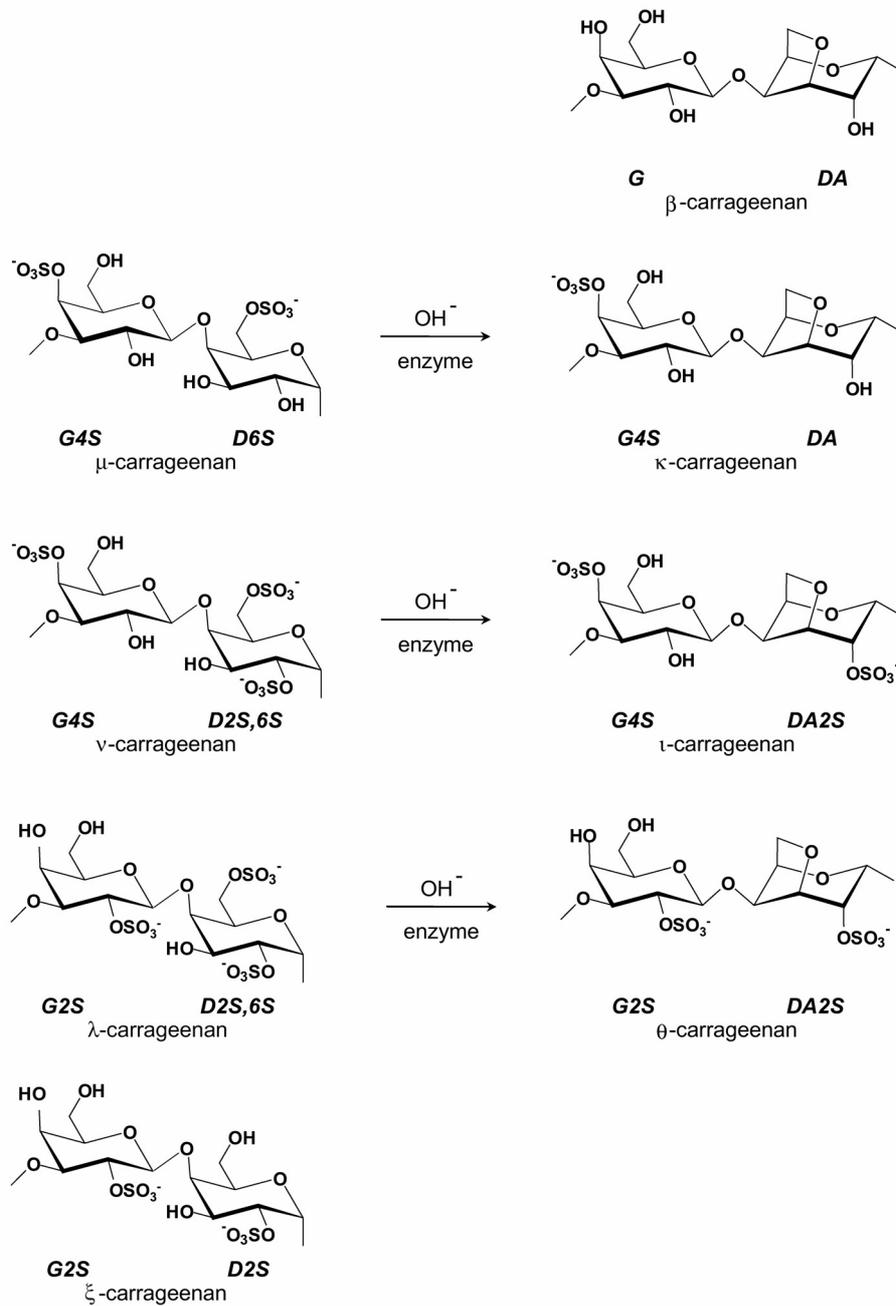


Figure 10 : Représentation schématique des unités répétitives des principaux types de carraghénanes (adapté de Van de Velde et al., 2004, 2002). Le code lettre est celui préconisé par Knutsen et col. (1994) développé dans le tableau 2.

que la diversité des motifs structuraux observée ne s'y inscrit pas entièrement. C'est ce qui a motivé Knutsen et ses collaborateurs à proposer en 1994 un nouveau système de nomenclature plus rationnel tenant compte de la nature hybride des carraghénanes (Tab. 2). Ce système considère les galactanes dans leur ensemble satisfaisant de cette façon les observations faisant part de la présence de L-galactose dans les carraghénanes (Parolis, 1978).

2.3.2. La famille des λ -carraghénanes

Chez les carraghénophytes, le gamétophyte (mâle ou femelle) contient un mélange de κ/ι -carraghénanes (famille κ), alors que les carraghénanes de la famille λ caractérisent en général le tétrasporophyte (Craigie, 1990). La famille des λ -carraghénanes regroupe les carraghénanes dont l'unité G (β -D-galactosyl) est toujours sulfatée en C2. Elle comprend donc selon la nomenclature traditionnelle:

- le **λ -carraghénane** défini idéalement comme étant le $[\rightarrow 3)\text{-}\beta\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow 4)\text{-}\alpha\text{-D-galactopyranosyl-2,6-sulfate-(1}\rightarrow]_n$ (ou G2S-D2S,6S),
- le **ξ -carraghénane** défini comme étant le $[\rightarrow 3)\text{-}\beta\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow 4)\text{-}\alpha\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow]_n$ (ou G2S-D2S) à qui il manque donc le sulfate en C6,
- le **π -carraghénane** défini comme étant le $[\rightarrow 3)\text{-}4,6\text{-carboxyethylidène-}\beta\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow 4)\text{-}\alpha\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow]_n$ (ou GP,2S-D2S) qui dérive du ξ -carraghénane via un pyruvate pontant les positions C4 et C6,
- le **θ -carraghénane** défini comme étant le $[\rightarrow 3)\text{-}\beta\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow 4)\text{-}3,6\text{-anhydro-}\alpha\text{-D-galactopyranosyl-2-sulfate-(1}\rightarrow]_n$ (ou G2S-DA2S) qui est une structure à ce jour rarement observée in vivo et découlant artificiellement du λ -carraghénane suite à un traitement alcalin au NaBH_4 .

A l'exception du λ -carraghénane, tous les carraghénanes sulfatés en C6 sur le résidu D (les γ -, δ -, ψ -, ν -, et μ -carraghénanes) sont des précurseurs de biosynthèse

conduisant sous l'action d'une sulfohydrolase à une désulfatation en C6 couplée à la formation du pont 3,6 anhydre (Craigie et Wong, 1978). En effet le λ -carraghénane n'est pas substrat de l'enzyme (sulfurylase) catalysant cette conversion chez le κ - et le ι -carraghénane (Craigie et Wong, 1978). En outre, le traitement alcalin mimant cette cyclisation est 20 à 30 fois plus rapide avec les carraghénanes de la famille κ (Ciancia et al., 1993). Ces auteurs suggèrent que dans le λ -carraghénane, les interactions électrostatiques (liaison hydrogène) entre l'hydroxyle en C3 du résidu D et le groupe sulfate C2 du résidu G sont responsables de la chute de vitesse de cette réaction conduisant au θ -carraghénane. L'hypothèse de cette interaction est appuyée par les modèles moléculaires et a de surcroît été décrite dans l'héparine sulfate (Ciancia et al., 1993).

2.3.3. Composition des λ -carraghénanes

Le λ -carraghénane est essentiellement isolé des sporophytes de Gigartinales (ordre). Depuis, les quinze dernières années les investigations concernant les différents genres de cet ordre se sont intensifiées, faisant suite aux travaux préliminaires/pionniers des années 60/70. Ces derniers ont initié les études fines concernant le profil de sulfatation du λ -carraghénane. C'est en effet à cette époque que l'étude du λ -carraghénane de *Chondrus crispus* (Dolan et Rees, 1965 ; McCandless et al., 1979) a révélé la présence de 30% de résidu G désulfatés en C2 (δ -carraghénane, famille β). Par ailleurs, en 1973, Penman et Rees sont les premiers à décrire le ξ -carraghénane dans *Gigartina atropurpurea*, *G. canaliculata*, et *G. chamissoi* suite à des analyses par méthylation. Plus récemment, l'équipe argentine de Cerezo, Noseda, Matulewicz, Ciancia et Stortz s'est montrée particulièrement active, au même titre que Furneaux et Falshaw en Nouvelle Zélande.

Plusieurs études employant des approches type fractionation au KCl, méthylation, et spectroscopie IR, ont montré l'hétérogénéité des extraits de *Gigartina skottsbergii* contenant du λ -carraghénane (37% p/p de l'algue ; 45% p/p des polysaccharides totaux) avec un taux molaire de sulfatation de 1,12 par galactose (Matulewicz et al., 1989, 1990) et un poids moléculaire estimé entre 80000 et

120000Da (Carlucci et al., 1997). Si certaines des fractions obtenues semblent relativement pures, elles indiquent cependant la présence de 3 résidu DA pour 100 unité galactose.

La haute viscosité des solutions de λ -carraghénane même à chaud rend difficile l'obtention de spectres ^{13}C -RMN bien résolus. McCandless et Gretz (1984) obtinrent de spectres simples et bien résolus après avoir traité le λ -carraghénane extrait de *I. cordata* par une λ -carraghénase. Rochas et Lahaye, en 1989, utilisèrent l'ultrasonication pour réduire la longueur des chaîne, et consécutivement la viscosité d'une solution de λ -carraghénane extrait de *G. canaliculata*, avant l'acquisition de spectres RMN. Ces derniers sont bien résolus mais leur complexité indique la nature hybride du polymère observé. Une approche identique par spectroscopie ^{13}C -RMN menée sur du λ -carraghénane soniqué extrait de *Iridaea undulosa* montre un λ -carraghénane très homogène et très proche du λ idéal contenant toutefois des motifs minoritaires portant un sulfate supplémentaire en C4 ou C6 du résidu G non répertorié dans la nomenclature usuelle par lettres grecques (Stortz et al., 1994 ; Stortz et Cerezo, 1993). A l'inverse le λ -carraghénane commercial fourni par Sigma s'avère très hétérogène puisqu'il y est détecté du ξ - et du π -carraghénane (Stevenson et Furneaux, 1991). Les investigations très complètes (méthylation, désulfatation, traitement alcalin, spectroscopie IR et ^{13}C -RMN) menées sur *G. decipiens* (Falshaw et Furneaux, 1994) montrent la prépondérance du motif λ idéal (G2S-D2S,6S), accompagné toutefois de G2S6S-D2S6S (estimé à 15%), alors que quasiment tout les résidus G et D sont sulfatés respectivement en C2 (G2S), et en C2 et C6 (D2S,6S). Les mêmes auteurs ont analysé similairement les très complexes λ -carraghénanes de *G. clavifera* et *G. alveata* qui s'éloignent franchement du λ idéal (Falshaw et Furneaux, 1995). Ils montrent la prédominance dans ces deux algues du ξ -carraghénane et la sulfatation constante du résidu G en C2 (G2S). Néanmoins, un résidu G2S sur six porte un acétal pyruvate (GP2S) caractéristique du π -carraghénane. Ce même résidu G2S apparaît faiblement mais significativement substitué par un sulfate en C6 (G2S,6S). Un résidu D sur six est complètement dépourvu de sulfate alors qu'une proportion un peu plus forte est sous la forme « classique » D2S,6S. Enfin, une proportion faible mais significative de résidus D porte un pont 3,6-anhydro et un sulfate en C2 (DA2S du θ -carraghénane). Les

mêmes auteurs toujours, avec la même méthodologie, se sont attachés à décrire le λ -carraghénanes de *G. lanceata* et *G. chapmanii* (Falshaw et Furneaux, 1998). La première espèce possède un λ -carraghénane quasiment idéal alors que celui de la seconde est principalement composé de G2S-D2S (ξ -carraghénane), avec un résidu G sur sept sous forme GP (π -carraghénane), un résidu D sur quatre est D2S,6S (λ -carraghénane idéal), et une proportion faible mais significative de DA2S (θ -carraghénane). Des extraits de λ -carraghénane provenant de *Stenogramme interrupta* (Phylloporaceae, Gigartinales) ont été fractionnés au KCL, auto-hydrolysés et analysés par spectroscopies IR et ^{13}C -RMN (Càceres et al., 2000). Ils contiennent du G2S, du D2S (ξ -carraghénane) et du D2S,6S (λ -carraghénane), aucun GP n'est détecté (π -carraghénane), et de façon surprenante 16% des galactose serait sous forme anhydre soit donc 32% des résidus D sous forme DA. En outre, le ratio molaire de sulfate est bas (0,60 par unité galactose). La dernière étude en date concerne le λ -carraghénane de *G. pistillata* (Amimi et al., 2001). Se basant majoritairement sur l'analyse de dérivé alditols acétates méthylés par spectrométrie de masse, elle illustre parfaitement l'hétérogénéité de ce polymère. En effet, seulement 20% des disaccharides adopteraient le motif λ idéal G2S-D2S,6S. De surcroît, elle révèle une importante et inusuelle proportion du motif G6S,2S-D2S,6S (près de la moitié des disaccharides). Enfin du G2S,4S, du GP2S, et du DA2,S sont détectés en faible proportion.

2.3.4. Structures secondaires, tertiaire et quaternaire des carraghénanes : leur impact sur la solubilité et la gélification

Une des principales caractéristiques de ces hydrocolloïdes est leur capacité à faire des gels à basse concentration (pour l'agarose la concentration limite est de 0,2% w/w). En plus d'être thermoréversibles les gels de carraghénane sont iono-réversibles du fait de l'intervention de cations spécifiques dans les interactions entre doubles hélices. A haute température le ι -carraghénane, le κ -carraghénane et l'agarose existent en solution sous forme de pelotes statistiques, alors que, après refroidissement une structure rigide et ordonnée apparaît selon une transition coopérative (Fig. 11). Il est admis que les gels formés par ces composés sont réticulés par des agrégats ordonnés

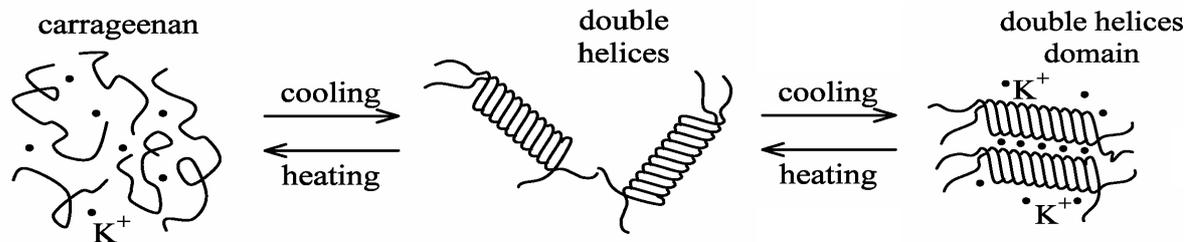


Figure 11 : Schématisation des transitions sol-gel et gel-sol du κ -carraghénane en solution de KCl (d'après Kara et al., 2003). Ce schéma s'applique également au ι -carraghénane en présence d'ions bivalents (Ca^{2+}).

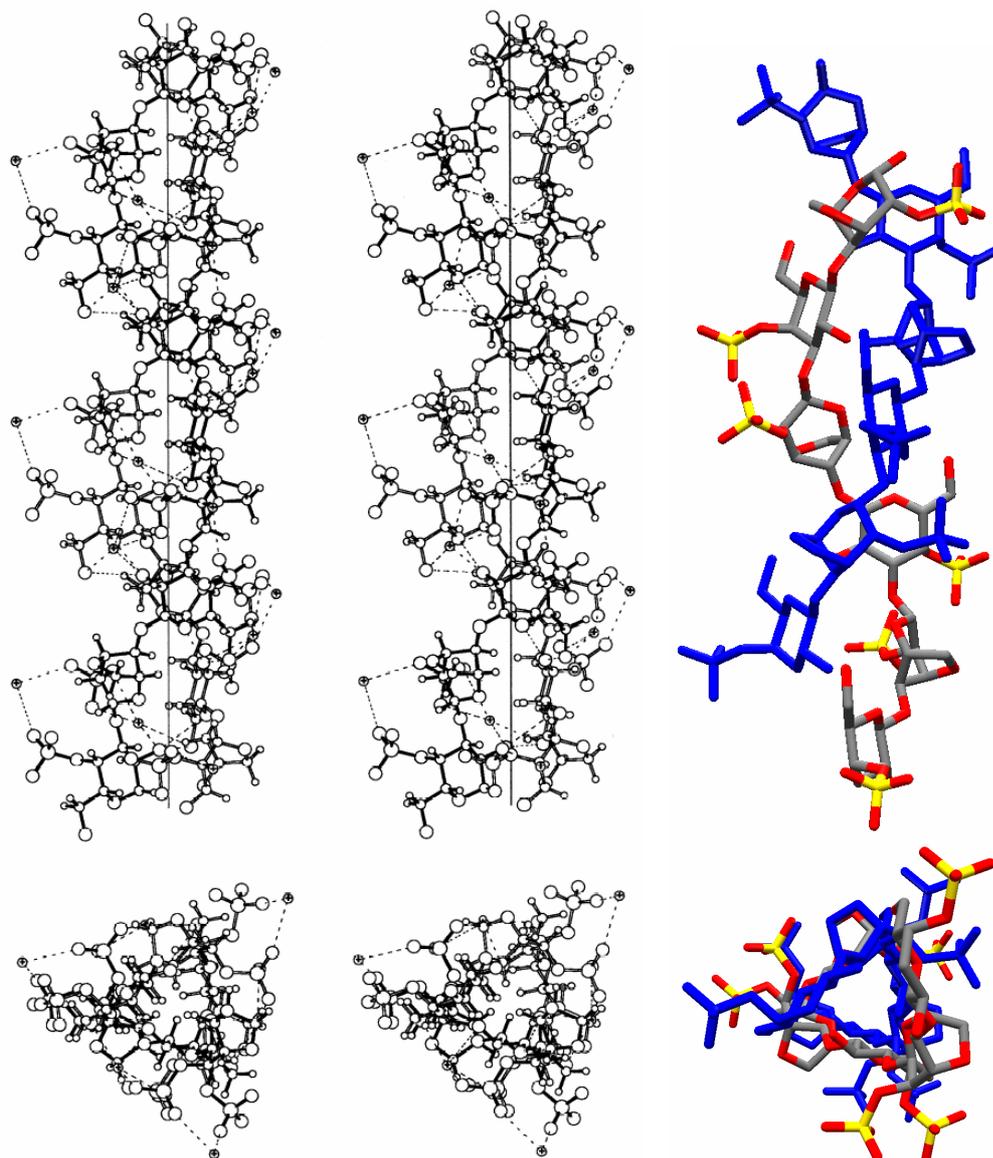


Figure 12 : Représentation stéréographique des deux vues perpendiculaires d'une double hélice de ι -carraghénane dont la structure a été obtenue par diffraction de fibres (extrait de Janaswamy et Chandrasekaran, 2002). Elle est stabilisée par des liaisons hydrogènes interchaînes (en pointillés). Les hydrogènes figurent sur ces structures ainsi que les molécules d'eau structurantes (cercles gris). Dans le but de faciliter la compréhension de la structure, à droite est présentée une double hélice dont un brin est coloré par type atomique (gris, carbone ; rouge, oxygène ; jaune, soufre).

de doubles hélices (impliquant jusqu'à dix chaînes). Des résidus irréguliers dont le pont 3,6 anhydro est manquant induisent la formation de « kinks ». En effet, la présence de la cyclisation sur le résidu D, impose la conformation 1C_4 au lieu de 4C_1 (Rees et al., 1982) ; son absence occasionne donc un changement géométrique du squelette. Ces kinks limitant la longueur des hélices limitent celle des zones de jonctions. En trop grand nombre ils bloquent la mise en place d'un réseau tridimensionnel. Mais en leur absence la connectivité inter-chaîne chute ce qui empêche également la gélification (Van de Velde et al., 2002A).

La densité de charge du polysaccharide (ι -carraghénane > κ -carraghénane > agarose) affecte la solubilité et les propriétés gélifiantes. Plus il y a de substituants sulfates meilleure est la solubilité. Ainsi l'agarose forme des gels forts, rigides, et sujets à la synérèse (exsudation de solvant). Le κ -carraghénane produit également des gels rigides en présence de potassium mais plus élastiques. Cependant avec des ions calcium ces gels de κ -carraghénane deviennent moins flexibles, cassant et sujet à synérèse. D'une manière générale, le ι -carraghénane génère lui des gels plus élastiques et plus fragiles. Leur intérêt repose notamment sur leurs propriétés thixotropiques et l'absence de synérèse. Ils sont bien plus performants en présence d'ions calciums que sodium. L'augmentation de la densité de charge a pour effet d'étendre les agrégations et d'augmenter l'hystérésis (voir ci-dessous). La concentration critique nécessaire à la gélification diminue alors que les gels deviennent plus fragiles (Perez et Kouwijzer, 1999). Notons par ailleurs que la nature des contre-ions des groupements sulfates affecte également le comportement de ces polysaccharides en solution.

Les hélices du κ -carraghénane et de l'agarose sont stabilisées par leur agrégation. Le chauffage entraîne la fusion de ces agrégats et la perte associée de la structure du gel. Cette fusion se fait à une température supérieure à celle à laquelle la gélification a lieu lors du refroidissement (boucle d'hystérésis). Dans les ι -carraghénane et κ -carraghénane, le mécanisme de gélification est le suivant (Fig 11). A des températures supérieures à un seuil critique T_A ces polymères sont sous forme de pelotes statistiques. A la température T_A , des doubles hélices commencent à se former (transition pelotes-hélices). Néanmoins elles ne sont pas encore assez nombreuses pour former un réseau de structure tertiaire. A la température T_B ($<T_A$), leur quantité devient

critique et la réticulation débute. A la température T_C ($<T_B$), les doubles hélices commencent à s'agréger (en présence de cations) pour former des structures quaternaires de type fibres. En deçà d'une température T_D ($<T_C$), la formation des doubles hélices est plus importante que leur agrégation. Ensuite la température diminuant, l'importance des deux phénomènes s'inverse. La gélification des carraghénanes a lieu essentiellement lors de cette étape d'agrégation.

Parmi cette famille de polysaccharides, le ι -carraghénane est le plus caractérisé à l'état solide. Sa structure ordonnée a été résolue par la diffraction aux rayons X de fibres cristallines comme pour le κ -carraghénane (Bailey, 1955 ; Anderson et al., 1969 ; Janaswamy and Chandrasekaran, 2001, 2002). Il forme des doubles hélices droites au sein desquelles les brins sont disposés parallèlement (Fig. 12). Il y a trois disaccharides par tour. Les hydroxyles non substitués O6 et O2 des G4S, sont impliqués dans des liaisons hydrogènes (O2-H \cdots O5 et O6-H \cdots O2) inter-brins, parallèles à l'axe de l'hélice, qui assurent la cohésion de la structure tertiaire (Janaswamy and Chandrasekaran, 2001, 2002). Aucune liaison hydrogène intra-chaîne n'est observée. La section de la double hélice est triangulaire portant à ses coins les groupements sulfates qui pointent vers l'extérieur (Fig. 12). Ces études structurales sur fibres orientées montrent qu'avec des ions sodium, les doubles hélices forment un réseau trigonal stabilisé par des interactions électrostatiques $\text{SO}_3^- \text{-Na}^+ \text{-SO}_3^-$ ou $\text{SO}_3^- \text{-Na}^+ \text{-H}_2\text{O} \text{-Na}^+ \text{-SO}_3^-$. Avec des ions calcium, les interactions entre doubles hélices sont de type $\text{SO}_3^- \text{-Ca}^{2+} \text{-SO}_3^-$. Malgré qu'elles soient plus fortes, la géométrie des hélices est exceptionnellement similaire puisque seule l'orientation du sulfate en C4 change d'orientation.

Le λ -carraghénane ne gélifie pas ; il est d'une part plus sulfaté que les ι - et κ -carraghénane, d'autre part il ne porte pas le pont 3,6 anhydro, et enfin sa structure primaire est plus irrégulière. Il se différencie par la seule position du sulfate sur le résidu G (2 au lieu de 4) du précurseur ν -carraghénane dont on sait qu'il constitue les zones amorphes du ι -carraghénane. Ce polymère ne montre pas de transition pelote-hélice. En solution, et il adopterait une structure type pelote statistique désordonnée sous des conditions variables de température et d'environnement ionique (Perez et Kouwijzer, 1999). Peu d'études expérimentales sont disponibles concernant la

conformation du λ -carraghénane en solution. Des tentatives de diffraction de fibre aux rayons X n'ont pas apporté de résultats réellement exploitables (Bailey, 1955). Rees (1969) les interprète néanmoins et propose une structure type ruban plat étiré. Le sulfate supplémentaire a probablement pour effet de réduire l'espace conformationnel.

Les modélisations moléculaires réalisées par Le Questel et col. (1995) tendent à montrer que, à l'état de pelote statistique, les chaînes de ι - et κ -carraghénanes seraient plus flexibles que celles de λ -carraghénane. En outre les chaînes de λ -carraghénanes adopteraient en solution une forme plus étirée. Leurs calculs d'énergie de conformation montrent que la liaison $\alpha(1,3)$ est moins entravée par le sulfate en 2 que la liaison $\beta(1,4)$. Ces hypothèses sont en accord avec la forte viscosité des solutions de λ -carraghénanes.

2.4. Les fucoïdanes

Le premier isolement de fucoïdane à partir d'algues brunes remonte au début du XX^{ème} siècle (Kylin, 1913). La présence de ce polysaccharide chez des invertébrés marins sera révélée 35 ans plus tard (Vasseur, 1948). Il est facilement extrait des parois cellulaires d'algues brunes par dissolution dans l'eau chaude ou dans des solutions acides. Chez les invertébrés on le retrouve dans le manteau gélatineux des œufs d'oursins (Mulloy et al. 1994) et dans la cuticule des concombres de mer (Mourão et Bastos, 1987). Les fucoïdanes d'algues sont présent dans plusieurs ordres des Phaeophyceae, principalement les Fucales et les Laminariales, mais aussi les Chordariales, les Dictyotales, les Dictyosiphonales, les Ectocarpales et les Scytosiphonales. En fait on les retrouve dans toutes les algues brunes où on les a cherché à ce jour. Ils sont cependant absents des algues vertes (Chlorophyceae), rouges (Rhodophyceae) et dorées (Xanthophyceae), ainsi que des algues d'eau douces et des plantes terrestres. Cette exclusivité du milieu marin soulève la question de la fonction biologique de ce polymère. Mais force est de constater que les fonctions biologiques du fucoïdane demeurent faiblement documentées au regard des nombreuses données disponibles concernant son activité dans les systèmes de type mammifères. Chez les algues il est proposé qu'ils interviennent comme structurant pariétal bien qu'ils soient

dépourvus de caractéristiques rhéologiques particulières contrairement aux alginates. Une autre fonction leur est attribuée, de mon point de vue, plus consistante. Ils pourraient en effet intervenir dans les régulations hydrique et cationique de la paroi (Kloareg et Quatrano, 1988). En effet, plus les algues sont hautes sur l'estran, plus elles sont soumises à la dissécatation suite au phénomène de marée ce qui est corrélé par une teneur croissante en fucoïdane (Kloareg et Quatrano, 1988). En ce qui concerne les holothuries, Mourão et Bastos proposent en 1987 un rôle équivalent pour les fucoïdanes de leur tunique. Enfin dans les oeufs oursins, si les fucoïdanes se trouvent encore à l'interface milieu/organisme, il est maintenant prouvé qu'ils interviennent dans le processus de fécondation en permettant la spécificité de reconnaissance du spermatozoïde (la réaction acrosomiale) afin de constituer une barrière d'espèce (Alves et al., 1997).

Le fucoïdane a pour composant principale le L-fucose qui se différencie du L-galactose uniquement par l'absence d'un hydroxyle en C6. L'isomérisation L le distingue donc des carraghénanes. Toutefois on peut envisager un cousinage structural avec les galactanes composés de L-galactose. Ils partagent en outre la présence de substituants sulfate. Cependant aucune particularité rhéologique ne caractérise le fucoïdane, ni viscosité, ni pouvoir gélifiant. Je n'irais pas plus loin dans les tentatives de rapprocher les deux volets de ce travail à travers ses deux polysaccharides modèles.

2.4.1. Nomenclature

La complexité structurale de ce polymère est à l'origine dans la littérature d'une nomenclature relativement floue étroitement liée à son niveau de pureté. Je me propose donc ici de clarifier cette situation. Kylin (1913) baptise son polysaccharide « fucoïdin », 40 ans plus tard McNeely (1959) le modifie en « fucoïdan » afin d'être conforme à la nomenclature générale des polysaccharides. Soupçonnant la présence de résidus contaminants, différents auteurs proposèrent de supprimer le terme « fucoïdane » au profit de l'appellation générique « fucane » désignant un polysaccharides riche en L-fucose (Larsen, 1966). Les techniques d'analyses et de séparation devenant plus performantes, plusieurs familles de polysaccharides sulfatés

furent identifiées dans les préparations de fucanes (décrites au § 1.2.3.4.). Il s'agit des ascophyllanes (ou xylofucoglycuronanes), des sargassanes (ou glycuronofucoglycane), des homofucanes correspondant à la définition initiale des fucoïdanes. A l'heure actuelle, on retrouve dans la littérature le terme périmé « fucoïdin » ; dans certains cas des auteurs créent leur propre termes, tels que fucanesulfate ; et la confusion est entretenue par certains travaux évoquant le fucoïdane pour décrire des polysaccharides contenant moins de 50% de fucose.

Pour conclure, dans ce travail, conformément aux recommandations de l'IUPAC et en contradiction avec les descriptions précédentes, le terme « fucane » sera utilisé pour décrire un polymère de L-fucose très homogène contenant moins de 10% de résidus contaminants. On pourra donc qualifier de fucanes sulfatés les polysaccharides très homogènes extraits des invertébrés marins. Le terme « fucoïdane » sera réservé aux préparations de polymère contenant du L-fucose sulfaté, typiquement ceux d'algues, étant entendu qu'il est moins restrictif que « fucane » et qu'il tolère plus d'incertitudes quant à la composition des polysaccharides concernés. Néanmoins l'appellation fucoïdane sera dédiée à des polysaccharides majoritairement composés de L-fucose excluant donc sargassanes et ascophyllane.

2.4.2. Structure et composition des fucoïdanes

Des études structurales ont été menées pour les deux grands groupes de fucoïdanes (algues/échinodermes). Elles ont souvent été motivées par l'attrait d'applications biomédicales ce qui explique leur intensification au cours des 15 dernières années. Compte tenu de la diversité des fucoïdanes, plutôt que d'en dresser un catalogue écrit risquant d'être rébarbatif, j'ai préféré récapituler sous la forme d'un tableau les différents résultats connus à ce jour concernant leur structure (Tableau 3). Ce dernier répertorie espèce par espèce les constituants, les motifs structuraux, ainsi que le mode d'obtention des fucoïdanes décrits dans la littérature. Toutefois j'ai souhaité ne pas y inclure de travaux trop anciens se basant sur des méthodes analytiques trop imprécises.

Tableau 3 : Revue bibliographique des études structurales concernant les fucoïdanes.

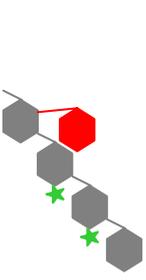
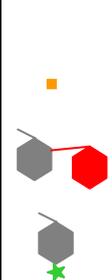
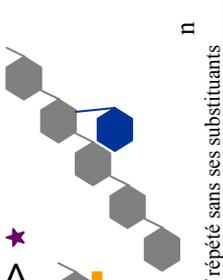
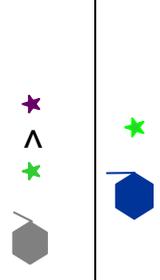
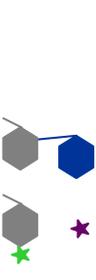
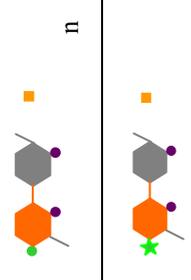
ESPECES (famille)	MOTIFS STRUCTURAUX IDENTIFIÉS	CARACTERISTIQUES GENERALES	OBTENTION DU FUCOÏDANE	REFERENCES
<i>Cladosiphon okamuranus</i> (Chordariaceae)	 <p>n</p>	<ul style="list-style-type: none"> - le motif contient un acétyl non localisé - produits de digestion enzymatique - n de 1 à 3 	<ul style="list-style-type: none"> - Extraction non exposée - Oligosaccharides issus de digestion enzymatique et purifiés par chromatographie à partir de 25% du fucoïdane total 	Sakai et al. (2003 A)
<i>Chorda filum</i> (Chordaceae)	 <p>n</p>	<ul style="list-style-type: none"> - DS=0.5 - DA=0.17 - 1 GU pour 6 F - Linéaire - F 64%, S 27%, GU 0% w/w - DF>0.95 - DA=0.20 - MW>100kDa - branchement élevé 	-	Nagaoka et al. (1999)
<i>Laminaria brasiliensis</i> (Laminariaceae)	 <p>motif répété sans ses substituents</p>	<ul style="list-style-type: none"> - DF=0.90 (10% Gal) - DS=0.90 - MW=150 kDa - branchement - DF=0,88 (12% Gal) - DS=0.77 	<ul style="list-style-type: none"> - Fucoïdane fractionné par extraction séquentielle et purifié par chromatographie - Fractionnement correspondant à 7% de la biomasse 	Chizhov et al. (1999)
<i>Adenocystis utricularis</i> (Adenocystaceae)		<ul style="list-style-type: none"> - DF=0.80 (20% Gal) - CH 60%, S 25%, GU 5% w/w - MW>100 kDa 	<ul style="list-style-type: none"> - Fucoïdane purifié par chromatographie - Fractionnement correspondant <5% du fucoïdane total 	Mulloy et al. (1994)
<i>Fucus distichus</i> (Fucaceae)	 <p>n</p>	<ul style="list-style-type: none"> - DS=1.21 - légère sous-sulfatation en C4 - DA=0.08 	<ul style="list-style-type: none"> - Fucoïdane purifié par chromatographie et fractionné - Fractionnement correspondant à 46% du fucoïdane extrait 	Bilan et al. (2004)
<i>Fucus evanescens</i> (Fucaceae)	 <p>n</p>	<ul style="list-style-type: none"> - DS=1.23 - Sulfatation en C4 dans 66% des cas - DA=0.36 - acétylation aléatoire en C3 et C4 	<ul style="list-style-type: none"> - Fucoïdane purifié par chromatographie et fractionné - Fractionnement correspondant à 47% du fucoïdane total 	Bilan et al. (2002)

Tableau 3 : Revue bibliographique des études structurales concernant les fucoidanes.

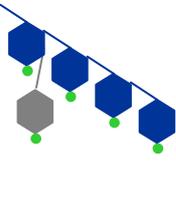
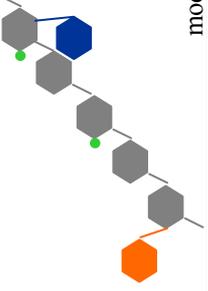
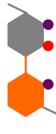
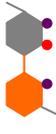
<i>Fucus vesiculosus</i> (Fucaeeae)		-		Percival et McDowell (1967)
	model 	- DS= 0.30 - MW=100 kDa - DF=0.97 - 30 à 50 % des F ramifiés par F en C2 et/ou C4		Patankar et al. (1993)
	model 	- DF=0.75 (25% Gal) - DS=0.40 - MW=40 kDa - branchement		Mulloy et al. (1994)
		- DS=0.90 - DF=1.00		Pereira et al. (1999)
	n 	- MW= 3 kDa - n de 4 à 7 - branchement en C4 possible		Chevolot et al. (2001)
<i>Ascophyllum nodosum</i> (Fucaeeae)	n 	- MW= 3 kDa - n de 4 à 7 - branchement en C4 possible		Chevolot et al. (2001)
		- DS=0,50 - DF=0.95 (5% Xyl) - 50% des F portent un branchement de F en 2		Marais et Joseleau (2001)
		- DS=0.93 - DF=1.00 - branchements suggérés		Perreira et al. (1999)
		- MW=25 kDa - sulfatation en C4 possible - F=47%, S=30%, GU=6% W/W		Daniel et al., (1999)

Tableau 3 : Revue bibliographique des études structurales concernant les fucoidanes.

<p><i>Sargassum stenophyllum</i> (Sargassaceae)</p>		<p>Squelette galactane sulfaté avec ramification de fucoidane.</p> <p>Type I</p> <ul style="list-style-type: none"> - S 21%, GU 10% w/w - DF=52% (23% Gal, 17% Man, 7% Xyl) - Branchement élevé <p>Type II</p> <ul style="list-style-type: none"> - S 29%, GU 2.5% w/w - DF=60% (32% Gal) - branchement faible <ul style="list-style-type: none"> - DS=1.46 - MW=290 kDa 	<ul style="list-style-type: none"> - Fucoidane fractionné par extraction séquentielle - Type I purifié par chromatographie correspondant à : 20% w/w du fucoidane total - Type II purifié par chromatographie correspondant à 24% du fucoidane total 	<p>Duarte et al. (2001)</p>
<p><i>Lytechinus variegatus</i></p>		<ul style="list-style-type: none"> - 	<ul style="list-style-type: none"> - Fucoidane purifié par chromatographie 	<p>Mulloy et al. (1994)</p>
<p><i>Arbatia lixula</i></p>		<ul style="list-style-type: none"> - 	<ul style="list-style-type: none"> - 	<p>Alves et al. (1997)</p>
<p><i>Strongylocentrotus purpuratus</i></p>		<ul style="list-style-type: none"> - type I: 80% des C4 porte un sulfate 	<ul style="list-style-type: none"> - 	<p>Alves et al. (1998)</p>
<p><i>Strongylocentrotus franciscanus</i></p>		<ul style="list-style-type: none"> - 	<ul style="list-style-type: none"> - 	<p>Vilela-Silva et al. (1999)</p>
<p><i>Strongylocentrotus droebachiensis</i></p>		<ul style="list-style-type: none"> - I: DS=0.53 - II: DS=0.80 - homopolymères linéaires 	<ul style="list-style-type: none"> - I et II appartiennent à des molécules distinctes - Fucoidane purifié par chromatographie 	<p>Vilela-Silva et al. (2002)</p>
<p><i>Strongylocentrotus pallidus</i></p>		<ul style="list-style-type: none"> - DS=1.30 - composants minoritaires <5% linéaire 	<ul style="list-style-type: none"> - Fucoidane purifié par chromatographie 	<p>Vilela-Silva et al. (2002)</p>

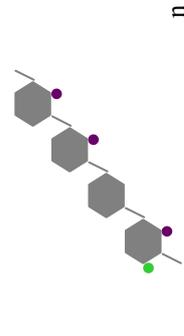
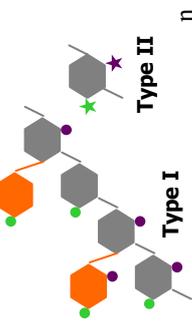
<p><i>Ludwigothurea grisea</i></p>	 <p style="text-align: center;">n</p>	<ul style="list-style-type: none"> - DS=1,10 - MW=30 kDa - linéaire et régulier 	<p>- Fucoïdane purifié par chromatographie</p>	<p>Ribeiro et al. (1994) Mulloy et al. (1994)</p>
<p><i>Stichopus japonicus</i></p>	 <p style="text-align: center;">n</p>	<ul style="list-style-type: none"> - Type I : DS=0.69 DF=1.00 MW=9 kDa - Type II : DS=0.79 DF=1.00 MW=32 kDa 	<p>- Fucoïdane purifié par chromatographie</p>	<p>Kariya et al. (2004)</p>
<p style="text-align: center;">LEGENDE</p> <div style="display: flex; justify-content: space-between;"> <div data-bbox="646 1758 909 1971"> <ul style="list-style-type: none"> ● Sulfate en 4 ● Sulfate en 2 ● Sulfate en 3 ★ Sulfatation aléatoire ■ Acétyle </div> <div data-bbox="662 1288 893 1691"> <ul style="list-style-type: none"> L-Fucose lié en alpha-1,2 L-Fucose lié en alpha-1,4 L-Fucose lié en alpha-1,3 Acide glucuronique </div> <div data-bbox="981 1288 1125 1870"> <p>4 ordres des Phaeophyceae (classe) (Algues brunes)</p> <ul style="list-style-type: none"> Chordariales Laminariales Ectocarpales Fucales </div> <div data-bbox="1013 347 1101 1086"> <p>2 classes des Echinozoa (superclasse) (Echinodermes)</p> <ul style="list-style-type: none"> Echinoidea (oursins) Holothuridea (concombre de mer) </div> </div> <p style="text-align: center;">n : signale le caractère répétitif des motifs présentés F : L-fucopyranose S : ester sulfate A : ester acetyl GU : acide D-glucuronique CH : carbohydrate DF : proportion molaire d'unité fucose (exprimé par unité osidique) DS : proportion molaire de substituant sulfate (exprimé par unité osidique) DA : proportion molaire de substituant acétyle (exprimé par unité osidique) Ci : carbone de position i dans un ose en partant du carbone anomérique</p>				

Tableau 3 : Revue bibliographique des études structurales concernant les fucoïdanes

La nature polydisperse, l'hétérogénéité des substitutions et les hauts poids moléculaires sont autant de freins aux investigations structurales de ce polysaccharide. La production de fractions de bas poids moléculaires par voie enzymatique semble un moyen privilégié pour palier à ces difficultés. Pour autant, la rareté d'enzymes dédiés limite ces approches. Toutefois, ces dernières années les efforts s'intensifient dans cette direction. Ainsi cinq travaux récents font rapport de leur usage (Daniel et al., 1999, 2001 ; Sakai et al., 2004, 2003 A et B), mais seulement deux d'entre eux ont abouti à des analyses structurales fines (Tab. 3).

Sur la base de l'ensemble de ces données quelques caractéristiques générales se dégagent. Le fucoïdane est un polymère de α -L-fucose identifié exclusivement dans l'environnement marin. Le C6 étant dépourvu d'hydroxyle, seule les positions C2, C3 et C4 sont susceptibles d'être engagées dans des liaisons O-glycosidiques ou des substitutions. Ces dernières concernent très majoritairement des esters sulfates (R-O-SO₃⁻) mais également des esters acétates (R-O-CO-CH₃). Dans les chaînes principales, les liaisons O-glycosidiques sont des liaisons α (1,3) et α (1,4). A ce propos, fondée sur plusieurs travaux récents convergents, une hypothèse émerge actuellement qui propose que les fucoïdanes de Fucales soient composés de disaccharides alternant les liaisons α (1,3) et α (1,4) s'opposant aux premiers modèles décrits (Bilan et al., 2004). Ils se distingueraient ainsi de fucoïdanes de laminaires dont ils semblent qu'ils soient principalement composés de α (1,3)-L-fucospyrannose. Par ailleurs, les branchements, fréquemment rapportés chez ces polysaccharides matriciels, se feraient sur le C2 des résidus du squelette via des résidus fucose ou glucuronate. Ces ramifications sont souvent constituées d'une seule unité osidique. D'une manière générale, chez les algues, à de rares exceptions près (Bilan et al., 2002, 2004), l'hétérogénéité des profils de substitutions masque fortement le degré de répétitions. En outre les fucoïdanes étudiés correspondent souvent à des fractions réduites des extraits bruts. Chez les échinodermes la situation est tout autre puisque les fucanes observés sont très homogènes, les répétitions très marquées (constance des profils de substitution), et les chaînes principales entièrement composées de la même liaison, α (1,3) ou α (1,4). Il arrive toutefois que deux types de fucoïdanes cohabitent au sein de la même espèce.

Aucune observation directe de la structure de ce polysaccharide par cristallographie ou microscopie de force atomique n'est à ce jour disponible. Cependant un travail récent (Andrade et al., 2004) propose une étude de l'ultrastructure du fucoïdane d'algue purifié via des approches de microscopie électronique (TEM, FESEM). Il révèle des structures particulières polycristallines de forme polygonale et associées à des sels qui sont quelque peu surprenantes. Ces particules sont néanmoins reconnues par des lectines spécifiques du α -L-fucose.

2.5. Activités biologiques des polysaccharides sulfatés

2.5.1. Activités sur les systèmes animaux

Les carraghénanes et les fucoïdanes peuvent déclencher ou influencer divers processus physiologiques au sein des cellules animales. La plupart des activités biologiques attribuées à ces polysaccharides sulfatés sont étroitement liées à leur profil de sulfatation et leur proximité structurale la famille de polysaccharides sulfatés de la matrice extracellulaire des animaux, les glycosaminoglycanes (ou GAGs, voir § 2.2.). Ces protéoglycanes sont en effet impliqués dans de nombreux mécanismes cruciaux. Il faut donc envisager derrière ces activités biologiques, un effet mimétique du à des analogies chimiques.

L'énorme littérature disponible concernant l'activité biologique de ces polysaccharides sulfatés chez les mammifères, illustre l'intérêt qu'ils suscitent en tant qu'agents thérapeutiques potentiels. Cependant, dans bien des cas, les études disponibles, se limitent à des rapports d'activité et les cibles biologiques ainsi que les mécanismes d'action demeurent inconnus. En outre la plupart d'entre elles emploient les extraits de polysaccharides non purifiés disponibles commercialement. Ce constat est très certainement lié à une connaissance insuffisante de la structure hétérogène et complexe des polysaccharides testés. D'une manière générale, ils sont souvent moins actifs que les GAGs ce qui ne constitue pas forcément un désavantage selon les applications envisagées. Enfin, l'activité est taille dépendante et elle est souvent associée à des molécules de bas poids moléculaires (inférieures à 50 résidus).

2.5.1.1. Propriétés anticoagulantes et antithrombotiques

La coagulation est un processus biologique complexe aboutissant à la formation d'un caillot insoluble composé de plaquettes et de fibrine. Il a pour but d'éviter les pertes massives de sang lors d'une brèche vasculaire. La thrombine est une protéase clé du phénomène puisqu'elle clive le fibrinogène en fibrine. L'héparine fut très tôt employée à des fins thérapeutiques (thrombose, prophylaxie) pour inhiber l'action de la thrombine et la coagulation. Brièvement, elle se lie et active la protéine antithrombine (AT) et à un cofacteur nommé HCII du plasma se qui conduit à au blocage de la thrombine et l'inhibition du facteur Xa via la formation de complexes ternaires (Petitou et al., 2003 ; Bourin et Lindhal, 1993). Cependant, compte tenu des effets secondaires de l'héparine et des sources animales dont elle est extraite, la demande de nouveaux agents anticoagulants a relancé les recherches sur les polysaccharides sulfatés d'origine végétale, à savoir, les carraghénanes et les fucoïdanes.

L'activité anticoagulante des galactanes sulfatés d'algues (Potin et al., 1992 ; Sen et al., 1994 ; Farias et al., 2000 ; Melo et al., 2004) ou d'invertébrés marins (Pereira et al., 2002 A) et des fucoïdanes d'algues (Church et al., 1989 ; Nishino et al., 1999 ; Pereira et al., 1999 ; Chevlot et al., 1999 ; Chaubet et al., 2000) ou d'échinodermes (Mourão et Pereira, 1999 ; Pereira et al., 2002 B) est maintenant bien établie. Il ressort de ces différentes études que : 1) l'interaction du polysaccharide avec les cofacteurs de coagulations et les protéases ciblées sont stéréospécifiques et ne dépendent pas uniquement de la densité de charges ou du contenu en sulfates; 2) les sites d'interactions sont différents de ceux de l'héparine ; 3) tous les galactanes et fucoïdanes sulfatés ne présentent pas le même potentiel.

2.5.1.2. Propriétés diverses

Depuis les quinze dernières années, le spectre des activités biologiques des carraghénanes et fucoïdanes sulfatés s'est passablement élargi. Les activités biologiques rapportées pour les fucoïdanes/carraghénanes reposent essentiellement sur des interférences qu'ils occasionnent dans les mécanismes moléculaires de reconnaissance cellule-cellule. Ces dernières leur permettent de revêtir des fonctions de type blocage/inhibition dans de nombreux processus biologiques. Les études

récentes semblent indiquer que les motifs structuraux impliqués sont différents de ceux actifs contre la coagulation sanguine (Haroun-Bouhedja et al., 2000).

Les propriétés biologiques des fucoïdanes dans les systèmes hétérologues ont été récapitulées de manière exhaustive dans la revue de Berteau et Mulloy (2003). Le fucoïdane se comporte notamment comme un compétiteur des ligand naturels (héparine et héparane sulfate) des sélectines (lectines se trouvant à la surface des leucocytes, plaquettes, et cellules endothéliales vasculaires) ce qui lui confère des propriétés anti-inflammatoires. Ils sont également à l'origine de perturbations de la reconnaissance par les macrophages des récepteurs disposés à la surface de leurs cibles (Hsu et al., 1998). Différentes études révèlent par ailleurs l'action antiprolifératives (Logeart et al., 1997) et antitumorales (Riou et al., 1996) de ces polysaccharides. Enfin, il a été montré qu'ils ont un effet protecteur vis à vis des infections des érythrocytes par *Plasmodium falciparum*, vecteur de la malaria (Xiao et al., 1996), et vis à vis des infections par les rétrovirus tels que l'*Herpes simplex* (Hoshino et al., 1998) ou l'HIV (McClure et al., 1992).

En tant que polysaccharides sulfatés, les carraghénanes présentent des propriétés similaires à l'exception des propriétés anti-inflammatoires. Néanmoins la connaissance de leur structure est mieux établie que pour les fucoïdanes. Ainsi Parish montre en 1987 que les propriétés antitumorales du λ -carraghénane sont plus performantes que celles des ι -carraghénane et κ -carraghénane. Le λ -carraghénane présente également un effet antiprolifératif probablement en inhibant la fixation de certains facteurs de croissances à leur récepteur membranaire dans diverses lignées cellulaires (Hoffman, 1993, 1995). Comme pour les fucoïdanes, plusieurs travaux leur attribuent des propriétés antivirales via un blocage de l'adsorption des virus (HIV, Herpes simplex, cytomegalovirus, virus de l'hépatite A) à la surface des cellules (Girond et al., 1991 ; Witvrouw et De Clercq, 1997 ; Carlucci et al., 1999). Concernant l'HIV, le ι -carraghénane est le plus prometteur vis à vis de la protection des muqueuses et il constitue à ce titre le principe actif de formulation actuellement en phase de test clinique (Elias et al., 1997). Mais les μ -, ν -, et λ -carraghénane non cyclisés seraient plus performants pour lutter contre les infections par les virus de l'herpes (Carlucci et al., 1997).

2.5.1. Activités sur les systèmes végétaux

Il faut bien reconnaître que face à la multitude des tests réalisés sur les modèles animaux, les études concernant l'activité biologique des carraghénanes et des fucoïdanes sur les végétaux supérieurs sont pour le moins restreintes. Cependant, de façon surprenante, le rôle des oligosaccharides dans la signalisation cellulaire et, en particulier, la résistance des plantes à leurs pathogènes est évoqué depuis une vingtaine d'année (Albersheim et Darvill, 1985). Il est désormais bien documenté (Rörhig et al., 1997 ; Potin et al. 1999 ; Vorwerk et al., 2004).

Parmi les rares travaux disponibles, l'étude de l'interaction entre l'algue rouge *Chondrus crispus* et son pathogène endophyte, l'algue verte filamenteuse *Achrochete operculata*, a mis en évidence que la virulence de ce dernier était gouvernée par la perception d'oligocarraghénanes (Bouarab et al., 1999). Ainsi le pathogène incubé en présence d'oligo- λ -carraghénanes (provenant du sporophyte) devient capable d'infecter le gamétophyte normalement résistant. Et à l'inverse, incubé avec des oligo- κ -carraghénanes (provenant du gamétophyte) sa virulence vis à vis du sporophyte habituellement sensible est très atténuée.

Les oligocarraghénanes et les oligofucoïdanes apparaissent comme des molécules ayant le potentiel de déclencher préventivement des réactions de défense chez les plantes terrestres (Potin et al. 1999). L'idée de leur utilisation pour la protection des cultures dans le cadre d'une agriculture raisonnée ne s'appuie néanmoins que sur de rares travaux. Le premier montre la capacité des oligo- κ -carraghénanes à induire chez *Rubus fruticosus* (la mûre) l'expression de laminarinase facteur de défense contre les champignons (Patier et al., 1995). Le second révèle la capacité du λ -carraghénane commercial (contrairement aux ι - et κ -carraghénanes) à éliciter un large panel de gènes impliqués dans les défenses de *Nicotiana tabacum* (tabac) à des concentrations de l'ordre de $10^2 \mu\text{g.ml}^{-1}$ (Mercier et al., 2001). Enfin Klarzynski et col. (2003) mettent clairement en lumière l'effet protecteur d'oligofucanes chez *Nicotiana tabacum* vis à vis du virus de la mosaïque via l'élicitation des réactions de défenses (Fig. 13).



Figure 13: Stimulation des réactions de défenses contre le virus de la mosaïque du tabac chez *Nicotiana tabacum* traitées par des oligofucoïdanes générés par la fucoïdanase de SW5 (adapté de Klarzynski et al., 2003). A) Plante non traitée. B) Plante traitée.

3. Dégradation enzymatique des carraghénanes et des fucoïdanes

3.1. Catabolisme enzymatique des carraghénanes

3.1.1. Recherches d'activités carraghénases

La première activité carraghénolytique a été identifiée au Japon dans l'appareil digestif d'invertébrés marins (Mori, 1939, 1943). Leur origine reste incertaine, pouvant être soit endogène, soit associée à la flore microbienne du tractus digestif.

Par la suite la dégradation des carraghénanes a essentiellement été caractérisée à partir de la décomposition microbienne des algues rouges consécutive à leur échouage. La flore bactérienne saprophytique joue un rôle écologique important dans le recyclage de la biomasse carbonée en milieu côtier. La première espèce bactérienne carraghénolytique isolée fut *Pseudoalteromonas carrageenovora* (anciennement nommée *Pseudomonas*, puis *Alteromonas carrageenovora*), une γ -Protéobactérie de l'ordre des Alteromodaceae (Yaphe and Baxter, 1955), sécrétant une κ -carraghénase (Weigl and Yaphe, 1966 ; McLean and Williamson, 1979 A ; Potin et al., 1995), et une λ -carraghénase (Johnston and McCandless, 1973).

D'autres bactéries marines carraghénolytiques ont été isolées depuis lors, notamment *Alteromonas fortis* (isolée de la surface de *Euchema spinosum*, Rhodophytes) qui produit une ι -carraghénase (Bellion et al., 1982 ; Greer et Yaphe, 1984). Il est regrettable que beaucoup d'entre elles n'aient pas reçu l'intérêt nécessaire à leur classification rigoureuse. Toutefois, *Zobellia galactinovorans* (anciennement *Cytophaga drobachiensis* souche Dsij), une Flavobacteriaceae, constitue un bon contre exemple (Barbeyron et al., 2001). Cette bactérie, isolée de l'algue rouge *Delesseria sanguinea*, sécrète une κ - et une ι -carraghénase (Potin et al., 1991, 1992 ; Barbeyron et al., 1998, 2000). L'ensemble des souches bactériennes carraghénolytiques connues sont rassemblées dans le tableau 4.

Récemment, une activité carraghénolytique a été démontrée chez un eucaryote, l'algue verte *Achrochete operculata* (Bouarab et al., 1999) parasite du sporophyte de

Chondrus crispus. Les extraits bruts de cette algue verte occasionnent une hydrolyse significative du carraghénane de ce sporophyte (λ -carraghénane). Par conséquent ce pathogène endophyte doit probablement exprimer de manière constitutive une λ -carraghénase. Par ailleurs, quand il est incubé avec des oligo- λ -carraghénanes ou des oligo- ι -carraghénanes, les extraits bruts deviennent actifs contre les carraghénanes du gamétophyte de *C. crispus*. L'expression d'une κ - et une ι -carraghénase serait donc inductible. Pour cette algue verte pathogène les carraghénases sont des facteurs de virulence permettant de dégrader les parois cellulaires de son hôte.

Le plus grand nombre de κ - et ι -carraghénases identifiées en comparaison d'autres carraghénases, s'explique au moins en partie par la facilité de leur criblage. En effet, similairement aux agarases, leur substrat gélifie. Il devient donc possible de s'en servir comme support d'étalement pour les isollements bactériens. Les souches carraghénolytiques sont alors révélées par l'apparition plus ou moins rapide d'un cratère de dégradation du gel autour des colonies.

La disponibilité des séquences de κ - et ι -carraghénases (Barbeyron et al., 1994, 1998, 2000), ajoutée à la multiplication des génomes bactériens entièrement séquencés, permet la prospection d'orthologues à grande échelle. Il faut néanmoins reconnaître que les génomes de bactéries marines ne sont pas légions. Malgré tout, à ce jour, aucune ι -carraghénase supplémentaire n'a été détectée par homologie de séquence. Les deux ι -carraghénases de *A. fortis* et *Z. galactanivorans* demeurent donc les uniques membres de la famille GH82 ; elles partagent 43% d'identité de séquences. En ce qui concerne les κ -carraghénases, seule une protéine de *Rhodospirellula baltica* (Planctomycetes) présente un score d'identité probant (46%) avec les deux uniques séquences de κ -carraghénases (provenant de *P. carrageenovora* et *Z. galactanivorans*, 45% d'identité de séquence entre leur domaine catalytique). Autrement dit, à l'heure actuelle, κ - et ι -carraghénases, appartenant à deux familles différentes de glycosidases, ne sont respectivement représentées que par 3 et 2 séquences. Afin de nuancer ce propos, on peut souligner que malgré l'absence d'orthologue dans le génome de *Microbulbifer degradans* 2-40, cette bactérie dégrade tout de même les carraghénanes (Gonzales et Weiner, 2000).

En guise de conclusion, j'aimerais faire remarquer que, dans les différentes sou-

Tableau 4 : Récapitulatif des différentes carraghénane hydrolases identifiées à ce jour.

Bactéries	Enzymes	Tailles (aa, kDa) *théorique °sauvage	Localisation, purification, expression	Produits limites <u>Substrat minimum</u>	Biochimie (liaison, mécanisme, mode d'action)	Identification (famille, structure)	Références		
Alteromonadaceae									
Souche 2	κ-carraghénase	nd	Inductible	DP1 et DP2 <u>DP3</u>	β(1,4)	nd	Bellion et al., 1982		
Souche 3/4	ι-carraghénase						Bellion et al., 1982		
	κ-carraghénase						Bellion et al., 1982		
	ι-carraghénase						Bellion et al., 1982		
	κ-carraghénase						Bellion et al., 1982		
Souche 7	κ-carraghénase						Bellion et al., 1982		
Souche 5	ι-carraghénase	59°	Extracellulaire Complète Inductible	DP2 et DP3 <u>DP4</u>	β(1,4)	GH82 ?	Greer, 1984		
	κ-carraghénase	36,5°					DP1 et DP2 <u>DP3</u>	GH16 ?	Greer, 1984
<i>Alteromonas fortis</i> (ancienne souche 1)	ι-carraghénase	<u>491</u> *, <u>465</u> 55*, 52°	Extracellulaire Complète Inductible	DP2 et DP3 <u>DP4</u>	β(1,4) Inverting Endo-processive	GH82 Hélice β	Greer and Yaphe, 1984		
<i>Pseudoalteromonas carrageenovora</i>	κ-carraghénase	<u>397</u> *, <u>295</u> 44*, 32°					β(1,4) Retaining Endo-processive	GH16 Sandwich β	McLean et williamson (1979 A) Potin et al., 1995 Michel et al., 2001 A
	λ-carraghénase	98°							Johnston and McCandless, 1973 Greer, 1984
Flavobacteriaceae									
<i>Zobellia galactovorans</i>	κ-carraghénase	<u>545</u> *, <u>281</u> 61*, 33°	Extracellulaire Complète Inductible	DP2 et DP3 <u>DP4</u>	β(1,4) Retaining Endo-processive	GH16 Sandwich β	Potin et al., 1991 Barbeyron et al., 1998		
	ι-carraghénase	<u>491</u> *, <u>472</u> 53*, 51°	Extracellulaire Partielle Inductible				β(1,4) Inverting Endo-processive	GH82 Hélice β	Barbeyron et al., 2000
Planctomycetes									
<i>Rhodospirillum baltica</i>	κ-carraghénase putative					GH16 Sandwich β			

ches bactériennes carraghénolytiques isolées, les présences des κ - et ι -carraghénase ne sont pas systématiquement conjointes (Tab. 4). Et d'autre part, leur présence dans des organismes aussi distants phylogénétiquement que les Flavobacteriaceae, les Alteromonadaceae ou les Planctomycetes, suggère fortement des évènements de transferts horizontaux (cf. Chapitre 3 §2.4.). Cette hypothèse est appuyée par la proximité écologique de ces souches.

3.1.2. Les scénarios de dégradation des carraghénanes

Des études biochimiques connues à ce jour (Tab. 4), il ressort deux propriétés générales : i) toutes les carraghénases identifiées et caractérisées clivent la liaison $\beta(1,4)$, engendrant des produits de la série néocarrabiose ; ii) elles affichent une stricte spécificité de substrat, tout du moins si l'on considère uniquement les trois grands types κ -, ι -, et λ -carraghénane.

Les carraghénases actuellement connues ne peuvent donc pas seules conduire à la décomposition complète des carraghénanes en D-galactose. Le système carraghénolytique le mieux caractérisé est celui de *P. carrageenovora* vis à vis du κ -carraghénane. Cette bactérie secrète une κ -carraghénase (McLean et Williamson, 1979 A) dont les produits majeurs sont le κ -néocarrabiose-sulfate (DA-G4S) κ -néocarratétraose-disulfate (DA-G4S-DA-G4S). La décomposition de ces produits est assurée par deux autres enzymes de *P. carrageenovora*.

Une glycosulfatase élimine le sulfate du κ -néocarrabiose-sulfate et un seul sulfate du κ -néocarratétraose-disulfate, pour donner respectivement le DA-G et le DA-G-DA-G4S (McLean et Williamson, 1979 B). Ce dernier est alors pris en charge par une néocarratétraose-sulfate hydrolase spécifique de la liaison $\beta(1,4)$ qui génère du DA-G et du DA-G4S reconnu par la glycosulfatase (McLean et Williamson, 1981). Le produit final de cette chaîne catabolique est donc le DA-G. Les enzymes intervenant de sa transformation en D-galactose ne sont pas encore identifiés. La κ -carraghénase de *Z. galactanivorans* a pour produit limite le néo- κ -carratétraose et le néo- κ -carrahéxaose (Potin et al., 1991). La présence de ce degré de polymérisation supplémentaire n'est pas rédhibitoire pour la transposition de ce scénario de dégradation.

On ignore si la dégradation des autres carraghénanes suit des schémas similaires. La présence d'un sulfate supplémentaire sur le ι -carraghénane (DA2S-G4S) augmente probablement les étapes de dégradation. Toutefois, chez *P. carrageenovora*, où aucune ι -carraghénase n'est détectée, une seule sulfatase pourrait permettre de le transformer en κ -carraghénane, et donc de rattraper le scénario précédent. Cependant les sulfatases connues à ce jour ont toutes des substrats de taille relativement modeste ce qui pourrait justifier la nécessité d'au moins une ι -carraghénase. Par ailleurs, les ι -carraghénases de *A. fortis* et *Z. galactanivorans* (Barbeyron et al., 2000) ont toutes deux pour produits limites les néo- ι -carratétraose et néo- ι -carrahéxaose.

La dégradation du λ -carraghénane doit probablement suivre un cheminement différent du fait de l'absence de pont 3,6-anhydre et d'une plus forte hétérogénéité dans le profil de sulfatation. Eventuellement, les sulfatases requises pour le sulfate en 2 pourraient également accepter le ι -carraghénane, alors que les sulfatases requises pour le sulfate en 6 pourraient également accepter les précurseurs ν - et μ -carraghénane. Il ne serait pas surprenant que la λ -carraghénase de *P. carrageenovora* s'accommode de certaines variations du profil de sulfatation afin de limiter le recours à un trop grand nombre de glycosides hydrolases plus spécifiques.

3.1.2. Les rapports structure/fonction des carraghénases

Récemment la résolution des structures de la κ -carraghénase de *P. carrageenovora* (Michel et al., 2001 A) et de la ι -carraghénase de *A. fortis* (Michel et al., 2001 B, 2003) a donné accès à l'intimité de ces deux enzymes (Fig. 14 et 15). Conformément aux nombreux enzymes de la famille 16, la κ -carraghénase adopte un repliement en β -sandwich à deux feuillets β antiparallèles de six et sept brins. Cependant elle est son unique représentant dont la topologie tunnel du site actif ait été démontrée (Fig. 14). Ce type de topologie est perçu comme une adaptation à la dégradation endo-processive d'un substrat solide ce qui vient confirmer l'observation biochimique. Elle permet en effet de maintenir la chaîne polysaccharidique entre deux coupures, après l'attaque initiale. Dans ce cas, le tunnel maintient plus le côté non réducteur de la chaîne polysaccharidique (sous-sites -5 à -1) que l'aglycone (sous-sites

+1 et +2). Cette asymétrie, explique le fait que cette κ -carraghénase ait pour substrat minimale un hexasaccharide et pour produits limites un néocarrabiose et un néocarratétraose. L'enzyme progresserait de l'extrémité réductrice vers l'extrémité non réductrice par saut de quatre résidus conduisant à la libération majoritaire de tétraose (Fig. 14). Enfin, le tunnel aurait également pour fonction d'éviter la réassociation de la chaîne de κ -carraghénane en double hélice. L'interaction avec le substrat implique d'une part des interactions électrostatiques entre les résidus basiques de l'enzyme et les sulfates des résidus G4S, et d'autre part, un environnement hydrophobe spécifique pour les résidus DA. Le sous-site -1 présente une fixation ferme du monosaccharide au site de clivage. L'incapacité des κ -carraghénanes à dégrader le ι -carraghénane s'explique par le volume restreint de la cavité et par une interaction électrostatique d'une histidine en -1 avec l'hydroxyle en C2 du résidu DA. Le sulfate que le ι -carraghénane possède à cette position provoque une gêne stérique et le mauvais positionnement de la liaison glycosidique à hydrolyser.

La structure tridimensionnelle de la ι -carraghénase est de type hélice β droite à dix tours comprenant quatre feuillets parallèles (Fig. 15). Ce repliement aisément identifiable mais peu répandu est très souvent associé à des enzymes interagissant avec des polysaccharides très acides (pectine, chondroïtine sulfate, alginate...) de type hydrolases, lyases, ou épimérasés. Ce domaine forme une grande crevasse encadrée par deux autres domaines C-terminaux, A et B, formant le site de fixation du substrat (Fig. 15). En l'absence d'oligosaccharide dans cette gorge le domaine A présente une forte flexibilité conformationnelle. Sa structure a été obtenue en diffusant dans le cristal des oligo- ι -carraghénanes (Michel et al ; 2003). Il s'agit d'un module α/β de fixation des polyanions que l'on retrouve dans les protéines se fixant aux acides nucléiques. La fixation d'un oligo- ι -carratétraose s'accompagne du basculement du domaine A vers la gorge ce qui a pour effet de créer une cavité de type tunnel où le substrat se trouve maintenu (Fig. 15). Ainsi la ι -carraghénase adopte deux conformations, un tunnel ouvert pour l'attaque initiale, et un tunnel fermé pour la dégradation en mode processif.

Contrairement au κ -carraghénane, le caractère hydrophobe du résidu GA est atténué par la présence d'un sulfate en C2. La gorge de la ι -carraghénase est quasiment

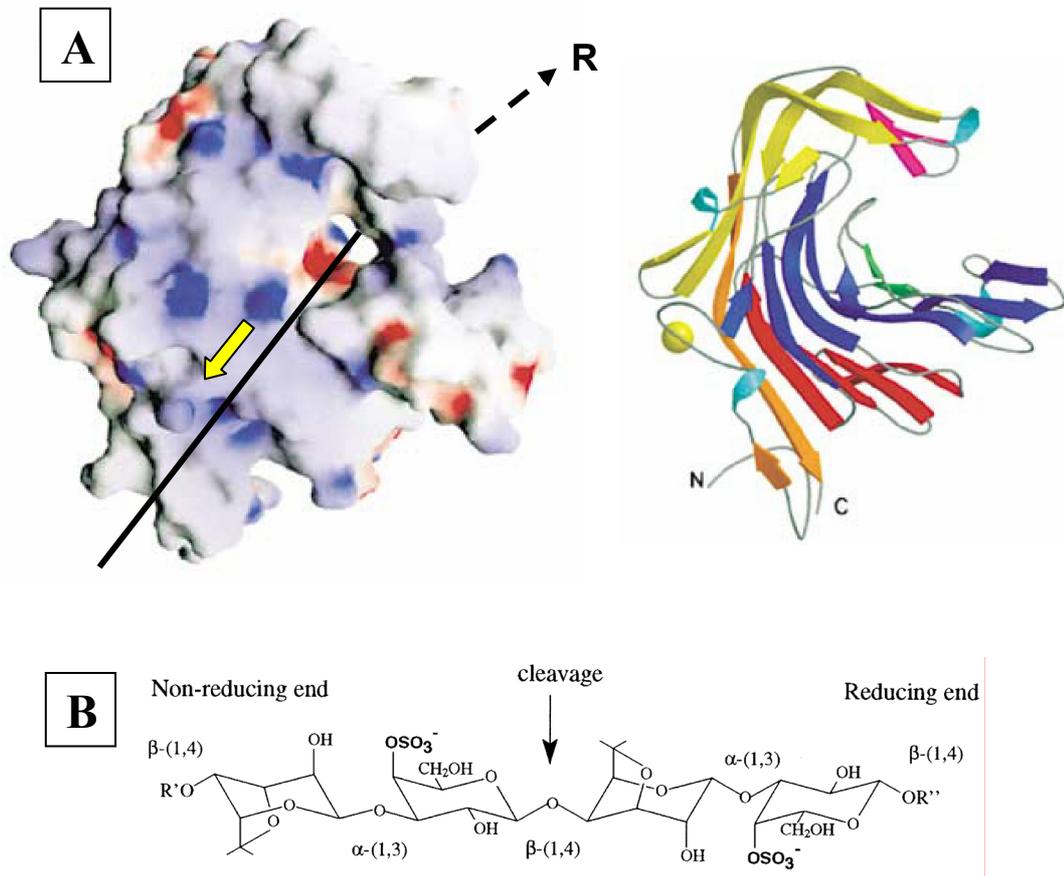


Figure 14 : La κ -carraghénase de *P. carrageenovora*, une glycosidase endo-processive adaptée à la dégradation de fibres polyanioniques en phase solide (adapté de Michel et al., 2001).

A) La structure de la κ -carraghénase présente une topologie en tunnel. La surface moléculaire est colorée selon le potentiel électrostatique, rouge -, bleu, + (R, extrémité réductrice).

B) Site de clivage du κ -carraghénane à la liaison $\beta(1,4)$.

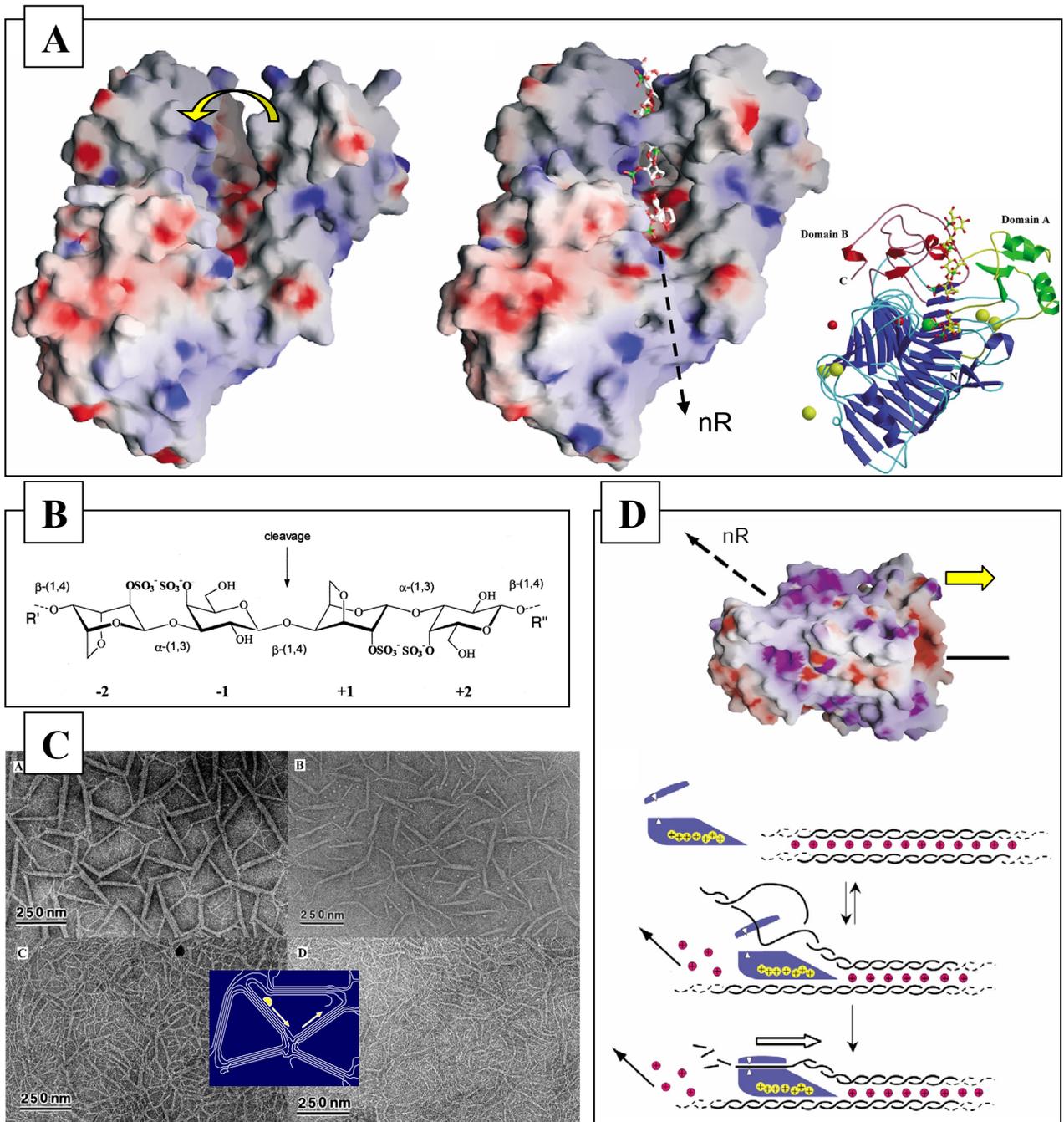


Figure 15 : La ι -carraghénase de *A. fortis*, une glycosidase endo-processive adaptée à la dégradation de fibres polyanioniques en phase solide (adapté de Michel et al., 2003).

A) La structure de la ι -carraghénase a une topologie ouverte qui se referme sur le substrat lors de l'attaque initiale via le basculement d'un domaine qui ferme le tunnel. La surface moléculaire est colorée selon le potentiel électrostatique, rouge -, bleu, + (nR, non réducteur).

B) Site de clivage du ι -carraghénane à la liaison $\beta(1,4)$.

C) Mise en évidence du comportement processif en phase solide par micrographie électronique de gel de ι -carraghénane. A, contrôle ; B, hydrolyse acide de 4 h dégradant les régions amorphes mais pas les fibres compactes, le réseau du gel est brisé ; C, après incubation de 2,5 h avec la ι -carraghénase, les fibres sont affinées mais le réseau est préservé ; D, accentuation du phénomène après 10 h d'incubation avec l'enzyme.

D) Modélisation de la dissociation des fibres polyanioniques de ι -carraghénane. L'enzyme processif progresse vers les extrémités réductrices en déplaçant les cations Ca^{2+} via son dos polycationique. L'ouverture de la double hélice n'est pas élucidée.

dépourvue de résidus aromatiques alors que plusieurs résidus basiques y sont conservés. Un néocarratétraose a donc été observé dans la gorge dans les sous-sites +1 (DA2S), +2 (G4S), +3 (DA2S) et +4 (G4S), et un néocarrabiose dans les sous-sites -4 (DA2S) et -3 (G4S). La reconnaissance de ces oligosaccharides par l'enzyme est dominée par les interactions électrostatiques entre sulfates et résidus basiques qui n'excluent toutefois pas des interactions de type « stacking » hydrophobe entre quelques résidus aromatique et des cycles du sucre. Le glissement de la chaîne polysaccharidique dans le tunnel pourrait être facilité par une répartition homogène des résidus basiques le long de la gorge, assurant une forte affinité globale dépourvue d'une trop forte sélectivité. Ce glissement doit avoir pour période la taille du produit limite, à savoir, un tétraose. Pour maintenir le substrat dans le tunnel entre deux événements catalytiques, il faut que la protéine se déplace de l'extrémité non réductrice de la chaîne polysaccharidique, vers l'extrémité réductrice, de façon à remplir les sous-sites -1 à -4 préalablement vidé du produit partant (Fig. 15).

La ι -carraghénase dégrade aussi bien les zones amorphes d'un gel de ι -carraghénane que les fibres cristallines (Fig. 15). Comme la gorge catalytique ne peut lier qu'une seule chaîne, la dégradation des fibres passe nécessairement par une déstabilisation des agrégats de doubles hélices (maintenues par des ions calcium) et l'ouverture d'une double hélice. Il s'avère que la face externe de la protéine opposée au site actif est fortement basique. Il est proposé que cette surface polycationique rentre en compétition avec les ions calcium, de manière à réaliser des interactions électrostatiques coopératives avec le polymère, participant ainsi à la dissociation des fibres par déplacement des Ca^{2+} entre les doubles hélices (Fig. 15). Le mécanisme d'ouverture de la double hélice demeure quant à lui bien mystérieux.

L'incapacité de la ι -carraghénase à hydrolyser le κ -carraghénane s'explique par une affinité moindre due principalement à sa densité de charge deux fois plus faible et à l'absence de sulfate en C2 du résidu DA. D'une part l'équilibre association-dissociation des hélices de κ -carraghénane/contre ions ne doit pas être efficacement déplacé. D'autre part, le positionnement productif de la liaison glycosidique cible ne doit pas être satisfaisant.

3.1.3. La λ -carraghénase de *P. carrageenovora*

La littérature au sujet de cette activité est pour le moins restreinte. Weigl et Yaphe (1966) sont les premiers à en faire part chez la bactérie marine *P. carrageenovora*. Poursuivant ces travaux, Johnston et McCandless (1973) évoquent l'existence, dans le surnageant de culture de *P. carrageenovora*, d'un complexe enzymatique hydrolysant le carraghénane soluble au KCl extrait de *Chondrus crispus*. Ils ont purifié partiellement trois fractions enzymatiques actives sur ce substrat faiblement caractérisé. Certaines réduisaient fortement la viscosité alors que d'autres généraient beaucoup d'extrémités réductrices. Il est probable que l'hétérogénéité de leur substrat ait occasionné une sensibilité à différents enzymes.

L'hypothèse de ce complexe enzymatique fut mise à bas par Greer (1984) qui purifia à homogénéité une protéine extracellulaire unique de 98 kDa. Malgré la capacité de l'enzyme à réduire drastiquement la taille du λ -carraghénane, l'auteur ne mesurait qu'une faible activité en sucres réducteurs et c'est donc par mesure de la décroissance de viscosité qu'il suivit la purification de cet enzyme. Il décrit l'enzyme comme ayant un pH optimum de 6,5, une température optimale de 30°C, une dépendance aux cations bivalents, et comme étant instable en l'absence d'une force ionique suffisante (NaCl 2 M). Il montra la spécificité de la protéine pour le λ -carraghénane, sur des extraits de trois sporophytes de Gigartinales (*G. canaliculata*, *G. pistillata*, et *C. crispus*), suggérant que le sulfate en C2 du résidu G était prépondérant dans la reconnaissance. Il distingua deux produits limites largement majoritaires sans toutefois les avoir caractérisés. Il suggéra par ailleurs que l'enzyme puisse être associé à la paroi de la bactérie (comme la ι -carraghénase de *A. fortis*). Reprenant ces travaux au laboratoire, Potin (1992) ne parvint pas, sur la base d'un protocole différent, à reproduire la purification de cette protéine.

3.2. Catabolisme enzymatique des fucoïdanes

Si l'on considère le fucoïdane comme un polymère composé majoritairement de α -L-fucose présentant de nombreuses sulfatations, trois grands types d'activité enzymatique peuvent théoriquement participer à sa décomposition en fucose

monomérique métabolisable. En premier lieu, une ou des activités sulfatases (ou sulfohydrolases ou encore sulfo-estérases) sont indispensables au « nettoyage » de ce polymère. Une activité α -L-fucosidase (ou fucose-hydrolase) sera également indispensable à la production de fucose monomérique. En outre, selon les scénarios de dégradation, une activité fucoïdanase de type hydrolase peut être nécessaire (mais pas indispensable) à la production d'oligosaccharides sulfatés, se plaçant alors en amont et/ou en aval d'activité sulfatase. Enfin, les branchements constatés dans ce polysaccharide requièrent probablement des glycosidases spécifiques.

Ces remarques très générales traduisent le flou qui règne autour de ce catabolisme (la biosynthèse est encore plus méconnue), tant les connaissances sont réduites sur les enzymes qui y participent. Bien qu'initiées il y a près de 50 ans (Yaphe et Morgan, 1959) avec des motivations pourtant très contemporaines, ces recherches ont visiblement du mal à aboutir, ce qui n'est pas sans rapport avec la complexité du fucoïdane. Cette hétérogénéité suggère que d'autres activités enzymatiques puissent être requises. La présence d'acides uroniques nécessite l'implication de lyases alors que les substitutions de type acétyles nécessitent probablement des activités estérases. Pour répondre à cette hétérogénéité et au masquage des répétitions structurales, deux grandes stratégies de dégradation s'opposent. Soit les enzymes qui y participent tolèrent les fluctuations structurales (notamment les profils de substitutions) via une spécificité de reconnaissance peu prononcée, soit au contraire elle nécessite un grand nombre d'enzymes jouant chacun un rôle très précis. Mais avant d'aborder cette question, un panorama des travaux actuellement disponibles me semble nécessaire.

Les recherches d'enzymes intervenant dans la dégradation des fucoïdanes ont été réalisées principalement dans des systèmes bactériens, et, dans une moindre mesure, dans les systèmes digestifs de mollusques marins tels que l'ormeau (*Haliotis* sp.), la coquille St jacques (*Pecten* sp.) ou le bigorneau (*Littorina* sp.). Si la plupart de ces études se cantonnent à des criblages d'activités, certaines d'entre elles sont néanmoins assez précises. La qualité et l'origine du fucoïdane employé ainsi que le degré de pureté des extraits enzymatiques limitent souvent la portée de ces travaux et entravent sérieusement leur comparaison.

3.2.1. Recherches d'activités fucoïdanases

Les deux bactéries marines *Pseudomonas atlantica* et *Pseudomonas carrageenovora*, connues pour leur capacité à dégrader les carraghénanes, ont été les deux premiers organismes où les activités enzymatiques de dégradation du fucoïdane ont été recherchées (Yaphe et Morgan, 1959). Légèrement plus tard, plusieurs espèces d'ormeau furent montrés comme contenant une endofucoïdanase capable de produire des oligosaccharides fucosylés (Thanassi et Nakada, 1967). La première recherche de grande envergure a porté sur le criblage de 396 souches bactériennes provenant de sédiments marins dont 32 présentaient une activité fucoïdanolytique avec la particularité de tous appartenir au genre *Vibrio* (Morigana et al., 1981). Dix ans plus tard, deux articles rapportent la purification d'une fucoïdanase produite par la souche *Vibrio* sp. N-5 (Furukawa et al., 1992 A et B). Plus récemment, une activité similaire a été obtenue à partir des deux bivalves *Patinopecten yessoensis* (Kitamura et al., 1992) et *Pecten maximus* (Daniel et al., 1999). Ces derniers auteurs ont confirmé la présence d'enzymes dégradant le fucoïdane dans la glande digestive de *Pecten maximus*. Ils ont ainsi montré la libération de L-fucose monomérique révélant l'activité d'une fucosidase. En outre, la forte décroissance du poids moléculaire apparent (passage de 25000 à 3000) trahissait la présence d'une endohydrolase détectable également par les sucres réducteurs produits. Cependant la nature du motif structural reconnu (résidus, liaisons, substitutions) n'y est pas rapportée. La même équipe a également étudié finement l'activité d'une sulfatase et de la fucosidase provenant du même organisme (voir ci-dessous ; Daniel et al., 2001 ; Berteau et al., 2002).

Une équipe russe dirigée par Zvyagintseva (laboratory of enzyme chemistry of the pacific institute of Bioorganic chemistry, Vladivostok) est à l'origine de nombreux cribles d'activités fucoïdanolytiques de provenance principalement bactérienne (Bakunina, et al., 2000, 2002 ; Alexeeva et al., 2002 ; Ivanova et al., 2003) mais aussi de mollusques (Kusaykin et al., 2003). Un crible de grande ampleur est rapporté concernant 25 souches bactériennes épiphytes de *Fucus evanescens* et *Chorda filum*, 53 souches isolées du concombre de mer *Apostichopus japonicus* (Bakunina, et al., 2000). Trois différentes sources de fucoïdanes ont été extraites de *F. evanescens*, *Laminaria cichorioides* et *L. japonica*, dont le pourcentage de fucose est

respectivement 87, 82, et 42% et le taux molaire de sulfate de 0,8, 1,7 et 0,9. Enfin l'activité fucoïdane hydrolase des lysats cellulaires a été suivie par dosage des sucres réducteurs. Parmi les 78 isolats, plus de 90% des Bacteroidetes (15 isolats, *Cytophaga/Flexibacter/Flavobacterium*) sont actifs et environ la moitié est actif sur au moins deux substrats; 75% des γ -Protéobactéries (43 isolats, *Alteromonas/Pseudoalteromonas/Pseudomonas/Vibrio*) sont actifs, et presque tous sont actifs sur au moins deux substrats. Ces résultats illustrent l'importance de ces deux phylums bactériens dans le catabolisme des polysaccharides marins et des fucoïdanes en particulier. Par ailleurs, les fucoïdanes de *L. cichorioides* et *L. japonica* génèrent les plus fortes activités ce qui chez le dernier peut être due à sa richesse en galactose (40%). En effet, ce type d'études est limité par le risque de mesurer l'activité d'autres enzymes. Une analyse plus restreinte sur l'espèce *Pseudoalteromonas citrea* a montré que parmi 12 souches (isolées de pleine eau, de mollusques, d'ascidies, d'échinodermes, d'éponges et d'algues brunes) seules les trois souches isolées d'algues brunes et d'holothurie présentent une activité fucoïdanolytique intracellulaire (Bakunina et al., 2002). Cela indique donc que la diversité des polysaccharidases contenues dans une même espèce de *Pseudoalteromonas* est étroitement liée à son habitat. La même équipe a réalisé à des criblages identiques de glycosidases diverses, dont des fucoïdane hydrolases, sur des mollusques variés. L'un d'entre eux, le gastéropode *Littorina kurila*, a fait l'objet de recherches plus approfondies (Kusaykin et al., 2003). Elles ont notamment porté sur la purification de fucoïdanases, sans toutefois y parvenir de manière convaincante.

L'entreprise japonaise Takara Bio Inc. fait partie des rares groupes travaillant à l'isolement d'enzymes fucoïdanolytiques. Dans ce contexte, elle est à l'origine de travaux très intéressants dont l'objectif est la production d'oligofucoïdanes bioactifs. Leur stratégie a également consisté à rechercher ce type d'activité dans des isolats bactériens. Leur première publication (Sakai et al., 2002) fait part de la sélection (parmi 120 souches) d'une Flavobacteriaceae, *Fucobacter marina* (SA-0082) sur sa capacité à métaboliser les fucoïdanes extrait de trois Laminariales (*Kjellmaniella crassifolia*, *Undaria pinnatifida*, et *Lessonia nigrescens*). Cette souche ne métabolise pas ou peu le fucoïdane extrait des Fucales *Fucus vesiculosus* et *Ascophyllum*

nodosum. Approfondissant ces travaux (Sakai, et al., 2003 B), les enzymes du surnageant de culture ont été employés (après concentration par précipitation au sulfate d'ammonium) pour digérer du fucoïdane extrait de *K. crassifolia* (taux molaire de L-fucose 0,84, de D-mannose 0,1, de D-galactose 0,05, de D-glucuronate 0,07). L'implication de lyases a été montrée par le développement durant la digestion d'absorbance à 232nm. Trois trisaccharides ont été purifiés à l'issue de la digestion (Fig. 16). Ils sont effectivement les produits d'au moins une lyase, mais ils appartiennent à la famille des fucoglucuronomannanes sulfatés. Ils résultent de la rupture de la liaison $\alpha(1,4)$ dans le polymère suivant $[\rightarrow 4\text{-}\beta\text{-D-glucuronyl-1}\rightarrow 2\text{-}\alpha\text{-D-mannosyl-6,R-3,}(\alpha\text{-L-fucosyl-2,R-3,R-4,R)-1}\rightarrow]_n$, avec R= H ou SO₃H, structure assez éloignée du fucoïdane.

Deux lyases de cette souche (nommées Fdl1 et Fdl2) donnant des produits similaires, ont fait l'objet d'un brevet exposant leur clonage, séquençage, surexpression et produits (Takayama et al., 2002). Ce brevet présente également les clonage, séquençage, et surexpression de deux fucoïdanes hydrolases isolées d'une souche d'*Alteromonas* (SN-1009) dont les produits ne sont pas révélés. Les gènes de ces deux fucoïdanases, nommées Fda1 et 2, sont contigus et très similaire suggérant une duplication récente. Mon analyse de leur séquence révèle une très forte similitude (70 % d'identité), mais Fda2 possède une extension N-terminale de 62 résidus riche en proline qui fait penser au bras espaceur que l'on trouve chez des enzymes ancrés dans les membranes plasmiques. Leur activité sur le fucoïdane a été suivi par chute de la taille moyenne du polymère et non par dosages des sucres réducteurs. Ce brevet évoque une préférence de ces fucoïdanases pour le fucoïdane pauvre en acide uronique contrairement aux deux lyases décrites ci-dessus. Une publication récente (Sakai et al., 2004) fait part de la caractérisation de cette souche SN-1009 et de la purification d'une fucoïdane endohydrolase extracellulaire. Sans qu'il en soit fait mention dans cet article, tout porte à croire qu'il s'agisse de Fda1. La purification n'est pas menée à l'homogénéité mais l'espèce protéique majoritaire est estimée à 100 kDa ce qui est compatible avec les deux enzymes cités dans le brevet. L'activité de cet enzyme a également été suivi par chute de la taille moyenne du polymère et non par dosages des sucres réducteurs. Son pH optimum se situe entre 6,5 et 8,0 et sa température optimale

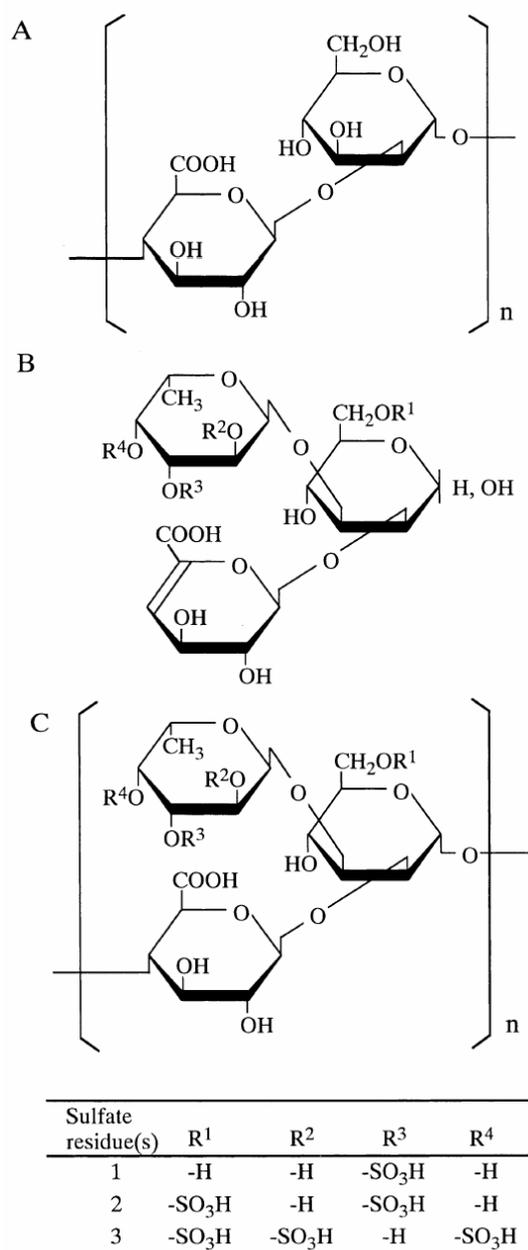


Figure 16 : Dégradation des glucuronomannanes (A) par les fucoïdanes lyases Fdl1 et Fdl2 (*Fucobacter marina*) en trois produits majoritaires (B), suggérant l'existence de chaînes de fucoglucuronomannanes sulfatés (C) dans les extraits de fucoïdane de la Laminariale *K. crassifolia* (extrait de Sakai et al., 2003 B).

entre 30 et 35°C. Son activité est activée par la présence de Ca^{2+} jusqu'à 30 mM et inhibé par Cu^{2+} et Zn^{2+} . Enfin, cette hydrolase ne serait active que sur le fucoïdane hautement sulfaté extrait de Laminariales et pas sur celui de Fucales. Compte tenu des différences structurales des deux fucoïdanes, j'en déduis que cette fucoïdanase pourrait couper la liaison $\alpha(1,3)$ mais pas en alternance avec la liaison $\alpha(1,4)$. Aucune structure de produits n'est révélée pour répondre à cette hypothèse.

Le même groupe a isolé d'un holothurie la souche bactérienne SU-1234, de l'espèce *Fucophylus fucoidanolyticus* (phylum des Verrucomicrobia), sur la base de son pouvoir de dégradation de fucoïdanes indépendamment de leur provenance de Laminariales ou de Fucales (Sakai et al., 2003 C). Les enzymes intracellulaires ont été employés pour générer des oligosaccharides à partir du fucoïdane de *Cladosiphon okamuranus* (Sakai et al., 2003 A). Ceux d'entre eux qui ont été purifiés et caractérisés sont constitués du motif pentasaccharidique suivant $[\rightarrow 3\text{-}\alpha\text{-L-fucosyl-1}\rightarrow 3\text{-}\alpha\text{-L-fucosyl-4,sulfate-1}\rightarrow 3\text{-}\alpha\text{-L-fucosyl-4,sulfate-1}\rightarrow 3\text{-}\alpha\text{-L-fucosyl-2,R-1}\rightarrow]_n$ répété de 1 à 4 fois avec $\text{R}=\text{H}$ ou $\alpha\text{-D-glucuronyl}$ dans 75% des cas. La liaison $\alpha(1,3)$ hydrolysée sépare deux résidus fucoses non sulfatés. La fucoïdane hydrolase qui la coupe peut donc être qualifiée d'endofucosidase. La présence de substituants acétyles est affirmée ainsi que l'enzyme qui les supprime (deacétylase ou acétyl estérase), justifiant ainsi leur non positionnement. L'identification de variants désulfatés des oligofucoïdanes décrits ci-dessus suggère également la présence de sulfatases intracellulaires spécifiques. Enfin, la détection de glucuronate libre après la digestion trahie l'existence d'une $\alpha\text{-D-glucuronidase}$. Ce système bactérien est donc très prometteur mais aucun enzyme n'a encore été isolé.

3.2.2. Les $\alpha\text{-L-fucosidases}$

Toutes les $\alpha\text{-L-fucosidases}$ répertoriées à ce jour appartiennent à la famille 29 des glycosides hydrolases. Ces enzymes ont reçu beaucoup d'attention du fait du rôle central joué par les oligosaccharides glycoconjugués contenant du fucose dans de nombreux processus biologiques des mammifères. Elles catalysent la libération de $\alpha\text{-L-fucose}$ à l'extrémité non réductrice d'un polysaccharide qu'il soit lié en $\alpha(1,2)$,

$\alpha(1,3)$, $\alpha(1,4)$ ou $\alpha(1,6)$ et sont donc des exoglycosidases. Elles partagent vraisemblablement le même mécanisme moléculaire d'hydrolyse par rétention de configuration qui a été démontré chez les fucosidase de *Sulfolobus solfataricus* (Archeae) et de *Thermotoga maritima* (Eubacteria) dont les structures (TIM barrel) sont maintenant disponibles (Sulzenbacher et al., 2004).

Berteau et col. (2002) ont caractérisé une α -L-fucosidase isolée de l'hépatopancréas du mollusque marin *Pecten maximus* (coquille St Jacques) active sur le fucoïdane d'*Ascophyllum nodosum*. Il s'agit d'une protéine d'environ 50 kDa tétramérique avec un pI 6,3. Avec le substrat artificiel p-nitrophényl- α -L-fucoside, elle affiche un pH optimum de 4,0, une température optimale de 60°C à ce pH, un Km de 650 μ M, et une activité spécifique de 85 μ mol.min⁻¹.mg⁻¹. Elle semble très stable, et elle procède par rétention de configuration. Elle est très spécifique du fucose mais apparemment pas de la liaison glycosidique conformément à la plupart des fucosidases caractérisées. Son activité sur le fucoïdane natif d'*Ascophyllum nodosum* est significative quoique assez réduite (elle est curieusement inactive sur le fucoïdane de *Fucus vesiculosus*). De manière intéressante, son activité est renforcée en présence d'une sulfatase partiellement pure isolée du même mollusque (Daniel et al., 2001 ; cf § 3.2.2.) ce qui suggère une synergie entre ces deux enzymes et une limitation de l'activité de cet enzyme due aux sulfates. C'est à ma connaissance l'unique α -L-fucosidase agissant sur le fucoïdane qui ait été purifiée à homogénéité et caractérisée biochimiquement.

3.2.3. Les sulfatases

Les bases structurales des propriétés biologique et physiologiques des fucoïdanes dépendent en grande partie de la densité et du positionnement des groupes sulfates le long de la chaîne (Mulloy et al., 2000 ; Pereira et al., 2002 B). A ce jour la seule possibilité de cerner leur importance et de comparer des fucoïdanes entre eux en dépit de leur hétérogénéité (Chevolot et al., 1999 ; Pereira et al., 1999), la désulfatation chimique étant une alternative trop imprécise. L'obtention d'activités sulfatases spécifiques de ce type de polysaccharides revêt donc un intérêt tout particulier.

Les sulfatases font partie d'une famille de gène très conservée partageant d'une part de fortes homologues de séquence, et d'autre part une modification post traductionnelle d'un résidu sérine ou cystéine en formylglycine impliqué dans la catalyse. A ce jour quatre structures sont disponibles : 1hdh, 1fsu, 1p49 et 1e3c (3 sulfatases humaines et une bactérienne). Le repliement est de type $\alpha/\beta/\alpha$ dont le feuillet est principalement composé de huit brins parallèles. Les arylsulfatases on en plus un domaine C-terminal d'environ 75 résidus constitué d'un feuillet β (quatre brins) décoré d'une à trois hélices α .

L'identification biochimique de ces protéines est facilitée par leur capacité à tolérer des substrats synthétiques aromatiques permettant des détections par fluorescence (ex. 4-nitrophenyl sulfate ou nitrocatechol sulfate) bien que leur substrat naturel puisse être différent (Kertesz, 1999). De ce fait, si beaucoup de sulfatases ont été identifiées, elles sont souvent qualifiées d'arylsulfatase par défaut, sans que leur spécificité de substrat ne soit élucidée. Il apparaît cependant que ces arylsulfatases puissent cacher en réalité des carbohydrates sulfatases (Kertesz, 1999).

A ma connaissance, seule une étude a clairement montré l'action spécifique d'une sulfatase sur le fucoïdane (Daniel et al., 2001). Cet enzyme, extrait du bivalve *P. maximus*, est capable d'hydrolyser l'ester sulfate spécifiquement en C2 du α -L-fucopyranoside. Si son action sur le fucoïdane d'*A. nodosum* est avérée, il n'est toutefois pas purifié à l'homogénéité.

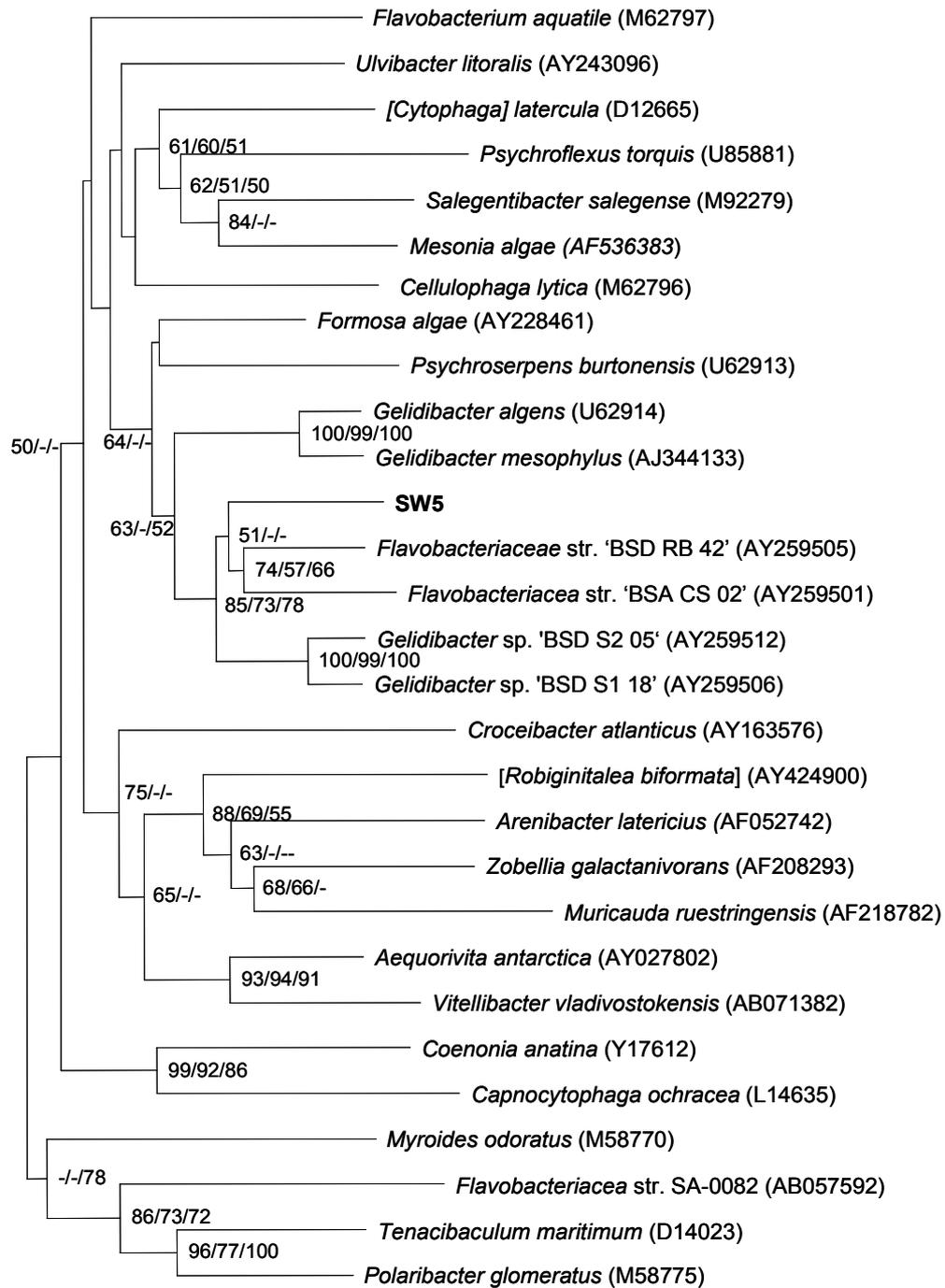
3.2.4. La α -L-fucoïdanase de SW5, FcnA

Le travail concernant cet enzyme a été initié par Descamps-Lepagnol au cours de sa thèse (1998). Elle a procédé à un criblage de bactéries en provenance de la station d'épuration des effluents d'une usine d'extraction d'alginate (Sobalg, Danisco Ingrédients, Landerneau, France). La souche SW5 a été sélectionnée parmi 51 souches isolées des boues d'épuration. Elle est déposée dans la collection DSMZ sous le numéro 12171. Elle présente une activité fucoïdanolytique extracellulaire maximale après cinq jours de cultures en milieu Zobell enrichi de fucoïdane (test de précipitation de la BSA par le fucoïdane non digéré). Sur la base de tests biochimiques et

physiologiques ainsi que de la séquence de son ARN16S, la souche SW5 appartient selon toutes vraisemblances au phylum des Bacteroidetes, et plus précisément au taxon des Flavobacteriaceae (Fig. 17; cf. Article 2). Elle présente une mobilité au contact de phase solide de type « gliding » (Barbeyron, communication personnelle) qui suggère une adaptation aux écosystèmes de type les biofilms microbiens. Notons ici qu'elle métabolise notamment le L-fucose, le D-galactose, le D-mannose, le D-glucuronate. La fraction protéique intracellulaire soluble ne produit pas d'oligosaccharides de manière significative (contrairement au lysat cellulaire total, résultats personnels).

Un enzyme extracellulaire capable d'hydrolyser du fucoïdane extrait de *Pelvetia canaliculata* a été purifié à l'homogénéité (cf. Article 3). Cette protéine (FcnA) a une masse moléculaire estimée à 105 kDa sur gel SDS-PAGE et à 143 kDa par gel filtration résolutive. Il est remarquable que son activité à l'homogénéité ne soit pas mesurable par le dosage des sucres réducteurs alors que du signal est détecté en début de purification. Plus la purification progresse, plus le signal diminue. Autrement dit ce dernier n'est pas directement le fait de la fucoïdanase purifiée. En outre, cela suggère l'existence d'activités fucoïdanolytiques complémentaires dans le surnageant de culture de SW5. Les pH optimum (7,5) et température optimale (20 à 25°C) mesurés par cette méthode, sur un pool d'enzymes semi-purifiés, ne sont donc pas fiables.

Ce même pool d'enzymes semi-purifiés (fractionnement au sulfate d'ammonium du surnageant de culture) a été employé pour synthétiser des oligo-fucoïdanes (cf. Article 2) à partir du fucoïdane de *P. canaliculata*. Deux des produits limites d'hydrolyse ont été purifiés et analysés d'un point de vue structural par RMN, sans que l'on puisse rigoureusement les attribuer à la seule activité de FcnA. Il s'agit d'un tétra- et d'un hexasaccharide comportant le motif répété suivant : $[-\rightarrow 4-\alpha\text{-L-fucopyranose-2,3-disulfate-1}\rightarrow 3-\alpha\text{-L-fucopyranose-2-sulfate-1}\rightarrow]_n$. Leur structure révèle l'hydrolyse de la liaison $\alpha(1,4)$ au sein de chaînes linéaires et répétitives. L'hydrolyse conduit à un rendement en masse de 66% d'oligomère de taille inférieure à 10 kDa. Et le motif structural décrit précédemment correspond à environ 10% du polymère initial. Ce lot enzymatique est également actif sur les fucoïdanes de *A. nodosum* et *F. spiralis*, mais il n'a pas été testé sur du fucoïdane de Laminariales.



0.01

Figure 17: Relations phylogénétiques de la souche bactérienne SW5 avec des représentants de Flavobacteriaceae marines (Erwan Corre, communication personnelle). L'arbre est basé sur la comparaison des séquences complètes d'ARN16S dont les numéros d'accès sont indiqués entre parenthèses. La topologie présentée est l'arbre obtenu par la méthode de neighbour-joining (correction de distance de Jukes et Cantor). A chaque nœud sont données les valeurs de « bootstrap » (100 réplicats) obtenues par les analyses de distances, de maximum de vraisemblance, et de maximum de parcimonie, respectivement (les tirets indiquent l'absence du nœud dans l'analyse correspondante). L'échelle représente le nombre attendu de changement par position nucléotidique.

3.3. Les glycosides hydrolases

Les glycosidases catalysent l'hydrolyse des liaisons glycosidiques et accélèrent la réaction d'un facteur 10^{17} en moyenne (Rye et Withers, 2000). Les carbohydrates exhibent une exceptionnelle diversité fonctionnelle et structurale. En réponse à cette diversité, les enzymes responsables de l'hydrolyse spécifique de liaison O-glycosidique (les O-glycosidases ; EC 3.2.1.X.) affichent eux aussi une large variété. Cette constatation se renforce chaque jour davantage du fait de l'accroissement du nombre des séquences et des structures tridimensionnelles disponibles. Le besoin d'une classification alternative à la nomenclature des chiffres EC de l'IUBMB (International Union of Biochemistry and Molecular Biology), illustrant plus les relations évolutives que la fonction précise de ces protéines, a poussé Henrissat (1991) à la mise en place, il y a une quinzaine d'année, d'une base de données originale (accessible sur le serveur web CAZy). Le schéma de cette nouvelle classification a pour principe l'intégration des caractéristiques structurales et mécanistiques à travers le regroupement en famille de glycosidases partageant des similarités de séquences. Ces similitudes de séquence reflètent des analogies structurales via l'existence d'un ancêtre commun et pas nécessairement la spécificité de substrat. Les familles ainsi créées sont donc généralement « poly-spécifiques ».

Au fil des années, cette classification, mise à jour régulièrement (Henrissat et Bairoch, 1993, 1996), voit le nombre de ces familles s'agrandir, mais de plus en plus lentement (Bourne et Henrissat, 2001). Si elle a débuté avec une quarantaine de familles, on convoite aujourd'hui la 100^{ème}. Quatre d'entre elles ont cependant été supprimées (GH21, 40, 41 et 60) suite à des remaniements. Puisque les structures sont mieux conservées que les séquences, le groupement de familles en clans (ou superfamilles) a été introduit (Henrissat et Davies, 1997). A ce jour (février 2005), 68 familles (14355 séquences) ont leur repliement identifié (Tab. 5) soit à partir de la structure d'un de leur membre (59), soit d'après l'affiliation à un clan dont le repliement est connu (9). Le repliement le plus représenté est incontestablement le TIM barrel ou tonneau (β/α)₈ avec 28 familles (7085 séquences).

Tableau 5 : Récapitulatif des familles de glycosides hydrolases répertoriées dans le serveur CAZy dont le repliement est identifié (février 2005). Les familles dont aucun membre n'a de structure résolue mais appartenant à un clan caractérisé sont également incorporées au tableau. Ce dernier concerne 14355 séquences répertoriées dans le serveur CAZy. Les familles présentant un intérêt direct dans ce travail sont colorées en jaunes. Les repliements colorés en gris sont évoqués dans les chapitres de résultats.

Clan	Famille	Statistiques (nb. séq.)	Mécanisme	Repliement	Représentants	Code PDB
A	1	569	Rétention	(β/α) ₈	myrosinase	2MYR
	2	591	Rétention	(β/α) ₈	β -galactosidase	1BGL
	5	552	Rétention	(β/α) ₈	endoglucanase	1ECE
	10	229	Rétention	(β/α) ₈	xylanase	1XAS
	17	357	Rétention	(β/α) ₈	β -glucanase	1GHR
	26	59	Rétention	(β/α) ₈	mannanase	1ODZ
	30	30	Rétention	(β/α) ₈	glucocerebrosidase	1OGS
	35	162	Rétention	(β/α) ₈	β -galactosidase	1TG7
	39	34	Rétention	(β/α) ₈	β -xylosidase	1PX8
	42	62	Rétention	(β/α) ₈	β -galactosidase	1KWG
	50	15			β -agarase	
	51	73	Rétention	(β/α) ₈	arabinofuranosidase	1PZ2
	53	36	Rétention	(β/α) ₈	endo-1,4- β -galactanase	1FHL
	59	10	Rétention		galactocerebrosidase	
	72	70	Rétention		glucanosyltransglycosylase	
	79	21	Rétention		heparanase	
86	6			β -agarase		
B	7	127	Rétention	β -jelly roll	cellobiohydrolase	1CEL
	16	412	Rétention	β -jelly roll	κ -carrageenase β -agarase	1DYP 1O4Y
C	11	200	Rétention	β -jelly roll	xylanase	1BCX
	12	74	Rétention	β -jelly roll	endoglucanase	2NLR
D	27	115	Rétention	(β/α) ₈	α -galactosidase	1SZN
	36	122	Rétention		α -galactosidase	
E	33	153	Rétention	6- β propeller	sialidase	2SIL
	34	1384	Rétention	6- β propeller	neuraminidase	1NSB
	83	238	Rétention	6- β propeller	neuraminidase	1E8T
F	43	206	Inversion	5- β propeller	arabinanase A	1GYD
	62	17			arabinofuranosidase	
H	13	2014	Rétention	(β/α) ₈	α -amylase	6TAA
	70	44	Rétention		dextranucrase	
	77	119	Rétention	(β/α) ₈	amylomaltase	1CWY
I	24	155	Inversion	$\alpha + \beta$	lysozyme	1LYD
	46	19	Inversion	$\alpha + \beta$	chitosanase	1CHK
	80	14	Inversion		chitosanases	

J	32	443	Rétention	5- β propeller	invertase	1UTW
	68	49	Rétention	5- β propeller	levansucrase	1OYG
K	18	1103	Rétention	(β/α) ₈	chitinase	1CTN
	20	202	Rétention	(β/α) ₈	chitibiase	1QBA
L	15	136	Inversion	(α/α) ₆	glucoamylase	3GLY
	65	83	Inversion	(α/α) ₆	maltose phosphorylase	1H54
	94	49	Inversion	(α/α) ₆	chitobiose phosphorylase	1V7V
M	48	20	Inversion	(α/α) ₆	cellobiohydrolase	1FCE
	8	77	Inversion	(α/α) ₆	chitosanase	1V5C
N	28	415	Inversion	hélice β	rhamnogalacturonase	1RMG
	49	8	Inversion	hélice β	dextranase	1OGM
Familles orphelines	3	503	Rétention		exoglucanase	1EX1
	4	125			α -glucosidase	1NRH
	6	103	Inversion		cellobiohydrolase	3CBH
	9	239	Inversion	(α/α) ₆	endoglucanase	1CLC
	14	118	Inversion	(β/α) ₈	β -amylase	1BTC
	19	497	Inversion	tout α	chitinase	1BAA
	22	208	Rétention	$\alpha + \beta$	lysozyme	1HEL
	23	631	Inversion	$\alpha + \beta$	lysozyme	153L
	25	179	Inversion	(β/α) ₅ (β) ₃	lysozyme	1JFX
	29	80	Rétention	(β/α) ₈	α -L-fucosidase	1HL8
	31	266	Rétention	(β/α) ₈	α -xylosidase	1WE5
	38	136	Rétention	(β/α) ₇	α -mannosidase	1O7D
	45	65	Inversion	β ₆ barrel	endoglucanase	1ENG
	47	121	Inversion	(α/α) ₇	α -mannosidase	1DL2
	54	17	Rétention		L-arabinofuranosidase	1WD3
	56	46	Rétention	(β/α) ₇	hyaluronidase	1FCQ
	57	52	Rétention	(β/α) ₇	glucanotransferase	1K1W
	58	5		6- β propeller	endo-sialidase	1V0E
	67	26	Inversion	(β/α) ₈	α -glucuronidase	1GQI
	74	20	Inversion	7- β propeller	cellobiohydrolase	1SQJ
82	2	Inversion	β hélice	ι -carrageenase	1H80	
88	36		(α/α) ₆	glucuronyl hydrolase	1VD5	
90	6	Inversion	β hélice	endorhamnosidase	1CLW	

Compte tenu de la modularité fréquente des « cazymes », le server CAZy recense également différentes familles de domaines non catalytiques se fixant aux carbohydrates, les « Carbohydrate Binding Modules » ou CBMs.

3.3.1. Deux mécanismes enzymatiques

Malgré la diversité structurale des glycosidases, seulement deux mécanismes d'hydrolyses sont décrits. L'hydrolyse enzymatique de la liaison glycosidique s'effectue par une catalyse acide nécessitant deux acides aminés essentiels : un catalyseur acide/base et un nucléophile. Ces deux résidus sont quasi exclusivement des aspartates et/ou des glutamates (Davies et Henrissat, 1995). Les deux mécanismes évoqués conduisent soit à la rétention soit à l'inversion de la configuration du carbone anomérique (Fig. 18).

Depuis les dix dernières années des avancées considérables ont été réalisées dans la compréhension fine du cheminement catalytique des glycosidases. Ces progrès sont imputables d'une part à l'augmentation des structures protéiques disponibles sous formes natives, mutée, en complexe avec leur substrat, voire même d'intermédiaire glycosyl-enzyme (Rye et Withers, 2000 ; Vassela et al., 2002). D'autre part, l'emploi de récents inhibiteurs mimant les états de transitions s'est avéré tout aussi pertinent. Ces analyses structurales et cinétiques montrent l'importance capitale des contraintes exercées sur la conformation du résidu osidique au sous-site -1, dans la bonne orientation de la liaison glycosidique cible par rapport aux fonctions catalytiques (Rye et Withers, 2000 ; Vassela et al., 2002).

3.3.1.1. La rétention de configuration

Le mécanisme qui retient la configuration anomérique se décompose en deux substitutions nucléophiles successives (Fig. 18). L'oxygène de la liaison glycosidique est protoné par le catalyseur acide/base, alors que le nucléophile attaque le carbone anomérique déclenchant ainsi le départ de l'aglycone. Cette première étape conduit au clivage de la liaison glycosidique et à la formation d'un intermédiaire réactionnel covalent glycosyl-enzyme. Dans un second temps, le catalyseur acide/base tente de récupérer son proton d'une molécule d'eau qui, ainsi activée, attaque alors le carbone

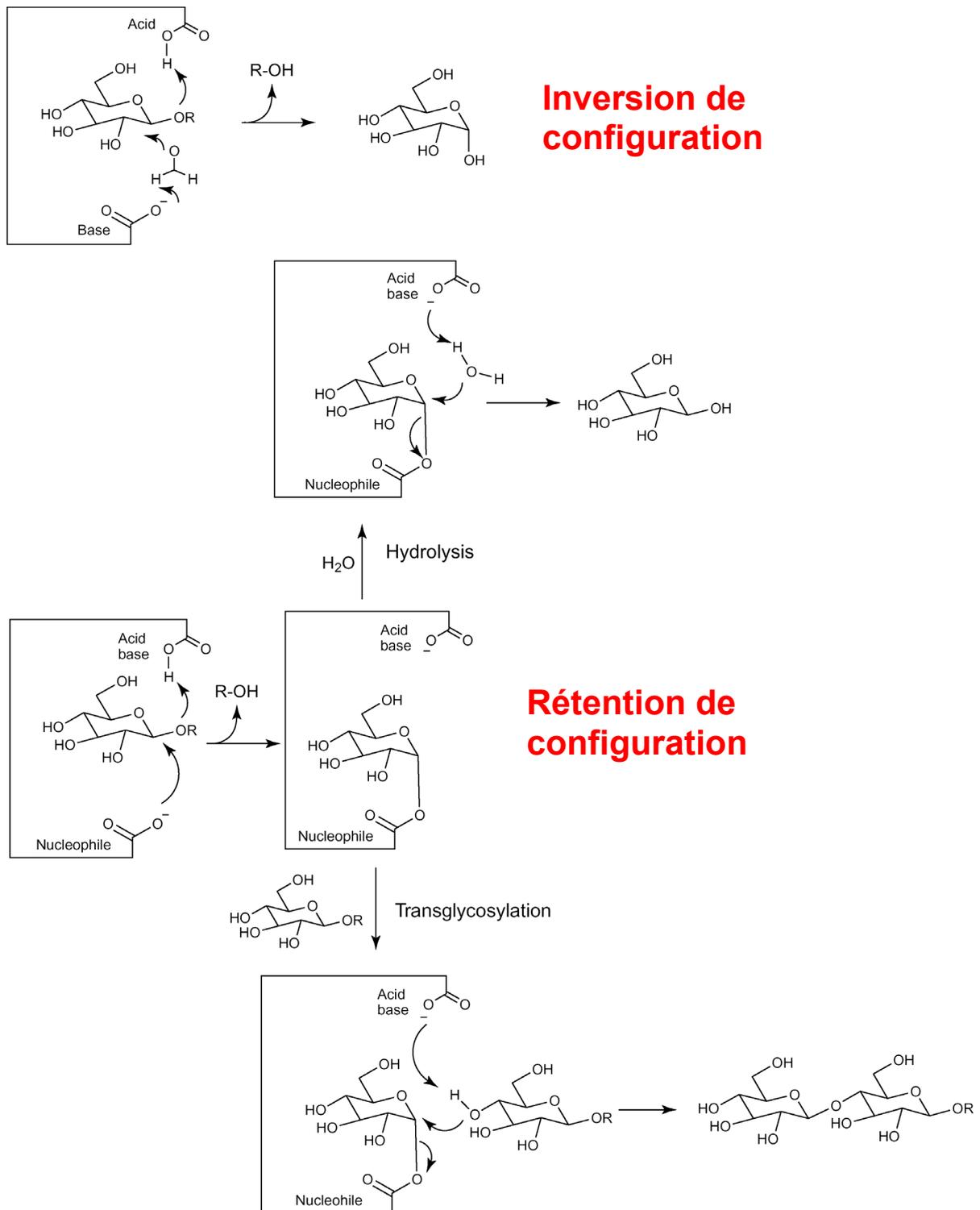


Figure 18 : Mécanismes réactionnels des glycosidases (adapté de Perugino et al., 2004). Les glycosidases fonctionnent soit par inversion soit par rétention de configuration suivant qu'elle conserve ou non la stéréochimie du carbone anomérique à la libération des produits. Les enzymes retenant la configuration sont potentiellement sujets aux réactions secondaires de transglycosylation.

anomérique par son oxygène. La liaison glycosyl-enzyme se rompt et le sucre est libéré. Les états de transition de ces deux étapes sont de type ions oxocarbonium. Enfin, chez les enzymes adoptant ce mécanisme, la distance moyenne entre les fonctions carboxyliques des deux résidus acides catalytiques est de 5,5 Å.

3.3.1.2. L'inversion de configuration

Contrairement au précédent mécanisme, celui qui induit l'inversion de configuration anomérique, ne comporte qu'une seule étape de substitution nucléophile (Fig. 18). L'oxygène de la liaison glycosidique est protoné par le catalyseur acide ce qui entraîne le départ de l'aglycone. Cette protonation est concomitante à l'attaque, à l'opposé, du carbone anomérique par une molécule d'eau activée par le résidu basique. Cette substitution nucléophile unique conduit donc à la libération d'une extrémité réductrice dont la stéréochimie est opposée à celle de la liaison hydrolysée. Comme pour le mécanisme de rétention de configuration, le catalyseur acide/base doit être à une distance de l'oxygène de la liaison glycosidique, compatible avec la formation d'une liaison hydrogène. Par contre, de l'autre côté du sucre, le résidu basique est généralement plus éloigné du carbone anomérique de façon à permettre la présence d'une molécule d'eau entre ce dernier et la fonction carboxylique catalytique. Ainsi, la distance moyenne, entre les deux fonctions carboxylique des résidus catalytiques est en moyenne de 10 Å. Les familles GH28 et GH82 (α -carraghénase), deux hélices β , constituent toutefois deux exceptions notoires puisque cette distance y est réduite à 5-6 Å.

3.3.2. Mode d'action et topologie du site actif

3.3.2.1. Les modes d'action

Initialement, étaient opposés deux comportements des glycosidases vis à vis des polysaccharides. Il y avait celles qui attaquaient par les extrémités, dites « exo » et celles qui attaquaient à l'intérieure de la chaîne, dites « endo ». A cette notion il faut en ajouter une seconde : la processivité.

Les enzymes intervenant sur les polymères se caractérisent par quatre grands types de comportements. On distingue ceux qui appréhendent la molécule par une extrémité, comportement « exo », de ceux ayant la capacité à agir au cœur de la chaîne, comportement « endo » (Fig. 19). Par ailleurs, indépendamment de la notion précédente, l'enzyme peut soit se décrocher du substrat après chaque acte catalytique, ce qui engendre des rencontres enzyme/substrat aléatoires, soit se maintenir à son contact et enchaîner les actes catalytiques via une translation de l'enzyme, comportement dit multi-attaque ou processif (Fig. 19). Cette dernière notion contrairement à la première n'est pas de type tout ou rien. Ainsi la processivité d'un enzyme sera plus ou moins marquée en fonction de la stabilité du complexe enzyme-substrat en dehors du processus catalytique proprement dit. La processivité implique un mouvement unidirectionnel de l'enzyme par rapport à son substrat qui selon les cas est orienté vers l'extrémité soit réductrice soit non réductrice. Un comportement endo stricte aura pour effet de réduire rapidement le degré de polymérisation d'un polysaccharide par opposition à un comportement exo (Fig. 19). Ce dernier, qu'il soit processif ou non, se caractérise par une diminution fatalement plus lente des chaînes. Après l'acte initial, le comportement enzymatique endo-processif ou exo-processif est similaire de ce point de vue (Fig. 19).

Le caractère processif traduit souvent la présence de domaine non catalytique type « carbohydrate binding domain » permettant de maintenir le contact avec le substrat pendant le déplacement du domaine catalytique. Toutefois ce n'est pas toujours le cas et d'autres solutions structurales autorisent ce comportement (ex. tunnel catalytique) qui illustre le rôle fonctionnel primordial assuré par la topologie du site actif.

3.3.2.2. Les topologies de sites actifs

Les données structurales ont permis de faire le lien entre ces modes d'actions et trois topologies de site actif dites poche, gorge et tunnel (Fig. 20) (Davies et Henrissat, 1995).

La poche est une cavité « aveugle » adaptée à la reconnaissance d'une extrémité de chaîne. Elle limite en outre le nombre de résidus osidiques pouvant y pénétrer.

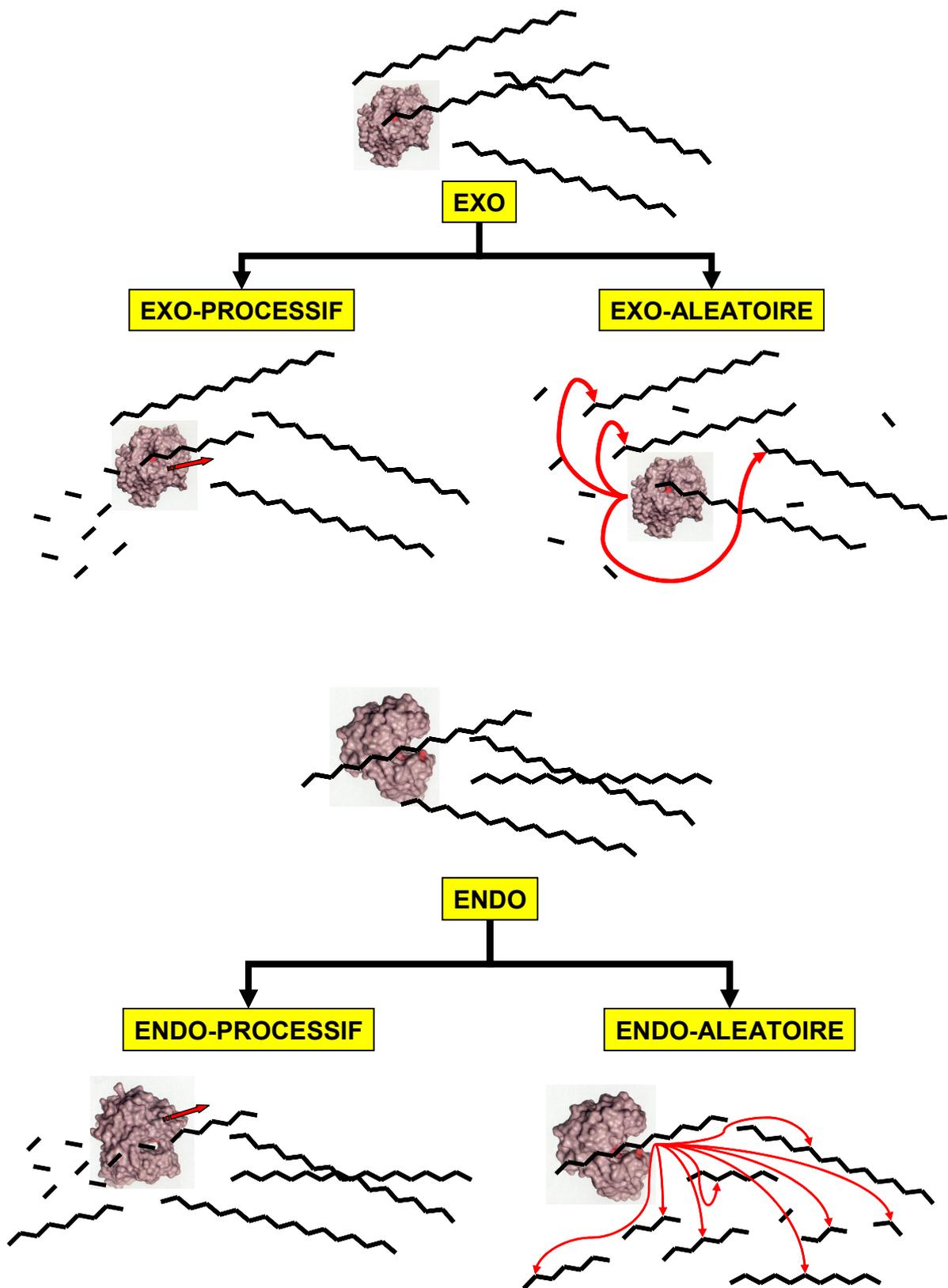


Figure 19 : Les quatre grands modes d'action des glycosidases intervenant sur les polysaccharides.

Cette topologie se rencontre principalement dans les monosaccharidases (ex. β -galactosidase) et les exo-polysaccharidases libérant des mono- ou disaccharides (β -amylase). Elle est spécialement efficace sur les polysaccharides simples présentant un grand nombre d'extrémités (ex. le grain d'amidon dont la structure ramifiée radiale présente les extrémités de chaînes à sa surface). Elle n'est donc pas adaptée aux substrats fibreux comme la cellulose, qui ne possède que très peu d'extrémités de chaînes libres. Enfin, elle est souvent associée à des enzymes intervenant dans la phase terminale de processus multi-étapes de dégradation.

La gorge est une crevasse ouverte qui procure une surface d'interaction avec le substrat plus importante. C'est une topologie typique de glycosidases intervenant en endo sur des polysaccharides. Elle permet d'appréhender des unités de répétitions plus complexes.

Le tunnel enfin dérive de la précédente dans la mesure où il s'agit d'une gorge refermée dont le rôle de glissière est nettement accentué. Elle permet de maintenir un contact étroit avec le substrat linéaire. On retrouve de ce fait cette topologie dans des enzymes endo-processifs et qui plus est sur des polysaccharides fibreux en phase solide (cellulose cristalline, fibre de carraghénane). L'ouverture/fermeture du tunnel par changement de conformation est indispensable au comportement endo. Les changements conformationnels sont le fait de boucles ou de domaines mobiles.

3.4. Intérêt des enzymes en glycobiologie

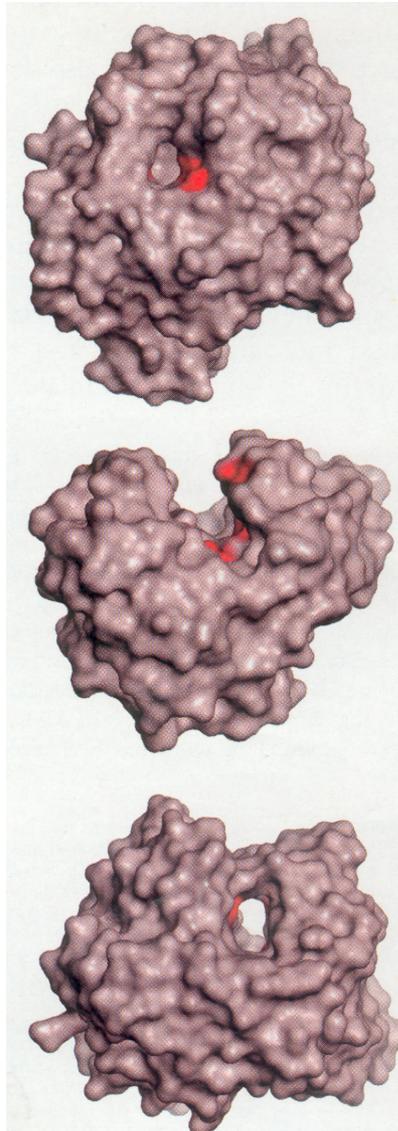
3.4.1. Analyse structurale

Pour répondre à la complexité et la diversité structurale des polysaccharides, denombreux procédés chimiques ont été développés. La composition osidique peut être déterminée par hydrolyse acide et analyse des produits par chromatographie (sur couche mince, ou en phase liquide) couplée à des méthodes de détection réfractométrique, UV, fluorimétrique ou électrochimique. Le positionnement des liaisons glycosidiques peut être déterminé par méthylation des hydroxyles libres et analyse des dérivés par chromatographie phase gazeuse couplée à la spectrométrie de

masse (GC-MS). Enfin la séquence osidique peut être déduite d'expérience de méthylation, de la sensibilité du polysaccharide à divers traitements tels que l'oxydation/réduction au périodate, acétolyse, hydrolyse acide, dégradation alcaline et surtout la spectrométrie RMN. Cependant, ces méthodes se réalisant dans des conditions souvent assez dures peuvent être biaisées par la résistance ou la fragilité de certains type de liaisons ou de substitutions. Les réactions secondaires sont multiples. Ainsi, l'hydrolyse acide peut s'accompagner de la perte de sulfate. La méthylation du fucoïdane peut entraîner la perte de substituants acétyles (Chizov et al, 1999). Sa désulfatation chimique, en plus d'être fréquemment incomplète, occasionne la rupture de liaisons glycosidiques (Mulloy et al., 1994). Enfin la rupture de ces liaisons par acétolyse s'accompagne parfois de réactions de condensation, créant des liaisons absentes du polysaccharide natif.

Face à la complexité et l'imprécision de ces méthodes analytiques, l'utilisation des enzymes présentent des avantages indiscutables du fait de la spécificité stéréochimique des réactions qu'elles catalysent en conditions généralement très douces (sous réserve que leur spécificité de reconnaissance soit maîtrisée). Concernant les polysaccharides on recherchera des enzymes rompant les liaisons glycosidiques (glycoside hydrolases, lyases) ou les créant (glycosyl transférases), modifiant les profils de substitution (ex. sulfatases, sulfotransférases, ou sulfurylase), ou épimérisant certains résidus (ex. C5-épimérase des alginates).

L'analyse des produits de réactions enzymatiques (à l'exception de la synthèse) est valorisable sous deux angles complémentaires selon que l'on souhaite observer ce que je qualifierai de produits typiques (ou légitimes), ou ce qui « résiste ». Dans le cas de polysaccharides hétérogènes, la première approche, plus directe, consiste donc à vérifier la présence de tel ou tel contexte structural (plus ou moins précis selon la spécificité de l'enzyme) dans le polysaccharide étudié. La seconde approche exploitera la capacité d'enzymes à homogénéiser/simplifier un substrat en éliminant certaines liaisons/substitutions minoritaires (comme par exemple des branchements), ou en complétant des substitutions manquantes (ex. sulfurylases et pont 3,6 anhydro des galactanes). Couplés à des techniques séparatives, des traitements enzymatiques judicieux permettent ainsi de préparer des fractions de polysaccharides ou d'oligosac-



Poche

- monosaccharidases
- exopolysaccharidases

Crevasse (ou gorge)

- endopolysaccharidases strictes

Tunnel

- enzymes endoprocessives

Figure 20 : Topologie des sites actifs de glycosidases (d'après Davies et Henrissat, 1995). Les résidus catalytiques sont colorés en rouge. Le site actif type « poche » de la glucoamylase de *Aspergillus awamori* ; le site actif type « crevasse » de l'endoglucanase E2 de *Thermomonospora fusca* ; le site actif type « tunnel » de la cellobiohydrolase II de *Trichoderma reesei*.

charides dont l'homogénéité facilite des analyses structurales ultérieures (structure primaire, mais aussi secondaire et tertiaire).

Une des toutes premières applications a été d'employer les exoglycosidases pour réaliser du séquençage d'oligosaccharide (Kobata, 1979 ; Harvey, 1999). Couplée avec des techniques analytiques récentes telles que la RMN, le Biacore, l'électrophorèse capillaire ou la spectrométrie de masse, cette méthode d'analyse connaît un essor important bien que limité par la diversité des enzymes disponibles.

En dehors de l'exploration structurale, les enzymes intervenant sur les carbohydrates présentent également une certaine utilité dans la production de protoplastes (Le Gall et al., 1990) ou dans la compréhension de l'agencement des parois cellulaires. Ils présentent enfin un intérêt dans l'analyse des activités biologiques des sucres via les modifications fines de la taille des oligomères ou de certains de leurs groupements chimiques, voire la synthèse d'oligosaccharides de novo.

3.4.2. Synthèse d'oligosaccharides

Si les enzymes permettent l'obtention d'oligosaccharides spécifiques par dégradation, il est également possible de les utiliser pour la synthèse de ces derniers. Ils constituent en effet une alternative à la synthèse organique souvent délicate et toujours compliquée. Trois classes d'enzymes sont envisageables, les glycosyl transférases responsables de la biosynthèse naturelle des glycanes, les glycosidases qui dans certaines conditions peuvent fonctionner « à l'envers », et enfin les glycosynthase, catalyseurs non naturels issus des glycosidases par génie génétique. Les glycosyl transférases présentent l'inconvénient d'avoir pour substrat des glycosyl-esters de nucléosides polyphosphatés compliqués à synthétisés et coûteux à purifier. L'utilisation des glycosides hydrolases pour la synthèse est connue depuis près d'un siècle (Bourquelot et Bridel, 1913). C'est d'ailleurs cette observation qui a permis de mettre en évidence la réversibilité des réactions enzymatiques. Comme toutes les enzymes, les glycosidases accélèrent la mise en équilibre d'une réaction chimique, en l'occurrence ici l'hydrolyse de la liaison O-glycosidique :



Un contrôle thermodynamique de l'équilibre permet de favoriser l'un ou l'autre des sens de la réaction. Ainsi, si l'enzyme est en présence d'un rapport Ose-O-R/Ose-OH très faible, accompagné d'une réduction de la disponibilité de l'eau ([Ose-OH + R-OH] à 80-90% p/v du milieu réactionnel, plus ajout de co-solvant ou co-solutés), il est alors possible d'utiliser les glycosidases pour accélérer la réaction dans le sens de la synthèse (Johansson et al., 1986).

Un contrôle cinétique de l'équilibre est également envisageable par les enzymes adoptant le mécanisme de rétention de configuration (« retaining enzymes », cf. § 3.3.) via une réaction secondaire dite de transglycosylation (Fig. 18, Crout et Vic, 1998). Dans ce mode de synthèse la première étape de l'hydrolyse a lieu, s'accompagnant de la formation d'un intermédiaire covalent glycosyl-enzyme. L'astuce consiste alors à faire attaquer l'intermédiaire non pas par une molécule d'eau mais par une autre molécule comportant un groupement hydroxyle (ose, alcool, ou même acides aminés). Pour ce faire, ce composé dit accepteur est ajouté à forte concentration dans le milieu réactionnel. Bien que nettement moins représenté que l'eau (100 mM environ contre 55 M), ce composé entre en compétition avec l'eau. A titre d'exemple la α -L-fucosidase de *Penicillium multicolor* permet d'obtenir par transglycosylation du disaccharide lié en $\alpha(1,3)$ (Ajisaka et al., 1998). L'hydrolyse du composé néoformé peut survenir mais dans un cycle catalytique ultérieur. Pour prévenir ce phénomène et améliorer le rendement, des dispositifs ont été développés pour piéger le disaccharide synthétisé, bien que dans la majorité des cas l'excès de substrat suffise à permettre l'accumulation des produits de transglycosylation.

Contrairement aux glycosyltransférases, les glycosidases sont peu spécifiques pour la molécule acceptrice. Cependant s'il est relativement aisé d'obtenir des disaccharides, la production de degrés de polymérisation (DP) supérieurs demeure problématique. L'emploi d'endoglycosidases tolérant des di- ou tétrasaccharides pourrait être une solution (Kobayashi, 1999).

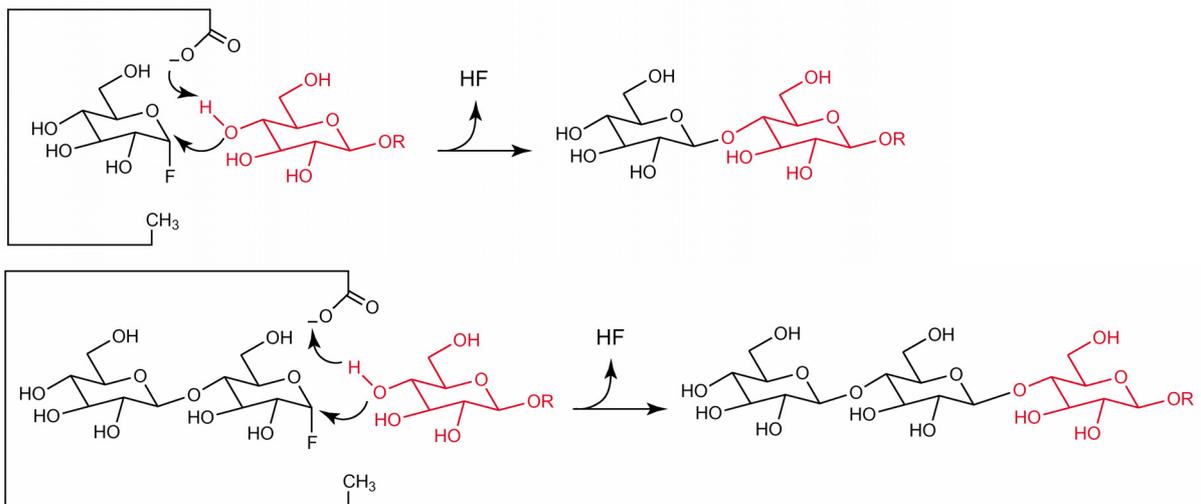


Figure 21 : Mécanisme réactionnel courant des glycosynthases (adapté de Perugino et al., 2004). Le degré de polymérisation (DP) des produits dépend de la propension de l'enzyme à accepter des longs substrats. Les glycosynthases dérivant d'endoglycosidases engendrent donc plus favorablement des DP importants.

Un autre problème inhérent à l'usage des glycosidases est le faible rendement de la transglycosylation. Pour y palier, une nouvelle classe de catalyseurs non naturels a été développée à partir des glycosidases de type « retaining » (Mackenzie et al., 1998). Ces enzymes sont obtenus en remplaçant par mutagenèse dirigée le résidu acide nucléophile responsable de la formation de l'intermédiaire covalent glycosyl-enzyme par un résidu plus neutre (glycine, alanine, serine...). La mutation a pour effet de réduire presque entièrement l'activité hydrolytique. Cependant le site actif modifié peut accepter un ose activé (ex. de type fluoro-glycoside ; William et Withers, 2000) dont l'anométrie est identique à celle de l'intermédiaire glycosyl-enzyme naturel (Fig. 21). Les glycosynthases sont capables de synthétiser de façon efficace des oligosaccharides puisqu'il est possible de convertir jusqu'à 92% de l'accepteur (Mackenzie et al., 1998). Il semble donc que ce soit des outils d'avenir.

4. Présentation du travail

A l'issue de cette introduction, je me propose de dégager quelques points intéressants. Tout d'abord, il s'avère que les polysaccharides sulfatés ne sont pas répartis de façon ubiquiste dans le vivant. Ils sont présents chez les Métazoaires et les végétaux. Dans les deux cas, ils se répartissent essentiellement à l'interface environnement-cytoplasme. A l'exception des glycosaminoglycanes animaux, il semble que la présence de polysaccharides sulfatés soit étrangement associée à l'environnement marin. Dans ce dernier, les macroalgues marines rouges et brunes sont des producteurs majeurs de polysaccharides sulfatés, quantitativement parlant. Cette production représente une source de carbone importante donnant lieu à un recyclage par les organismes marins brouteurs et par les microorganismes notamment bactérien. Ces derniers y jouent un rôle capital. Ils constituent de ce fait une source précieuse d'enzymes participant au catabolisme de ces polysaccharides sulfatés. Ainsi, les deux glycoside hydrolases concernées par ce travail ont été isolées de deux bactéries marines.

Le chapitre 2 est consacré aux travaux que j'ai entrepris pour surexprimer la λ -carraghénase de *P. carrageenovora*. Les résultats obtenus sont pour l'essentiel rassemblés dans un projet d'article (Article 1). Au commencement de ce travail, deux

κ -carraghénases et deux ι -carraghénase étaient caractérisées biochimiquement et structuralement. L'intérêt fondamental porté à la λ -carraghénase reposait donc sur la comparaison avec ces quatre prédécesseurs, mais aussi sur l'acquisition d'un nouvel outil d'investigations des carraghénanes. Il s'agissait également de compléter l'arsenal enzymatique utilisés pour l'obtention d'oligosaccharides sulfatés bioactifs.

Le chapitre 3 concerne les travaux portant sur la surexpression de la fucoïdanase de la bactérie SW5. Il comprend deux articles (Article 2 et 3). Le premier reprend principalement les travaux réalisés par Valérie Descamps lors de sa thèse pour isoler et caractériser la souche SW5. Il figure dans ce chapitre du fait de ma participation à sa rédaction et aux résultats concernant la culture de la souche en fermenteur et le suivi de l'activité fucoïdanase. En outre, il constitue un complément indispensable à la lecture du deuxième article. Ce dernier correspond davantage à mon travail exception faite de la purification de l'enzyme sauvage. Les recherches portant sur cet enzyme sont motivées d'une part, par l'absence de catalyseur caractérisé permettant l'obtention d'oligofucanes, et d'autre part, par la forte activité biologique de ces produits sur les systèmes végétaux. La caractérisation de cette enzyme apportera en outre des éléments de comparaison précieux pour la compréhension de l'enzymologie des polysaccharides sulfatés.

A l'exception de la méthodologie employée, il ne se dégage pas de dénominateur commun suffisamment consistant entre ces deux grandes parties de mon travail. C'est la raison pour laquelle j'ai choisi de ne pas réaliser de chapitre spécifiquement dédié aux conclusions de ces travaux de thèse. Dès lors, chacun de ces deux volets de mon travail a donné lieu à une conclusion et une mise en perspective détaillées à la fin des chapitres 2 et 3.

Enfin, ces travaux ont donné lieux à la rédaction et aux dépôts de deux demandes de brevets d'invention :

- n° 04 06667 : « Gène isolé codant pour l'endofucoïdanase et endofucoïdanases recombinantes obtenues à l'aide de ce gène » (référence : « endofucanase recombinante »), dépôt sur la France en date du 18 juin 2004),
- n° 04 08201 : « Gène isolé codant pour la lambda-carraghénase et la lambda-carraghénase recombinante obtenue à l'aide de ce gène » (référence : « carraghénase type lambda »), dépôt sur la France en date du 23 juillet 2004.

CHAPITRE 2

LA λ -CARRAGHÉNASE :
PURIFICATION DE LA PROTÉINE,
CLONAGE & SÉQUENÇAGE DU GÈNE,
ANALYSE DE LA SÉQUENCE,
SUREXPRESSIION DE L'ENZYME.

L'UMR 7139 s'est attachée depuis maintenant une quinzaine d'années à la description d'enzymes de modification des polysaccharides pariétaux de macroalgues. Un intérêt particulier a été porté sur des enzymes procaryotiques de dégradation des galactanes neutres (agarose) ou sulfatés (κ - et ι -carraghénane). Plusieurs d'entre eux ont été finement caractérisés ainsi que leurs produits, tant d'un point de vue biochimique, que structural. Si les deux β -agarases (Allouch et al. 2003, 2004) et les deux κ -carraghénases (Michel et al 2001A ; Barbeyron et al, 1998, 1994) décrites appartiennent à la famille GH16 des glycosides hydrolases, les deux ι -carraghénases identifiées (Michel et al 2003, 2001B ; Barbeyron et al, 2000) constituent la famille GH82. Ces trois polysaccharides matriciels sont gélifiants et portent de zéro à deux substitutions sulfates par disaccharides. Le λ -carraghénane contient trois groupes sulfates par répétitions et l'absence du pont 3,6-anhydro sur le résidu galactose lié en 4 inverse sa conformation (4C_1 -conformation) comparée à celles des κ - et ι -carraghénane (1C_4 conformation). Ceci a pour effet d'empêcher sa gélification au profit d'une forte viscosité. Cette distinction du λ -carraghénane a motivé l'étude d'un enzyme participant à sa dégradation. En outre ce polysaccharide est d'une manière générale beaucoup moins homogène que le κ - et ι -carraghénane, et son étude gagnerait à pouvoir bénéficier d'outils enzymatiques caractérisés permettant d'isoler des fractions résistantes.

Ce chapitre a pour objet la description de l'ensemble des résultats obtenus concernant un enzyme dégradant le λ -carraghénane. Cet enzyme a été purifié de la bactérie marine *Pseudoalteromonas carrageenovora*. La présence d'une activité extracellulaire de dégradation du λ -carraghénane produite par cette γ -protéobactérie avait été initialement reportée par Weigl et Yaphe (1966). Poursuivant ces travaux, Johnston et McCandless, purifiaient partiellement l'enzyme en 1973 et avançaient l'hypothèse d'un complexe multienzymatique. Cette hypothèse fut mise à bas par Greer (1984) qui purifia à homogénéité une protéine extracellulaire unique de 98 kDa. Malgré la capacité de l'enzyme à réduire drastiquement la taille du polymère, l'auteur ne mesurait qu'une faible activité en sucres réducteurs et c'est donc par mesure de la

décroissance de viscosité qu'il suivit la purification de cet enzyme. Il montra la spécificité de la protéine pour le λ -carraghénane et distingua deux produits limites largement majoritaires sans toutefois les avoir caractérisés. Reprenant ces travaux au laboratoire, Potin (1992) ne parvint pas, sur la base d'un protocole différent, à reproduire la purification de cette protéine. Néanmoins Chabeau (1995) procéda à la synthèse d'une sonde nucléique censée cibler le gène de cette λ -carraghénase. L'objectif de mes travaux était de cloner et séquencer le gène codant pour cet enzyme, afin de procéder à la surexpression de cette activité enzymatique dans *E. coli*.

L'essentiel de ce travail a été rédigé sous forme d'un article disposé à la fin du chapitre (cf. Article 1). Néanmoins certains résultats complémentaires seront développés et discutés ci-dessous. La lecture sera facilitée si elle débute par cet article.

1. La purification à homogénéité de la λ -carraghénase de *P. carrageenovora*

Cet enzyme n'avait jamais été purifié à homogénéité au laboratoire. Cette purification a nécessité la mise au point d'un protocole original (Fig. 22 et 23) largement inspiré toutefois de la méthode de Greer (1984). Il est décrit dans son intégralité dans l'article 1. Greer est parvenu à purifier l'enzyme à homogénéité dès la chromatographie d'interactions hydrophobes. J'ai comme lui constaté que l'enzyme n'est pas retenu sur cette colonne mais simplement retardé. Le pool des fractions portant l'activité ne révéla qu'une seule bande, à la taille attendue (98 kDa), en gel SDS-PAGE coloré au bleu de Coomassie. Néanmoins une coloration au nitrate d'argent (Kit Pharmacia biotech) montra la présence de bandes surnuméraires de taille inférieure. Afin de lever toute ambiguïté une chromatographie d'exclusion de taille sur une colonne Superdex 200HR fut réalisée après concentration par ultrafiltration de ce pool. La protéine ainsi récoltée était parfaitement pure sur gel SDS-PAGE coloré au nitrate d'argent. Cependant l'étape d'ultrafiltration engendra une grosse perte de protéine et contribua au faible rendement de purification.

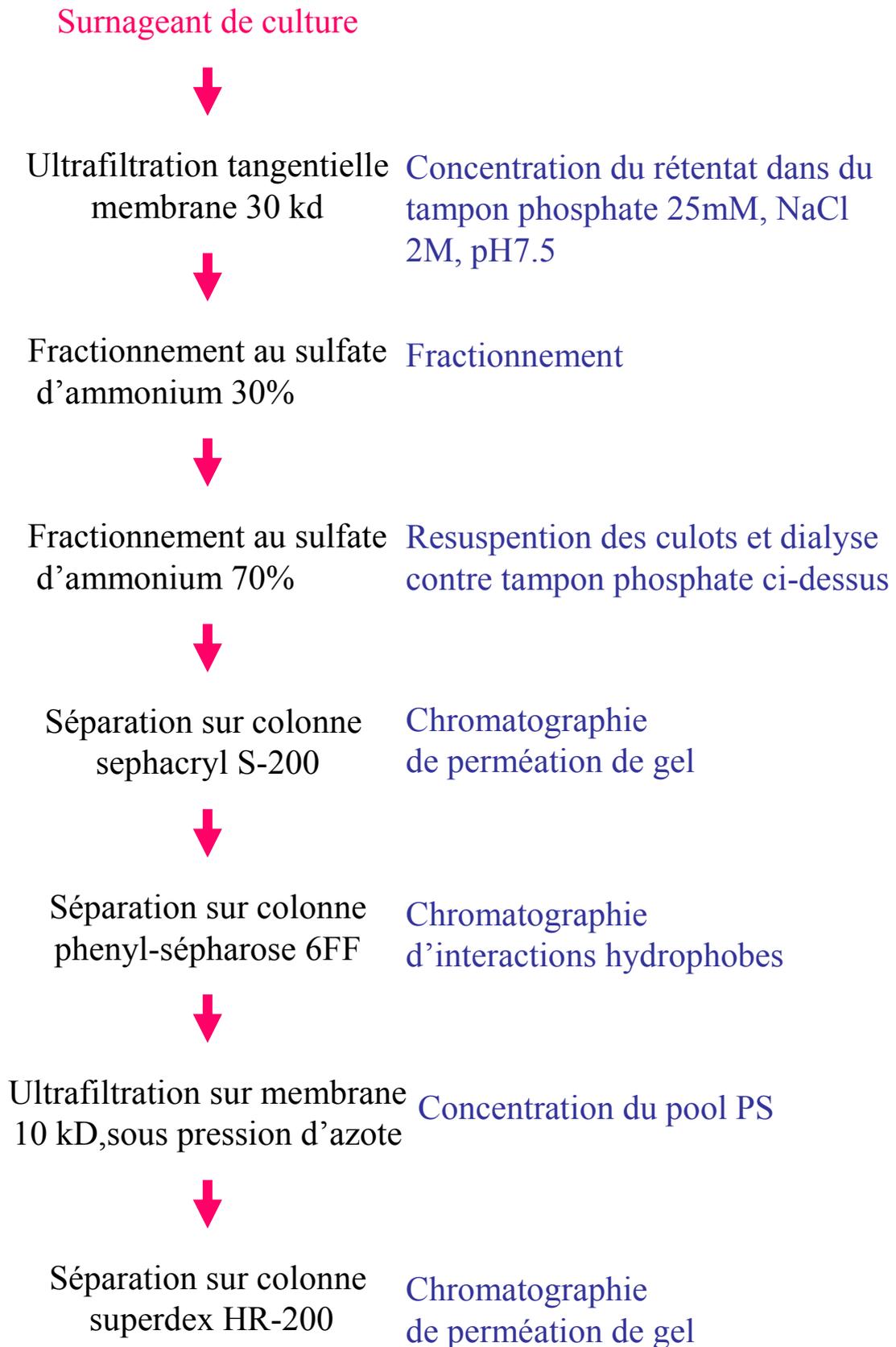


Figure 22 : Protocole schématique de la purification de la λ -carraghénase.

A bilan qualitatif

- | | |
|---|------------------------------------|
| 1- surnageant culture | 7- culot 70%SA |
| 2- filtrat 45µm | 8- pool sephacryl |
| 3- retentat ultrafiltration tangentielle 30kD | 9- pool phényl sépharose |
| 4- culot 30%SA | 10- rétentat ultrafiltration 10 kD |
| 5- surnageant 30% SA | 11- pool superdex |
| 6- surnageant 70%SA | |

B fractions superdex

- 1- pool phényl sépharose
- 2- retentat ultrafiltration 10 kD
- 3 à 7- fractions du pool superdex

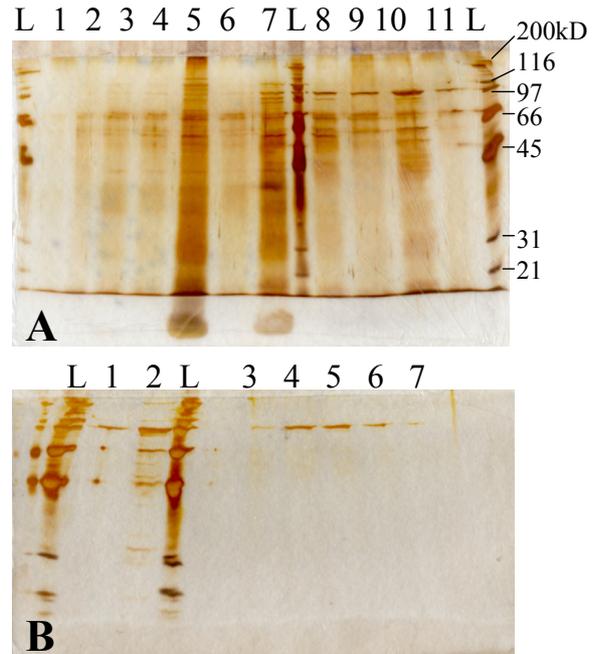


Figure 23 : Bilan qualitatif de la purification de la λ -carraghénase.

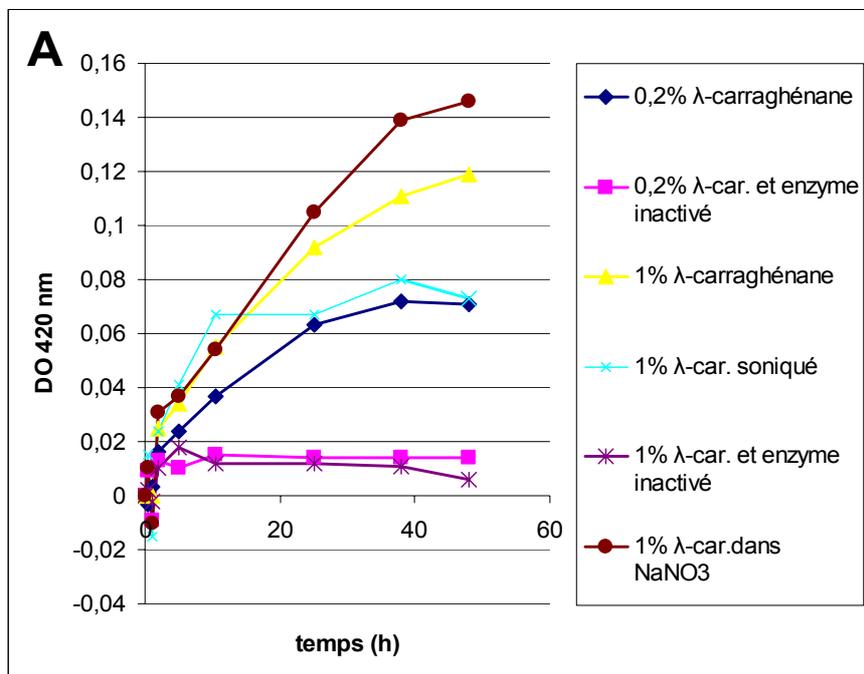


Figure 24 : Test de dosage des extrémités réductrices au ferricyanure (Kidby & Davidson, 1973). Cinétique d'hydrolyse par la λ -carraghénase de λ -carraghénane solubilisé dans différentes conditions (tampon phosphate 0,1M pH6,2 NaCl 0,1M, λ 0,2% p/v, 1%, ou 1% après sonication ; NaNO_3 0,1M, λ 1%). Les réactions sont menées à 30°C avec 1/100ème v/v d'enzyme.

2. Une méthode de dosage pour la λ -carraghénase

La forte viscosité du polymère engendre de fortes imprécisions de mesure de volume au delà de 0,2% (p/v) de concentration, ainsi que de gros problèmes d'homogénéisation des réactions d'hydrolyse. Les tests de dosages de l'activité d'hydrolyse de l'enzyme par une méthode classique de dosage des extrémités réductrices au ferricyanure (Kidby et Davidson, 1973) ont montré le manque de sensibilité de cette méthode appliquée à cette réaction. La figure 24 illustre en effet qu'à cette concentration, le maximum de signal, généré en fin de cinétique (autrement dit à complétion), ne dépasse pas 0,1 unité DO, empêchant toute mesure fiable de la vitesse initiale d'hydrolyse. Une solution reposait sur l'emploi de polymère à plus forte concentration afin d'allonger la phase linéaire initiale (le modèle michaelien est ici adopté par défaut bien que l'on puisse considérer la solution de polymère comme un système hétérogène à deux phases). Le problème de la viscosité a été contourné par sonication du polymère avant utilisation. Toutefois il apparaît que ce traitement engendre, à complétion, une variation de signal moindre que sans sonication, n'améliorant donc pas la sensibilité de la mesure. En d'autres termes, ce traitement a pour effet de réduire le nombre des actes catalytiques, probablement par suppression d'une partie des sites de coupure. Le meilleur résultat fut obtenu en remplaçant le tampon phosphate du substrat (0,1M pH 6,2, NaCl 0,1M) par du NaNO_3 0,1M. Le signal maximum n'excédait cependant pas 0,2 unité DO.

L'absence de méthode de dosage efficace des extrémités réductrices générées par l'enzyme m'a amené à mesurer indirectement son activité via la chute de viscosité se produisant lors de l'hydrolyse du λ -carraghénane. Cette méthode de mesure originale est développée dans l'article 1. La viscosité d'une solution de polymère dépend de plusieurs paramètres : i) la nature du polymère ; ii) la distribution en taille des molécules du polymère ; iii) les conditions physico-chimiques de la solution (température, pH, force ionique, nature des ions) ; iv) et enfin, les conditions de mesure dans le cas d'un liquide non newtonien (c'est le cas des solutions de λ -carraghénane). L'action de l'enzyme fait chuter la taille du polymère selon un mode d'action qui lui est propre, et consécutivement la viscosité décroît. Sa valeur est la

résultante macroscopique à un temps donné de la distribution en taille de l'ensemble des molécules de polymère. L'acte catalytique est donc appréhendé indirectement et au niveau macroscopique. Cette approche implique la conservation stricte des conditions de mesure et en premier lieu du lot de polysaccharide pour l'étalonnage et les dosages (chaque lot de λ -carraghénane variant en taille et homogénéité). Le suivi d'une cinétique enzymatique par viscosimétrie nécessite un équilibrage du système de mesure (rhéomètre à disque tournant à la surface du liquide), cette phase durant entre 30 sec et 1 min. L'idée développée ici consiste, par analogie à un dosage d'activité classique, à mesurer la vitesse initiale de décroissance de la viscosité (Fig. 25) qui devra être constante sur un intervalle de temps supérieur à celui de la de mesure. La cadence de points autorisée par le rhéomètre employé (jusqu'à un point par seconde) permet une approximation très fine de la vitesse instantanée de décroissance. Cette sensibilité rend ainsi possible la mesure des vitesses initiales sur un intervalle de temps supérieur ou égal à 30 sec. Dans ces conditions, une vitesse initiale n'est plus mesurable au-delà de $0,25 \text{ cP}\cdot\text{sec}^{-1}$.

L'étalonnage du système (Fig. 25) a été réalisé avec de l'enzyme pure à des dilutions variables. Le volume d'enzyme étant constant pour chaque réaction ($5\mu\text{l}$ dans $500 \mu\text{l}$ de milieu réactionnel), cela revenait à faire varier la concentration en enzyme. Il a montré que cette pente initiale (V_i) n'était pas proportionnelle à la concentration en enzyme ($[E]$), étant donné que $V_i = f([E])$ était mieux corrélé par une équation de type :

$$V_i = a \cdot [E]^b \quad (\text{I}), a \text{ et } b \text{ étant des constantes réelles} \quad (\text{Fig. 25})$$

Ainsi en posant $A = k \cdot [E]$, (I) donne :

$$A = a' \cdot V_i^{1/b} \quad (\text{II}), a' \text{ et } k \text{ sont des constantes réelles}$$

Une unité arbitraire d'activité (A) a été défini comme étant la quantité d'enzyme permettant une chute de viscosité de $0,1 \text{ cP}$ par seconde, ce qui permet de déduire a' , b étant défini empiriquement par l'étalonnage. La quantité d'activité d'un échantillon quelconque se déduit alors par la mesure de la vitesse initiale de décroissance de la viscosité dans les mêmes conditions, en reportant celle-ci sur l'équation II étalon (attention, ici la vitesse mesurée n'est pas proportionnelle au nombre d'unités d'activité).

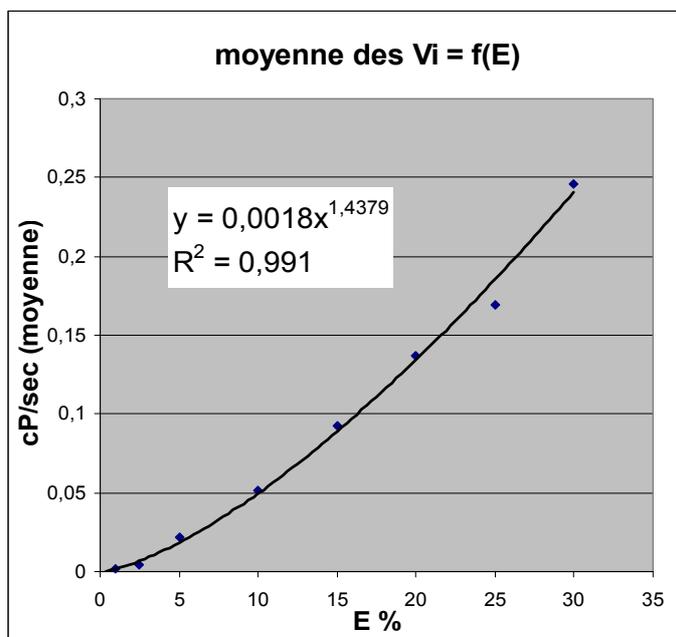
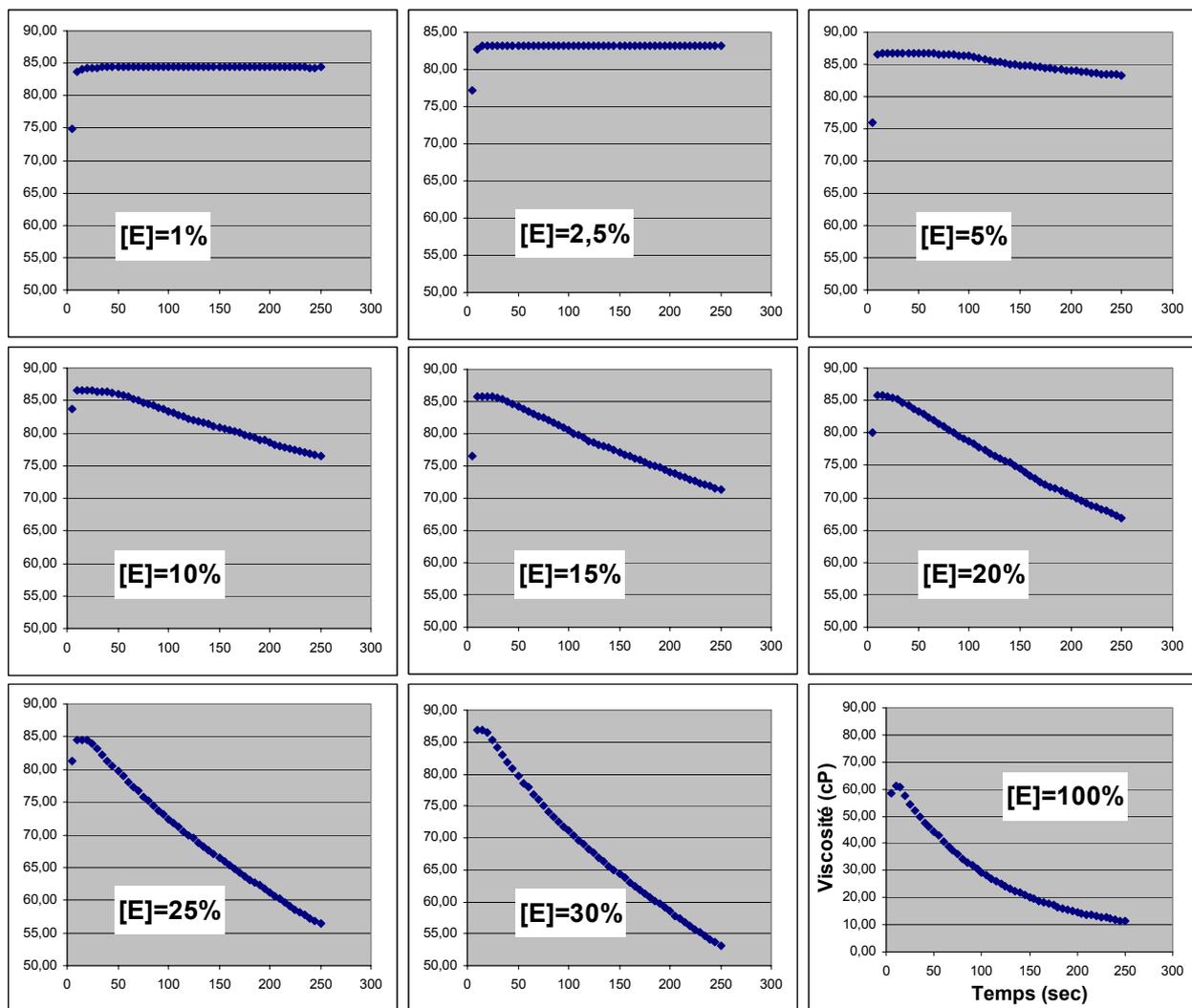


Figure 25 : Etalonnage d'une méthode de dosage de l'activité λ -carraghénase par viscosimétrie. Plusieurs cinétiques d'hydrolyse du λ -carraghénane (0,5% p/v) sont suivies par viscosimétrie, à raison d'un point toutes les 5 secondes. Les vitesses initiales de décroissance de la viscosité ont été mesurées après la phase d'équilibre (45 à 60 sec) pour 10 concentrations d'enzyme, de 0% à 30% (exprimées en % de la concentration maximale du lot utilisé). Cette vitesse est considérée comme initiale si elle est constante sur au moins 30 sec. Des moyennes sur trois répétitions ont été reportées en fonctions de la concentration en enzyme de façon à obtenir une courbe étalon.

Malgré les contraintes d'un tel système de mesure, il présente les avantages d'être rapide (<5 min/échantillon) et d'être très sensible ce qui économise les échantillons d'enzyme. Cependant, il ne permettra pas de définir les constantes catalytiques.

3. Etude de la spécificité de substrat

Afin de vérifier les spécificités de substrat et d'aborder les critères de reconnaissance de ce dernier par l'enzyme, j'ai mené plusieurs réactions d'hydrolyse sur du κ -carraghénane, du κ -carraghénane enrichi en μ -carraghénane, du ι -carraghénane, du ι -carraghénane enrichi en ν -carraghénane (*Euchema cottonii*). Plusieurs sources de λ -carraghénane ont également été comparées : 1, *Gigartina pistillata* et *acicularis* (Sigma); 2, *Gigartina skottsbergii* (CPKelco); 3, *Gigartina acicularis* (Sanofi Bioindustries); 4, *Chondrus crispus* (SKW biosystem); 5, *Gigartina pistillata* (extrait selon Amimi et al., 2001).

Il ressort de l'analyse par C-PAGE des produits d'hydrolyse (Fig. 26) que le κ -carraghénane, contenant ou pas du μ -carraghénane, n'est pas dégradé par l'enzyme. Par contre le ι -carraghénane utilisé contient des sites de coupures. La dégradation est beaucoup plus marquée quand la teneur en ν -carraghénane est plus importante. La séparation électrophorétique des produits révèle une échelle régulière de taille, suggérant la présence d'un motif élémentaire commun à tous ces produits, à des degrés de répétition divers. L'analyse des différents lots de λ -carraghénane montre la multitude des produits limites de faible taille (a priori < DP10). Cette grande variété de produits est principalement le fait d'une variation du degré de sulfatation couplée à des tailles légèrement différentes (DP2, DP3, DP4). Mais d'autres substitutions comme la présence de pyruvate (π -carraghénane) sont à considérer. Cette multitude illustre une certaine permissivité de reconnaissance de l'enzyme.

Le lot d'enzyme utilisé pour cette expérience correspond au pool issu de la chromatographie d'interactions hydrophobes concentré par ultrafiltration. On pourrait donc avancer que l'activité mise en évidence ne serait pas le fait de cet enzyme. Toutefois, je considère pour ma part, en accord avec la purification de Greer, que les

bandes supplémentaires visibles sur le gel SDS-PAGE de la figure 23 révèle une dégradation de la λ -carraghénase, étant donné que l'on retrouve des bandes similaires sur des échantillons d'enzyme pure après quelques semaines de conservation à 4°C. Par ailleurs, les profils de dégradation obtenus sur le λ -carraghénane de *G. skottsbergii* avec ce lot, l'enzyme pure ou l'enzyme surexprimé sont parfaitement identiques.

Le λ -carraghénane se distingue du ν -carraghénane uniquement par la position d'un substituant sulfate sur le galactose lié en 3, respectivement G2S et G4S. La dégradation du ν -carraghénane enrichi en λ -carraghénane par l'enzyme est de ce fait très intéressante, d'autant qu'elle semble générer des produits d'une série homogène. Aucun de ces derniers n'étant caractérisé, il m'est impossible de me prononcer sur la reconnaissance par la λ -carraghénase de motif ν -carraghénane. D'autant qu'on ne peut complètement exclure la présence de quelques motifs λ dans ce polymère. La répartition des motifs ν dans le ν -carraghénane, aléatoire ou en bloc, est, à l'heure actuelle, toujours l'objet de spéculation. Si la λ -carraghénase reconnaît effectivement le ν , la structure de ces produits pourrait apporter de précieux renseignements sur l'agencement de cet hétéropolymère largement exploité dans l'industrie.

4. Clonage et séquençage du gène de la λ -carraghénase

L'obtention de la sonde nucléique permettant le criblage de la banque génomique est également décrite intégralement dans l'article 1. La synthèse de cette sonde de 905 pb ainsi que son séquençage ont été réalisés antérieurement à ce travail de thèse par Chabeau (1995) au cours de son DEA. Cette synthèse repose sur une amplification par PCR d'une partie du gène cible à partir d'amorces dégénérées. Ces dernières sont déduites de micro-séquences peptidiques obtenues selon la méthode d'Edmann. La protéine micro-séquencée est extraite d'un profil de séparation SDS-PAGE par découpage de la bande correspondante. Ici cette bande avait été découpée à partir d'un profil de séparation d'un pool non pur de λ -carraghénase, profil dont aucune image n'est à disposition.

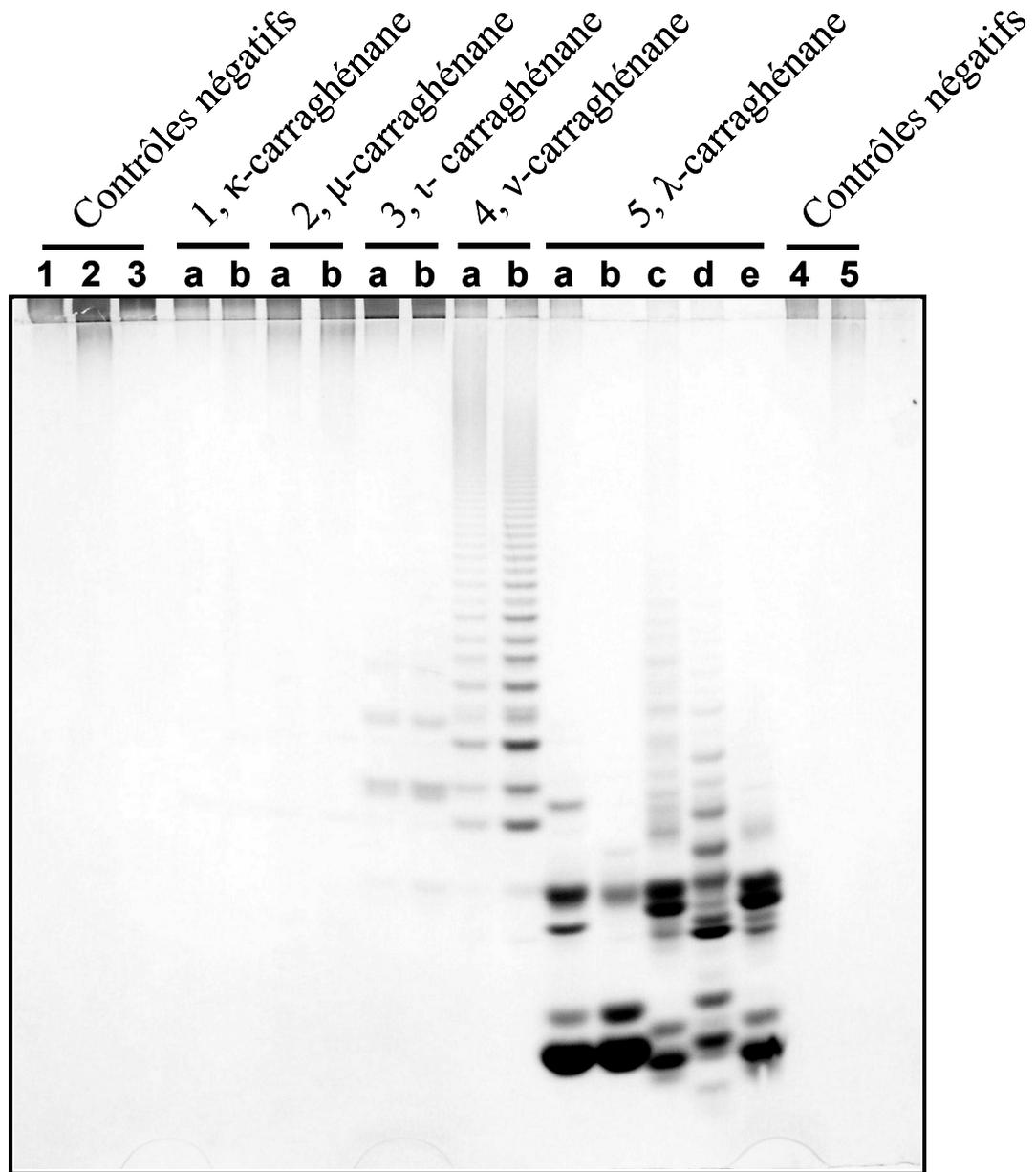


Figure 26 : Etude de la spécificité de substrat de la λ -carraghénase. Analyse par C-page des produits d'hydrolyse (a, 24 h ; b, 55 h) des substrats suivant (extrait de *Euchema cottonii*): 1, κ -carraghénane ; 2, κ -carraghénane enrichi en μ -carraghénane ; 3, ι -carraghénane ; 4, du ι -carraghénane enrichi en ν -carraghénane. Plusieurs sources de λ -carraghénane (5) ont également été comparées : a, *Gigartina pistillata* et *acicularis* (Sigma); b, *Gigartina skottsbergii* (CPKelco); c, *Gigartina acicularis* (Sanofi Bioindustries); d, *Chondrus crispus* (SKW biosystem); e, *Gigartina pistillata* (extrait selon Amimi et al., 2001). Les contrôles négatifs correspondent à des réactions menées avec de l'enzyme dénaturé par ébullition. Tous les substrats ont été incubés à 0,1 % p/v et solubilisés dans du NaNO_3 0,1 M, avec 2% v/v d'enzyme du pool concentré issue de la chromatographie d'interactions hydrophobes.

J'ai pris connaissance de ce fait après avoir isolé et séquencé le gène ciblé par cette sonde. Le risque important d'avoir utilisé une sonde aspécifique m'a contraint à mettre au point la purification à homogénéité de cet enzyme (c.f. ci-dessus), de façon à vérifier que le gène cloné correspondait effectivement à la bonne protéine. L'acquisition de la masse précise de cette protéine et/ou l'obtention de micro-séquences par spectrométrie de masse auraient pu confirmer la séquence obtenue. Cependant, la correspondance de taille (chromatographie de perméation de gel et gel SDS-PAGE) a été jugée suffisante bien que ce ne soit pas une preuve absolue de la pertinence de cette sonde nucléique. L'activité λ -carraghénase de la protéine recombinante nous a donné raison d'avoir pris ce risque.

Cette sonde m'a tout d'abord servi à chercher le nombre de copie du gène ciblé dans le génome de *Pseudoalteromonas carrageenovora*. Un Southern blot a été réalisé sur l'ADN génomique de *P. carrageenovora* préalablement digéré par diverses endonucléases. Il apparaît que la sonde s'hybride sur les fragments *HindIII/BamHI*, *EcoRV/HindIII* et *HindIII* d'environ 1500 pb et sur les fragments de 3000 pb *BamHI/EcoRV* et *EcoRV* (Fig. 27). Ce résultat suggère que le gène de la λ -carraghénase est présent en un seul exemplaire dans le génome de cette bactérie.

C'est cette même sonde que j'ai utilisé pour cribler une banque génomique de *P. carrageenovora*. Quatorze clones positifs (Tab. 6) ont été détectés dans la banque, soit 0,27% du nombre de clones total (approx. 6000). Après ré-isolément sur boîte, un second criblage de ces 14 clones a confirmé le premier criblage. Compte tenu de la taille de la λ -carraghénase (98 kDa, sur gel SDS-PAGE), son gène a été évalué à environ 3 kb. J'ai donc entrepris de cartographier physiquement les 3 plasmides recombinants, pLA38, pLA45 et pLA53, dont les inserts font respectivement 4, 6 et 5 kb (cartes de restriction, Fig. 28). Ces cartes mettent en évidence des régions communes. Les correspondances entre ces trois clones en font un ensemble cohérent et confirment la présence de tout ou partie du gène λ -carraghénase sur ces inserts.

Après avoir réalisé un Southern blot sur différentes digestions de pLA45 (Fig. 29), il apparaît que la sonde N-terminale de 905 pb s'hybride sur la région comprise entre les sites *HindIII* (3400) et *HindIII* (4800). Le plasmide pFU11 choisi comme contrôle négatif est issu de la banque génomique de la souche fucanolytique SW5. La

sonde n'étant pas orientée, seul le clone pLA45 offrait la certitude de contenir le gène entier. C'est donc par ce plasmide que j'ai décidé de débiter le séquençage. Afin de connaître l'orientation du gène sur l'insert, le fragment *HindIII* (3400) / *SpeI* (4300) a été sous-cloné (pALA45, Fig. 30). Le séquençage de ce fragment a permis d'orienter le gène, de confirmer l'intégralité de ce dernier dans pLA45 et de déterminer les autres sous-clones à réaliser (Fig. 30). Le sous-clone pELA45, trop long (1.8 kb) n'a pu être séquencé entièrement à partir de ces extrémités. J'ai alors établi des amorces internes qui m'ont servi à scinder cet insert en trois amplicons. La partie centrale de ce fragment ainsi que ces deux extrémités ont été amplifiées par PCR. Puis trois clones distincts de chaque produit d'amplification ont été séquencés. Par ailleurs, du fait de l'absence de chevauchement entre les fragments F et A puis B et D, se posait le problème de la pertinence de leur raboutage (Fig. 30). Ces deux régions ont fait l'objet d'un séquençage non automatique à chaud qui a démontré la parfaite juxtaposition des fragments F et A d'une part, et B et D d'autre part.

Les sous-clones pALA45 et pBLA45 ont été séquencés deux fois dans les deux sens, pDLA45 huit fois à partir de *HindIII* et quatre fois à partir *BstXI*, pFLA45 trois fois dans les deux sens, l'extrémité *PstI* de pELA45 cinq fois dans un sens et six fois dans l'autre, l'extrémité *ClaI* de pELA45 six fois dans un sens et sept dans l'autre, et le cœur de pELA45 neuf fois dans les 2 sens.

Enfin, l'insert de pLA45 a été séquencé dans son entier par « gene walking » de façon à compléter et confirmer le contig précédemment obtenu. Le détail de ces séquençages est présenté figure 31.

5. Analyse de la séquence

5.1. Analyse du contig

Une fois nettoyées et assemblées, ces séquences constituent un contig de 5665 bases (annexe 1) confirmant la carte de restriction établie. Ce dernier contient deux phases ouvertes de lecture (ORF : open reading frame ; Fig. 31) complètes. La première de 2826 bases (942 acides aminés) code pour la protéine d'intérêt, précédée des motifs

Tableau 6 : Tailles des inserts des clones positifs.

<i>clones</i>	<i>pLA5</i>	<i>PLA6</i>	<i>pLA16</i>	<i>pLA17</i>	<i>pLA20</i>	<i>pLA271</i>	<i>pLA272</i>
Taille de l'insert (kb)	20 à 15	11	21	8	10 à 15	5	6

<i>clones</i>	<i>pLA38</i>	<i>pLA45</i>	<i>pLA461</i>	<i>pLA462</i>	<i>pLA53</i>	<i>pLA55</i>	<i>pLA57</i>
Taille de l'insert (kb)	4	6	3.5	5	5	20	11

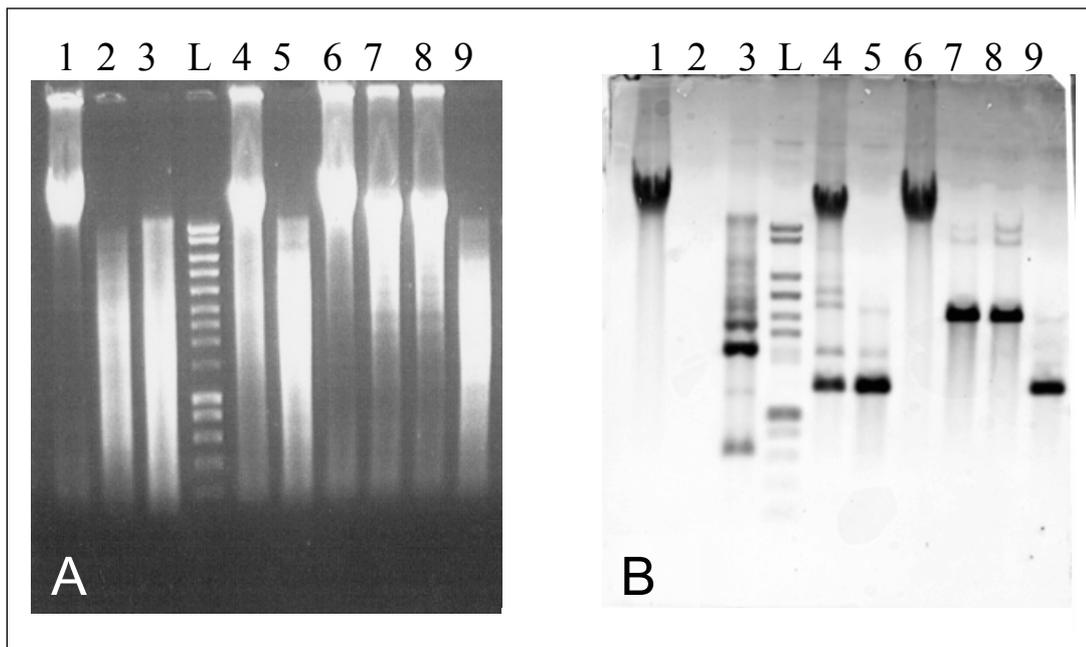


Figure 27 : Recherche du nombre de gène λ -carraghénase dans le génome de *Pseudoalteromonas carrageenovora* par Southern blot (A : gel ; B : membrane).

1 : ADN génomique non digéré ; **2** : control négatif par digestion de l'ADN génomique de la souche fucanolytique SW5 par *Sau3AI* ; **3** : digestion par *Sau3AI* ; **4** : digestion par *HindIII* ; **5** : double digestion par *HindIII/BamHI* ; **6** : digestion par *BamHI* ; **7** : double digestion par *BamHI/EcoRV* ; **8** : digestion par *EcoRV* ; **9** : double digestion par *EcoRV/HindIII* ; **L** : Smart ladder.

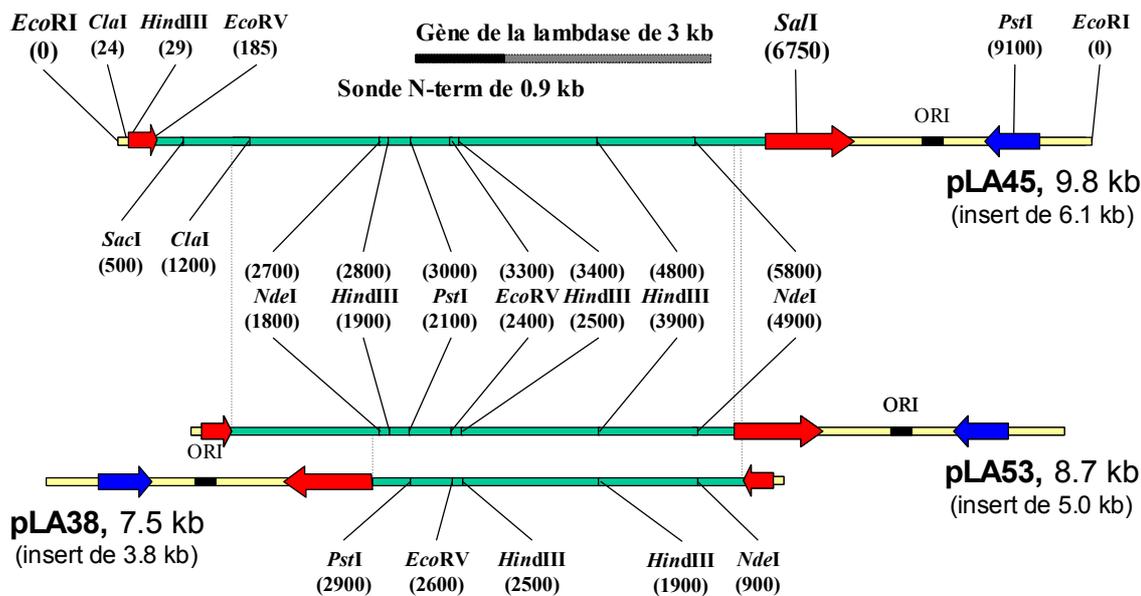


Figure 28 : Alignement des cartes de restriction des plasmides pLA38, pLA45 et pLA53. Les inserts sont représentés en vert, le jaune correspondant au vecteur pAT153 dont en bleu le gène de la β -lactamase et en rouge celui de la résistance à la tétracycline.

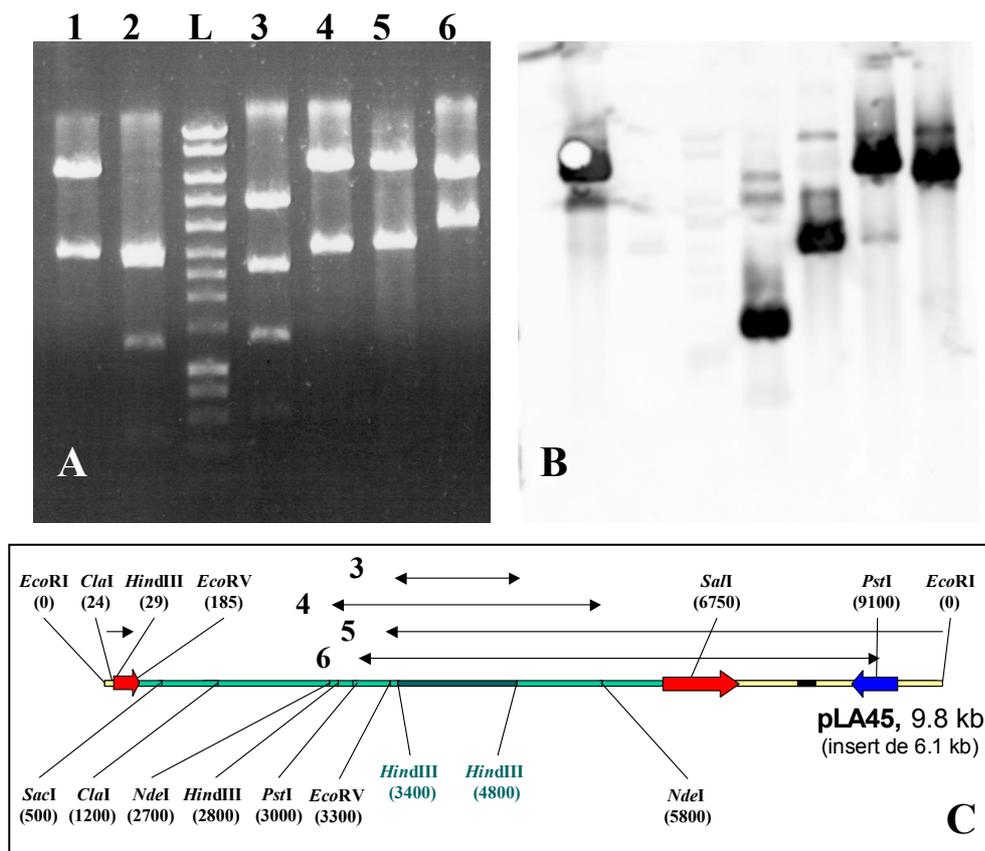


Figure 29 : Localisation de la sonde sur l'insert de pLA45 par Southern blot (A : gel ; B : membrane ; C : carte de restriction de pLA45).

1 control positif: pLA45 digéré par *EcoRI/SalI*; 2 control négatif: pFU11 digéré par *EcoRI/SalI*; 3 : pLA45 digéré par *HindIII*; 4 : pLA45 digéré par *NdeI*; 5 : pLA45 digéré par *EcoRV*; 6 : pLA45 digéré par *PstI*; L : Smart ladder. Les doubles flèches sur le schéma C indiquent les fragments de restrictions hybridant la sonde.

RBS (« Ribosomic Binding Site ») de type epsilon et Shine-Dalgarno. La seconde, située sur le brin opposé, code pour une protéine de 297 résidus (891 nucléotides). De toutes évidence (BLASTp sur Genbank ainsi que la recherche automatique de domaine du NCBI RPS-BLAST sur CDD) cette protéine appartient à la famille des régulateurs transcriptionnels LysR (COG 0583 ; PFAM PF00126 et PF03456), liés au métabolisme des acides aminés et donc sans lien direct avec notre gène cible. Les 250 meilleurs scores BLAST sur Genbank proviennent d'homologues de γ -protéobactéries (famille des Vibrionaceae).

Le début de l'insert de pLA45 présente une ORF amont supposée partielle de 227 résidus. Elle est colinéaire à celle de la λ -carraghénase mais la distance les séparant est relativement élevée (602 nt). En outre, BLASTp sur Genbank ainsi que la recherche automatique de domaine RPS-BLAST du NCBI sur CDD permet d'associer cette protéine à une famille d'hydratase/decarboxylases (COG 3971), pouvant intervenir dans le métabolisme bactérien de métabolites secondaires. L'hypothèse d'un opéron est donc peu probable. J'ai également recherché des séquences promotrice et terminatrice de la transcription de part et d'autre du cadre de lecture d'intérêt. Ces séquences putatives sont représentées en annexe 1.

5.2. Analyse de la séquence protéique de la λ -carraghénase

La traduction selon le code génétique universel de ce cadre de lecture donne la séquence d'une préprotéine de 105638 Da dont le pH isolélectrique serait a priori acide (variation selon les méthodes de calcul: Mwcalc donne 5.91, ProtParam tool donne 5.54, et Abi donne 5.37 ; outils disponibles sur le serveur INFOBIOGEN).

5.2.1. Le peptide signal

Cet enzyme étant présent dans les surnageants de culture, j'ai procédé à la recherche d'un peptide signal. Le programme de prédiction Psort V6.4 suggère fortement la présence d'un peptide signal de type prolipoprotéine avec un clivage entre Gly17 et Cys18. Cette séquence satisfait par ailleurs à tous les consensus propres à ce type de peptides signaux: i) coupure avant un résidu cystéine ; ii) présence des résidus

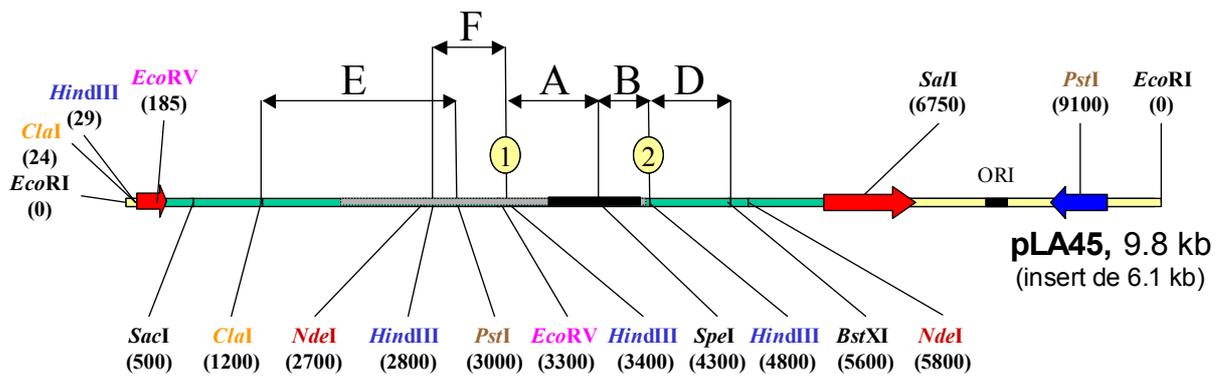
glycine en -1, isoleucine en -2, et leucine en -3 et -4 ; iii) absence de résidus chargés entre -5 et -11 (réf. Prosite: PDOC00013). En outre, le programme SignalP V1.1, qui ne considère pas spécifiquement ce type de peptides signaux en raison de leur trop forte spécificité, prévoit également un peptide signal dont le clivage aurait lieu devant le résidu 18. Dans ce cas, cet enzyme serait ancré à la membrane interne et/ou externe de la cellule via la greffe d'un glycéride sur la cystéine terminale durant la translocation après clivage du peptide signal (signal peptidase II). Un tel système d'adressage est connu chez plusieurs familles bactériennes appartenant notamment aux γ -protéobactéries.

Toutefois, cette protéine a été purifiée à partir du surnageant de culture. Et le microséquençage de son extrémité N-terminale a montré un premier résidu correspondant au résidu 25 de la séquence. De plus, la version V2.0 de SignalP prédit également un clivage à cet endroit, une solution qui était déjà envisagé par la version V1.1. Enfin, la longueur moyenne des peptides signaux identifiés chez les bactéries gram- est de 25.1. J'ai donc privilégié cette dernière hypothèse qui est validée par les résultats expérimentaux, mais sans toutefois pouvoir exclure une prise en charge de la préprotéine par deux systèmes d'excrétion différents. En effet, une expérience, que je ne développe pas ici faute d'avoir été répliquée, à consister à extraire les lipoprotéines d'une culture induite de *P. carrageenovora*. Cette fraction protéique a montré de l'activité λ -carraghénase contrairement à une extraction contrôle sur une fraction de l'enzyme purifiée du surnageant de culture.

5.2.2. Recherche des domaines de la λ -carraghénase

Cette séquence, confrontée dans son entier par le programme BLAST à GENBANKnr et PAT (serveur NCBI), ne présente aucune similitude franche avec les séquences qui y sont répertoriées à ce jour (E-value > 0,002, sur des hits < 150 résidus). Cette protéine ne semble donc pas avoir d'homologue connu. De même, la recherche par RPS-BLAST de domaines/motifs répertoriés dans CDD (NCBI) ou Interproscan sur Interpro (EBI) n'a donné aucun résultat probant.

Cette protéine de grande taille (942 résidus) est très vraisemblablement organisée



FRAGMENTS	A	B	D	E	F
SITES	<i>HindIII</i>	<i>SpeI</i>	<i>HindIII</i>	<i>ClaI</i>	<i>HindIII</i>
FLANQUANTS	<i>SpeI</i>	<i>HindIII</i>	<i>BstXI</i>	<i>PstI</i>	
TAILLES (kb)	0.9	0.5	0.8	1.8	0.6

Figure 30 : Présentation des sous-clones réalisés pour l'obtention de la séquence entière de la λ -carraghénase. Les régions 1 et 2 ont fait l'objet d'un séquençage non automatique.

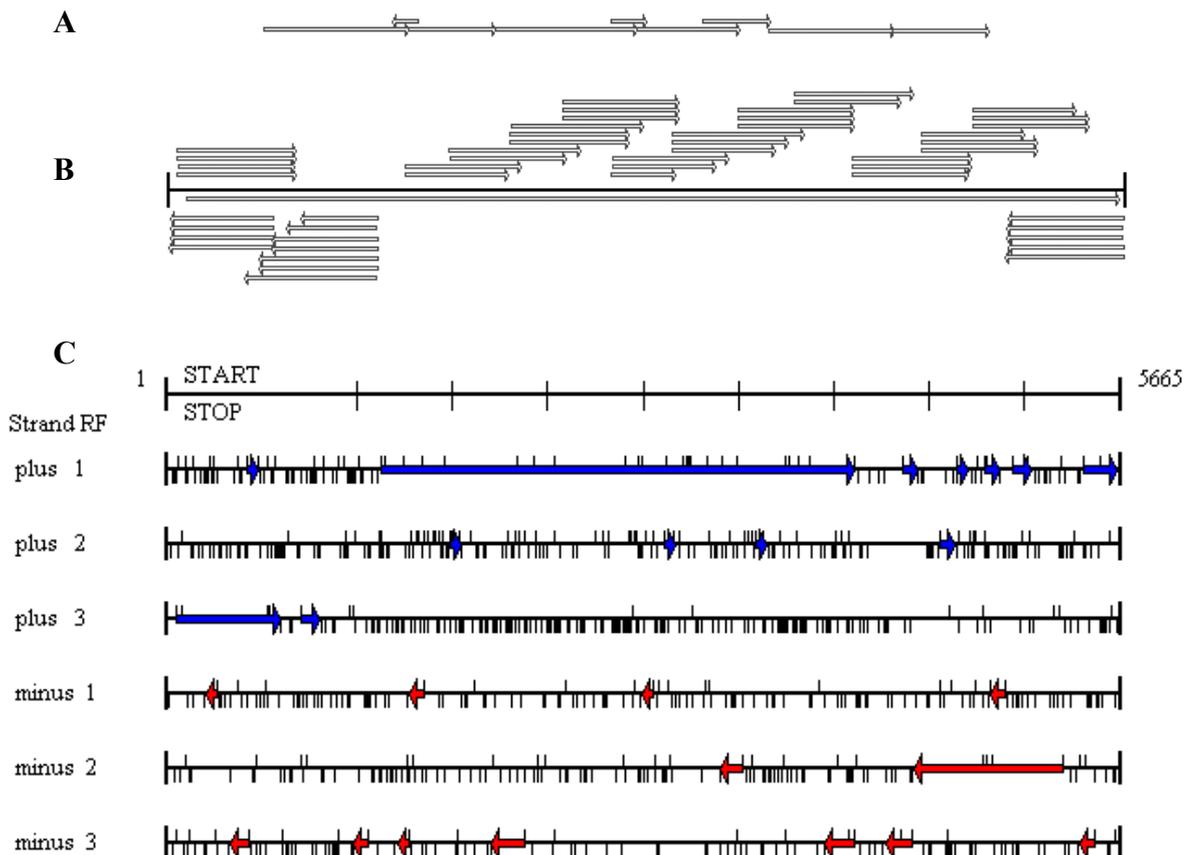


Figure 31 : Présentation des séquençages réalisés sur pLA45 et cadres ouverts de lectures (ORF) obtenus sur l'ensemble de son insert. A : séquences des sous-clones et de leurs jonctions ; B : séquençage en gene-walking ; C : traduction du contig entier (schéma obtenu à l'aide du logiciel DNAMAN en utilisant le code génétique universel, la taille des ORF étant limitée à 25 résidus).

en plusieurs domaines ou modules distincts. L'un des objectifs de l'analyse de séquence est d'identifier, ou du moins de localiser, ces domaines de manière à percevoir l'architecture de la protéine, l'attribution de fonctions biologiques à ces domaines ne venant qu'en second lieu. Pour ce faire j'ai réalisé des recherches de similitudes par PSI-BLAST sur des portions plus ou moins larges de la séquence.

Un hit c'est avéré intéressant sur l'extrémité C-terminale de la λ -carraghénase. Les 77 derniers résidus révèlent 38% d'identité de séquence avec l'extrémité C-terminale d'une cyclomaltodextrinase putative (87% d'identité avec un enzyme caractérisé) issue de *Thermotoga maritima* et membre de la famille GH13 des glycosidases. Tous les enzymes de cette famille possèdent un domaine tout bêta C-terminal (se repliant en clé grecque ; ref. Interpro IPR006048, SCOP b.71.1.1) dont la longueur varie de 70 à 100 résidus. Il suit immédiatement le domaine catalytique (ref. Interpro IPR006046) de type tonneau $(\alpha/\beta)_8$ de 350 à 400 résidus. C'est ce domaine C-terminal qui s'aligne avec celui de la λ -carraghénase. PSI-BLAST, ne partant que de deux séquences au départ, converge lentement mais de façon très homogène en 5 itérations sur ce domaine.

L'alignement de séquences homologues provenant d'enzymes de structure connue (Fig. 32) avec cette région C-terminale de la λ -carraghénase (89 résidus) met en évidence des analogies consistantes malgré un faible taux d'identité (9 à 20%). Différentes prédictions de structures secondaires (PsiPRED, SamT99-2D, ProfSec) révèlent, entre les résidus 850 et 930, un ensemble de 8 brins β se conformant à la position des éléments de structures secondaires dans cet alignement. Ce domaine a la particularité d'être systématiquement associé à un domaine catalytique de la famille GH13. Dans CglA, les 400 résidus amonts à cette région ne présentent aucune homologie détectable par BLAST ni avec cette famille ni avec aucune autre. Toutefois, la prédiction de structures secondaires y met en évidence une alternance plus ou moins régulière d'hélices α et de brins β , mais sans que les blocs α/β caractéristiques des TIM barrels n'y apparaissent clairement. La modélisation structurale des 480 derniers résidus réalisée via le meta-server 3D-JURY (homology modeling et/ou threading) désigne comme la plus compatible (avec un score légèrement significatif de 58 pour un seuil de 50) la structure de la α -fucosidase de *Thermotoga maritima* (1hl8), qui est

constituée d'un TIM barrel catalytique suivi, comme dans la famille GH13, d'un domaine tout bêta à 8 brins (Sulzenbacher et al., 2004). Parmi les meilleurs modèles proposés (mais avec des scores non significatifs) la majorité sont « enfilés » (threading modeling) sur des TIM barrels de glycosidases (ref. SCOP c.1.8) dont on sait qu'ils constituent le fold le plus courant parmi ces enzymes (66% des séquences, et 40% des familles ; Nagano et al., 2002; Tab. 5). Enfin, soumises séparément à modélisation structurale, ces deux régions procurent des résultats convergeant vers les mêmes repliements mais avec un degré de confiance moindre que lors de leur soumission groupée. En tout état de causes, je suis confronté à la limite de fiabilité de ces outils prédictifs. Pour conclure sur cette région, il est intéressant de rappeler que les β -agarases de la famille GH50 ont été prédites comme adoptant également un TIM barrel du fait de leurs similitudes de séquence avec la famille GH42 (Rigden et al., 2003).

L'étude de l'extrémité N-terminale de la λ -carraghénase a révélé des similitudes de séquence significatives avec des domaines de type β -propeller. Ces domaines sont constitués de 4 à 8 feuillets β de 3 à 5 brins antiparallèles chacun, la structure canonique ayant 6 feuillets à 4 brins (Fig. 33B et 33C). Ces structures circulaires se caractérisent généralement par la forte conservation de ces unités, ce qui permet de les considérer comme des répétitions d'en moyenne 50 résidus. La cohésion de l'ensemble est due au décalage qui existe entre la succession des brins sur la séquence et leur participation aux feuillets. C'est en effet souvent le premier brin de la séquence qui vient compléter le dernier feuillet, assurant par un effet « velcro » la fermeture de l'ensemble. Ce support structural est codé par des séquences très variées. Il est dédié à des fonctions diverses, catalytiques ou non (Guruprasad et Dhamayanthi, 2004 ; Jawad et Paoli, 2002 ; Fulop et Jones, 1999). Une protéine issue *Methanosarcina acetivorans* str. C2A (AAM04933), longue de 495 résidus, présente notamment une succession de trois alignements sur les 150 premiers résidus de la λ -carraghénase (peptide signal non inclus). Les autres alignements dégagés par BLAST se concentrent sur cette région avec des décalages d'environ 50 résidus dans les deux sens. En outre, les protéines mises en évidence, du fait qu'elles contiennent un domaine constitué de répétitions marquées, s'alignent sur cette région à plusieurs reprises. L'utilisation du programme PSI-BLAST donne une convergence après cinq itérations vers un pool de séquences

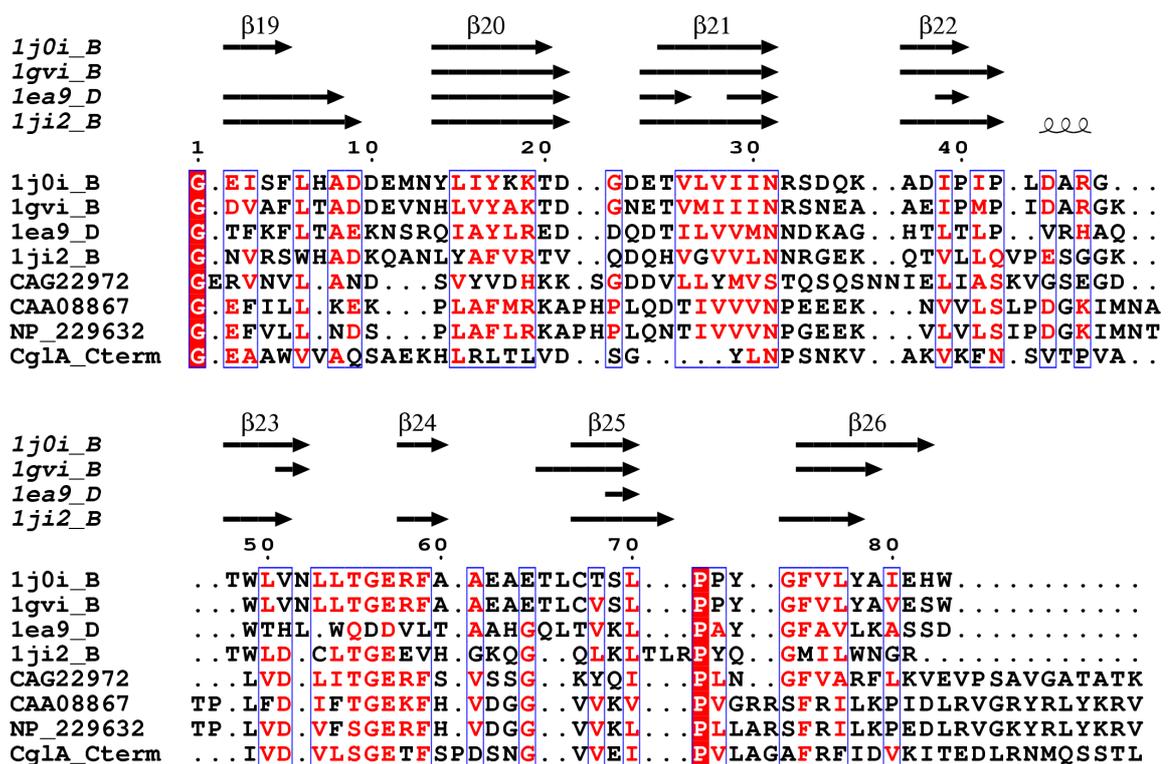


Figure 32 : Etude de l'extrémité C-terminale de la λ-carraghénase (CglA). Mise en évidence des similarités avec le domaine C-terminal qui suit le domaine catalytique de 7 enzymes de la famille GH13 des glycosidases. Cet alignement inclus les meilleurs hits fournis par BLAST contre Genbanks et la PDB. Les structures secondaires connues sont placées au dessus de l'alignement. (figure réalisée à l'aide de Esprit 2.2)

homogène contenant des β -propellers (type WD, FG-GAP, BNR, et PQQ). Ce résultat s'explique par les caractéristiques très marquées de ces séquences. Les alignements qui en découlent ne permettent cependant pas d'établir des caractéristiques communes franches avec la séquence de la λ -carraghénase. Par ailleurs, la recherche de répétitions à l'aide du programme RADAR n'a donné aucun résultat probant concernant ni cette région en particulier ni la séquence entière (les répétitions de bêta propeller pouvant être dispersées). La prédiction de structures secondaires sur ce fragment (les 370 premiers résidus) de la séquence donne une composition en brins β largement majoritaire, appuyant l'hypothèse d'une structure tout β comme le sont ces propellers. Soumis aux outils de reconnaissance de repliement, les 400 premiers résidus, fournissent des scores hautement significatifs (3D-JURY donne des scores >130) avec des modélisations se basant sur des β -propellers à 6 ou 7 feuillets. La figure 33A présente la recherche par HCA de ces feuillets. Six répétitions structurales hypothétiques y sont alignées.

Il existe un nombre réduit de repliement tout β de cette dimension (> 250 résidus) recensés dans CATH. Les β -propellers y prennent une large part, mais les hélices β (ex : ι -carraghénases de la famille GH82) et les β -jelly roll (ex : κ -carraghénases de la famille 16) en font également parties. Néanmoins, aucun des consensus propres aux hélices β n'a été retrouvé dans CglA. Par ailleurs, si la juxtaposition de deux ou trois domaines tout β est naturellement envisageable pour expliquer la taille de cette région, à mon sens, les outils d'analyses employés auraient été à même de les discriminer (cf. analyses des domaines cadherin like de la fucanase FcnA, Chapitre 3).

Si l'absence de répétitions aisément identifiables empêche une prédiction claire de ces propellers, elle n'est cependant pas rédhibitoire quant à leur existence. En effet, le siège de la diversité entre propellers se situe au niveau des boucles joignant les brins aux autres, puisqu'elles sont responsables de la fonctionnalité de ces structures. Celles-ci varient en composition et en longueur à tel point que certains β -propellers présentent des insertions entre les brins suffisamment importantes pour constituer soit des structures secondaires, type hélice α ou brin β , soit des domaines à part en entière. La dispersion des brins le long de la séquence, ainsi occasionnée, empêche l'identification

automatique, ou manuelle, de répétitions. Ainsi, la structure d'une glycoside hydrolase de la famille 43 (Arabinanase, clan F) révèle un bêta propeller de 5 feuillets (Nurizzo et al., 2002) qui est à ce point atypique, qu'il ne contient pas de répétitions, que le schéma d'assemblage des brins dans les feuillets n'a été observé que deux fois (absence de fermeture type velcro du propeller), et que les feuillets ne présentent pas tous le même nombre de brins. Par ailleurs, le repliement en β -propeller d'une nouvelle glycoside hydrolase de la famille 32 (invertase, clan J) a également été tout récemment démontré (Alberto et al., 2004). Il ne possède là encore aucune des répétitions de séquence qui les caractérisent habituellement. Enfin, l'unique structure de la famille 74 est un β -propeller à sept feuillets qui présentent de très faibles répétitions de séquence détectées par l'alignement structural des feuillets (Yaoi et al., 2004).

Trois clans (E, F, et J) et neuf familles de glycoside hydrolases présentent ce type de structure (33, 34, et 83 ; 43 et 62 ; 32 et 68 ; 58 ; 74 ; Tab. 5). Pour autant un tel repliement n'a pas nécessairement une fonction catalytique, et l'on pourrait alors envisager un rôle soit d'ancrage au substrat ou à la cellule avec une fonction de type lectine, soit de localisation à la membrane externe de la bactérie via des interactions protéines-protéines (de nombreux hits de BLAST sur cette région proviennent de larges protéines procaryotiques annotées comme « protéine de surface cellulaire »).

5.2.3. Conclusions de l'analyse de séquence

L'absence d'homologie de séquence flagrante (>20% d'identité) sur toute ou partie de la séquence suggère l'appartenance de cet enzyme à une nouvelle famille de glycoside hydrolase. En tout état de cause, étant donnée l'absence des consensus catalytiques, cette enzyme ne s'apparente pas aux familles ni 16 ni 82 des glycoside hydrolases qui contiennent les κ - et ι -carraghénases.

L'analyse de séquence n'a pas permis de dégager de façon suffisamment fiable les différents domaines qui composent cette protéine. Son architecture reste de ce fait relativement floue. La seule certitude est la présence à l'extrémité N-terminale d'une large région tout bêta qui pourrait former un domaine unique, ainsi qu'un domaine plus

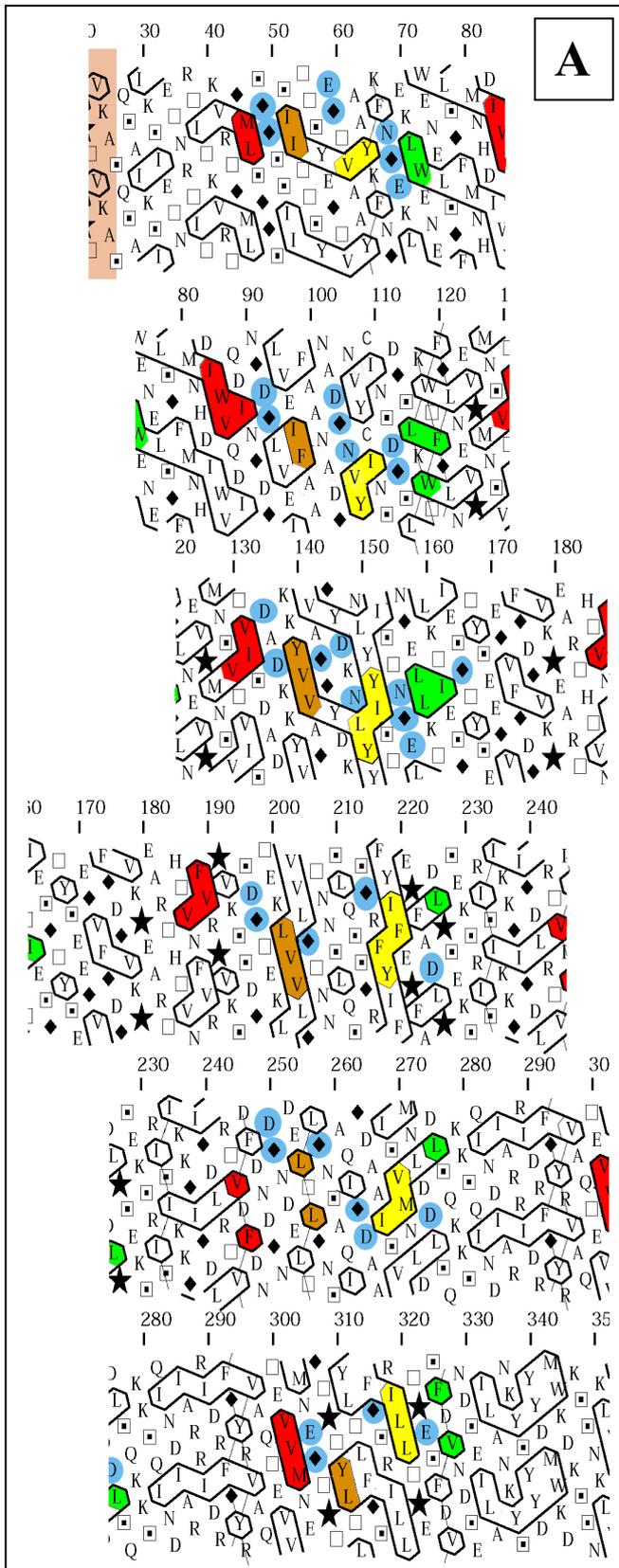
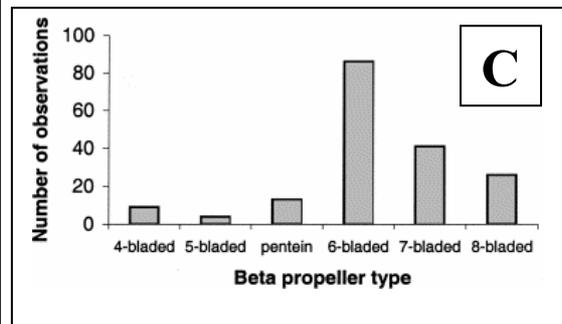
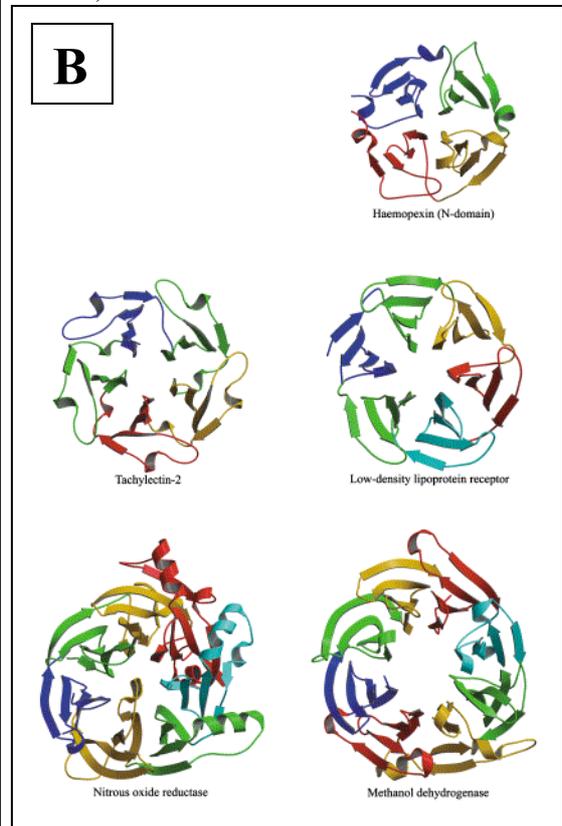


Figure 33 : Hypothèse d'un β -propeller à l'extrémité N-terminale de la λ -carraghénase A : Recherche par la méthode graphique HCA de répétitions structurales correspondant à des feuillets β de 4 brins. Six répétitions successives sont proposées. B : Exemple de β -propellers de 4, 5, 6, 7 et 8 feuillets (extrait de Jawad & Paoli, 2002). C : Répartition des différentes structures de β -propellers (extrait de Guruprasad & Dhamayanthi, 2004).



court tout bêta également en C-terminal. La région les séparant ne fournit aucun hit et serait composée d'un mélange α/β . C'est cette organisation tripartite qui sera retenue comme architecture la plus probable de CglA.

5. La surexpression de la λ -carraghénase

Les domaines de la protéine n'ayant pu être bornés convenablement les tentatives de surexpression ont essentiellement portées sur la protéine entière. Peu de constructions ont été réalisées (technologie Gateway d'Invitrogen) : i) la protéine native (pDEST14); ii) fusionnée avec tag histidine 6x de 22 résidus en N-terminal (pDEST17); iii) fusionnée avec un tag histidine 6X de 6 résidus en C-terminale.

De nombreux paramètres d'expressions ont été testés :

- souches : BL21 DE3, BL21 DE3 pLysS, Rosetta, Rosetta pLysS.
- milieux : LB, M9, autoinducteur (F. W. Studier, communication personnelle).
- températures : 12°C, 15°C, 20°C, 25°C, 30°C et 37°C.
- inducteur : IPTG à 0,1 0,2 0,5 et 1 mM
- additifs : éthanol (stimulation de l'activité des chaperones), glucose (apport d'énergie après induction et arrêt de l'induction), antibiotique (arrêt de la synthèse protéique après induction ; ex. chloramphénicol).

La protéine a systématiquement été produite sous forme de corps d'inclusions (vérification en microscope photonique par contraste de phase) avec une absence d'activité dans la phase soluble. Face à ce problème j'ai envisagé de renaturer la protéine insoluble après une solubilisation à l'urée. J'ai tout d'abord essayé sans succès de replier la protéine après l'avoir fixée sur une résine d'affinité pour le tag his. J'ai finalement réussi à rétablir de l'activité par dilution (50 fois) brutale de la protéine solubilisée dans l'urée dans le milieu réactionnel de l'enzyme à température ambiante. Le rôle du substrat (polymère polyanionique) dans le repliement correct de la protéine peut s'expliquer par des interactions spécifiques et/ou aspécifiques. L'utilisation de matrice de refolding à 96 conditions n'a pas fourni de meilleur résultat. Il est bien évident que si l'activité ainsi générée est suffisante pour valider le clonage du gène et

sa séquence, elle est largement insuffisante pour exploiter l'enzyme surexprimé d'un point de vue industriel, même à l'échelle pilote.

6. Conclusions générales et perspectives

6.1. La λ -carraghénase, un outil en devenir

Les travaux exposés dans ce chapitre permettent d'aborder l'intimité de ce nouvel enzyme. Mais l'étude biochimique de la λ -carraghénase souffre de l'absence de méthode de dosage suffisamment sensible pour permettre le calcul des constantes cinétiques. En outre, le substrat naturel est lui-même ni suffisamment caractérisé et ni assez homogène pour assurer la reproductibilité et la fiabilité de tels calculs. L'idéal serait évidemment de disposer d'un lot de polymère homogène en taille et surtout en composition ce qui nécessite un travail de purification à partir de pool partiellement hydrolysés très astreignant car de rendement faible. On peut enfin se poser la question de la pertinence de ces constantes cinétiques obtenues sur un système en solution, alors que *in vivo* le substrat est très vraisemblablement présent sous forme solide car étant partie prenante de débris végétaux ou d'agglomérats de matière organique.

La caractérisation du comportement et des produits de l'enzyme sauvage sont l'objet des travaux de Marion Guibet actuellement en thèse au laboratoire. Ses résultats ont permis d'identifier par RMN la nature de trois des produits limites purifiés par chromatographie (SEC et HPLC) et obtenus sous l'action de l'enzyme sauvage pure. Il s'agit, pour les deux produits majoritaires, du néo- λ -carratétraose et néo- λ -carrahexaose. Ces résultats ont confirmé l'hydrolyse de la liaison $\beta(1\rightarrow4)$ et la reconnaissance par l'enzyme du motif λ canonique. Le DP2 étant le plus petit produit ceci montre que l'enzyme doit satisfaire au moins huit sous-sites de spécificité pour engendrer une fixation efficace du substrat ce qui aboutit à l'accumulation du DP3 et à l'absence du DP1. Ce constat nécessite cependant d'être nuancé. Il semble en effet que, à complétion, alors que l'enzyme n'a quasiment plus de substrat de $DP \geq 4$, le DP3 accumulé puisse devenir compétitif et être très lentement dégradé. Le troisième produit caractérisé, minoritaire, est un néo- λ -carratétraose sulfaté D2S,6S-G4S,2S-

D2S,6S-G2S. Ce produit présente un profil de sulfatation intermédiaire entre le λ - et le ν -carraghénane (respectivement -G2S-D2S,6S- et -G4S-D2S,6S-) et vient renforcer l'idée que la λ -carraghénase puisse reconnaître le ν -carraghénane. Ce sulfate supplémentaire, absent du λ -carraghénane, n'est pas nécessaire à la reconnaissance du substrat. Il crée cependant une densité de charge et un encombrement stérique très important. La tolérance de ce sulfate illustre donc le large spectre de reconnaissance du site actif (nombreux produits limites). Ce sulfate peut se trouver aux sous-sites +2 et/ou -3 dont la tolérance est peut être plus souple que les sous sites -1 et +1.

Alors que le mode d'action est de toute évidence de type « endo », le degré de processivité n'est pas défini et le mécanisme d'hydrolyse est encore à l'étude. Concernant la processivité, il faut tout de même souligner que la proportion de DP3 générée (attention estimation sur gel) semblerait suffisamment importante pour donner un rapport DP2/DP3 en faveur d'un comportement endo majoritairement aléatoire. En effet, intuitivement on comprend bien que à chaque attaque endo, l'enzyme a approximativement une chance sur deux de créer un fragment de DP impaire. Si l'enzyme est strictement processive le seul cas de production de DP3 correspondra à une extrémité de chaîne de DP impaire. Mais moins il est processif, autrement dit plus il décroche pour réattaquer en endo aléatoirement, plus la probabilité de générer des fragments de DP impaire augmente et plus au final la proportion de DP3 sera forte. Dans un cas de processivité nulle, elle sera maximale.

Ces études sur la biochimie de l'enzyme, la structure de ses produits et l'activité biologique de ces derniers, sont évidemment pénalisées par le demi succès de la surexpression hétérologue. Pour palier à ce problème il est envisagé d'améliorer expérimentalement le bornage des différents domaines de la protéine sauvage de façon à faciliter la surexpression du domaine catalytique en l'isolant du reste de la séquence. Pour ce faire, la protéine sauvage purifiée devrait être hydrolysée en conditions douces par différentes combinaisons de protéases. Les produits seront séparés par chromatographie, analysés par spectrométrie de masse (taille et composition), et soumis à un microséquençage N-terminal. Postulant que les charnières inter-domaines seront préférentiellement accessibles et donc sensibles aux protéases, cette méthode

permettra peut être de séparer les domaines, voire de déterminer lequel assure la catalyse.

En l'état, nous sommes donc encore loin de l'exploitation de cet enzyme à l'échelle pilote.

6.2. La λ -carraghénase, une nouvelle famille de carraghénases

La séquence de la λ -carraghénase a été confirmée par l'obtention d'activité recombinante. Son analyse montre qu'il s'agit d'une glycosidase distincte des κ - et ι -carraghénases déjà identifiées et très vraisemblablement orpheline au sein de cette classe d'enzyme (EC.3.1). Si nous ignorons le repliement du domaine catalytique de la λ -carraghénase, compte tenu des tests bioinformatiques réalisés, il est toutefois très probable qu'il soit différent de ceux de la κ - et ι -carraghénases. Ceci signifie donc que au cours de l'évolution un repliement supplémentaire a été sélectionné pour assurer la dégradation des carraghénanes. Que faut-il voir dans cette apparente correspondance entre la diversité des substrats et celle des repliements des enzymes les dégradant : une coïncidence dans l'enchaînement des découvertes, un hasard évolutif, et/ou l'effet de réelles contraintes biochimiques ?

Les différences de motif au sein des carraghénanes semblent d'un prime abord assez ténues à l'échelle du disaccharide. On pourrait ainsi penser qu'une protéine dégradant l'un pourrait « facilement » évoluer de façon à reconnaître l'autre. De mon point de vue, s'accommoder de la présence/absence de groupes sulfates chargés et volumineux par changement de mode de fixation (stacking hydrophobe/interactions électrostatiques) ainsi que par modifications stériques du site catalytique, n'impose pas a priori la sélection d'un repliement différent. Cependant, les positions et la quantité des substituants sulfates ou pont 3,6 anhydro, engendrent des différences très marquées dans la conformation et le potentiel d'interactions électrostatiques des résidus, lesquelles se répercutent sur la nature et la stabilité des structures secondaires, tertiaires et quaternaires des sucres (cf. Chapitre 1). Et il faut à mon sens envisager l'hypothèse d'éventuelles contraintes biochimiques sous cet angle. In vivo, de telles différences de structure quaternaire de ces polysaccharides justifieraient plus

solidement la sélection d'architectures protéiques (repliement et combinaison de domaines) différentes. Alors, évoquer l'impact de contraintes physico-chimiques de ces substrats sur l'évolution de la λ -carraghénase pourrait se résumer par l'hypothèse suivante : l'évolution d'une glycosidase GHx dégradant des fibres/gels de carraghénane en une protéine dégradant un carraghénane monocaténaire (ou inversement) est moins probable que l'acquisition de cette activité à partir d'une autre famille (GH ou pas) hydrolysant par exemple un polymère monocaténaire polyanionique. C'est une hypothèse délicate à défendre pour plusieurs raisons.

Ces considérations sous-entendent tout d'abord l'émergence de plusieurs activités carraghénases au sein d'une même espèce bactérienne. Or *Zobellia galactanivorans* et *Pseudoalteromonas carrageenovora* sont les seuls exemples à ma connaissance où plus d'une famille de carraghénases ont été identifiées. Elles contiennent chacune deux activités carraghénases (respectivement, κ - et ι -carraghénases, et κ - et λ -carraghénases). En outre, l'apparition de ces différentes fonctions a pu se faire dans des clades bactériens monophylétiques distincts, partant de patrimoines génétiques différents, et offrant en conséquences des solutions adaptatives différentes. La propagation de la fonction par transfert horizontal dans une niche écologique commune, peut alors expliquer la présence au sein d'une même souche bactérienne de plusieurs familles de carraghénases, sans que ne soit impliqué un handicap évolutif de telle famille pour telle activité (cf. Chapitre 3 § 2.4.). Le transfert horizontal est un phénomène de brassage génétique constituant un moyen d'évolution rapide dont l'importance est maintenant bien établie (Ochman et al., 2000 ; Boucher et al., 2003).

Des β -agarases ont été identifiées dans la famille GH16 puis GH50 et GH86, chacune pouvant être isolée dans des taxons bactériens indépendants, bien que plusieurs d'entre elles cohabitent chez *Microscilla* ou *Pseudoalteromonas atlantica*. Pour expliquer cela, soit on reprend le concept du transfert horizontal, soit on admet qu'une espèce peut faire émerger des fonctions similaires à partir de familles protéiques indépendantes malgré la préexistence dans son patrimoine génétique de fonctions proches ou identiques. Considérer cette dernière solution reviendrait à reconnaître un opportunisme génétique gouvernant l'émergence de nouvelles

fonctions, reposant largement sur le hasard et en décalage avec une vision trop dirigiste de l'évolution. Il est donc permis théoriquement d'envisager la sélection de plusieurs familles de carraghénases, dans une même espèce, au gré des conjonctures écologiques, sans accuser les spécificités biochimiques de tel ou tel carraghénane d'avoir rendu incompatible telle ou telle branche évolutive pour un substrat cousin.

Il apparaît enfin d'autant plus délicat d'évoquer des contraintes physico-chimiques liées aux substrats ayant conduit à cette ségrégation tant il est vrai que nous ignorons si, par exemple, aucune activité λ -carraghénase n'est portée par un membre transfuge encore inconnu des familles 16 ou 82 et réciproquement. D'un strict point de vue statistique, nous manquons d'un nombre suffisamment important de séquences de protéines partageant ces activités carraghénases (trois κ -, deux ι -, et une λ -carraghénases) pour confirmer ou infirmer ce qui n'est pour l'instant qu'une ségrégation anecdotique. Il est donc trop tôt pour chercher les raisons de l'exclusivité d'une famille pour un type de carraghénane puisque l'on ignore si elle est avérée. En tout état de cause, l'hypothèse de travail, postulant de l'impact du substrat dans la diversité des familles de carraghénases, me semble des plus fragiles. Et, dans l'immédiat, tout l'intérêt à disposer de cette nouvelle famille de carraghénase réside dans l'enrichissement qu'apportera l'étude d'un modèle enzyme/carraghénane original sur la connaissance des carraghénanes et de l'enzymologie des polymères en général.

ARTICLE 1

λ-CARRAGEENASE, A NEW GLYCOSIDASE DISTINCT FROM κ- AND ι-CARRAGEENASES

Sébastien Colin¹, Tristan Barbeyron¹, Jean-Claude Yvin², Bernard Kloareg¹, and Gurvan Michel¹.

¹UMR 1931 (CNRS and Laboratoires Goëmar), Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France ; ²Laboratoires GOËMAR, z.a.c. La Madeleine, avenue Général Patton, 35400 S^t Malo, France.

Correspondence to G. MICHEL, UMR 1931 (CNRS and Laboratoires Goëmar), Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France. Fax: +33 (0)2 98 29 23 24, Phone: +33 (0)2 98 29 23 30, E-mail: gurvan@sb-roscoff.fr

INTRODUCTION

Carrageenans are the main components of the cell walls of various marine red algae (Rhodophyta), where they are laid out as highly ordered molecules, both associated to the cellulose microfibrils or in the microfibril-less intercellular matrix (Gordon-Mills et al., 1977, 1978). This large family of hydrocolloids is made up of linear chains of galactopyranose residues in the D-configuration, linked by alternating $\alpha(1\rightarrow3)$ and $\beta(1\rightarrow4)$ linkages. This regular structure is modified by the occurrence of 3,6-anhydro bridges in the α -linked galactose residues and by substitution with ester sulfate, methyl or pyruvate groups (Kloareg and Quatrano, 1988). The main industrially used carrageenans are substituted with either one (κ -carrageenan), two (ι -carrageenan), or three (λ -carrageenan) sulfate ester groups per disaccharide repeating unit. κ -Carrageenan is conventionally described as the repetition of the disaccharidic motif 4-sulfate-*O*- β -D-galactopyranosyl-(1,4)-3,6-anhydro- α -D-galactose, also referred to as κ -neocarrabiose-sulfate. ι -Carrageenan differs from κ -carrageenan by the presence of one additional sulfate substituent per repeating disaccharide, at C2 on the α -linked D-galactose residues (Kloareg and Quatrano, 1988). Idealised λ -carrageenan is defined as 2-sulfate-*O*-D-galactopyranosyl-(1,4)-2,6-sulfate- α -D-galactose and then does not share the 3,6-anhydro-bridge.

In general, carrageenan serves as a gelling, stabilising and viscosity-building agent. In aqueous solutions, κ - and ι -carrageenans form thermoreversible gels (Rees, 1969) unlike λ -carrageenan which behave as a thickener. Gelling properties involves a superhelical-type mechanism of gelation, which consists of the formation of double-helices of carrageenan chains readily followed by the aggregation of double helices (Morris et al., 1980; Viebke et al., 1994). Gelation is promoted by the presence of such cations as potassium for κ -carrageenan and calcium for ι -carrageenan. Sodium has also a promoting effect but in a lesser extent. By binding to the helices, either electrostatically or through specific site binding, these cations considerably reduce the helix charge density and enhance their tendency to aggregate (Piculell et al., 1994). The different rheological behaviour of λ -carrageenan results from the fact that its $\beta(1\rightarrow4)$ linked galactose adopt a 4C_1 -conformation whereas in κ - and ι -carrageenans, the 3,6-anhydro ring induces the 1C_4 conformation in those residue which allows the helical secondary structure. Occurrence of disaccharide units without the 3,6-anhydro-ring and having 4C_1 -conformation such as η -, μ - and λ -carrageenan, causes sterical clashes in the regular chain and prevent the helical strands formation.

The recycling of the red algal biomass by marine bacteria involves endo-hydrolases specific of the internal β (1 \rightarrow 4) linkages of carrageenans and yielding products of the neocarrabiose series. As enzymes which naturally attack concentrated carrageenan gels and which recognize the pattern of sulfation on the digalactose repeating units (Johnston and McCandless, 1973; Barbeyron et al., 1994, 1998, 2000), carrageenases constitute an interesting model to study the solid-phase degradation of sulfated polysaccharides. To this end, we have cloned several marine bacterial κ - and ι -carrageenases and undertaken their structural resolution. On the basis of secondary structure analyses, κ -carrageenases were shown to belong to family GH-16 (Barbeyron et al., 1994, 1998). The κ -carrageenase from *Pseudoalteromonas carrageenovora*, which catalyses the endo-processive degradation of κ -carrageenans, folds into a curved β -sandwich, characteristic of this family of glycoside hydrolases (Michel et al., 2001A). ι -Carrageenases constitute the family GH-82, a family unrelated to that of κ -carrageenases (Barbeyron et al., 2000) but behave also as endo-processive enzyme (Michel et al., 2003). This finding was confirmed by the crystal structure of *Alteromonas fortis* ι -carrageenase which consists of a right-handed parallel β -helix (Michel et al., 2001B).

An extracellular enzymatic activity from *P. carrageenovora* degrading λ -carrageenan mixture was detected by Weigl and Yaphe (1966) and latter described by Johnston and McCandless (1973), followed by Greer (1984). The enzyme was then termed “ λ -carrageenase”. We here report the complete purification of this enzyme and the description of its biochemical behaviour. Its gene was also cloned, sequenced and overexpressed in *E. coli*.

MATERIALS AND METHODS

Bacterial Strains and Culture Conditions

The marine bacterium *Pseudoalteromonas carrageenovora* (ATCC 43555, previously known as *Alteromonas carrageenovora*) was obtained from the American Type Culture Collection. It routinely grew in Zobell Medium (Zobell, 1941), but to purify the λ -carrageenase, we used the Y-2 modified medium (Weigl and Yaphe, 1966) composed as following: solution A (NaCl 25 g/l, MgSO₄·7 H₂O 5 g/l, CaCl₂·2H₂O 0.1 g/l, KCl 1 g/l, NaNO₃ 2 g/l, casamino acids (Difco) 2.5 g/l, yeast extract (Difco) 1 g/l, purified λ -carrageenan 2 g/l, milliQ water), supplemented with 60 ml/l of solution B (Na₂HPO₄ 0.2M, pH 7.2), and 330 μ l/l of solution C (10 % w/v FeSO₄·7H₂O). Solution A,B and C were

autoclaved separately prior to mixing. Cultures were incubated in a G-25 incubator (New Brunswick Scientific) at 22 °C and 250 rpm for 24-48 h.

Escherichia coli strains DH5 α (*recA*, *endA1*, *gyrA96*, *thi1*, *hsdR17* [*rk-mk+*], *supE44*, *relA1*, *lacZ* Δ M15), was used to maintain plasmid vectors and grown on Luria Bertani medium (LB), eventually supplemented with bacto agar 15 g/l (Difco), at 37 °C routinely. BL21 (DE3) pLysS (F⁻, *ompT*, *hsdS_B* [*r_B⁻*, *m_B⁻*], *gal*, *dcm*, DE3, pLysS [Cm^R]), Rosetta (DE3) pLysS (F⁻, *ompT*, *hsdS_B* [*r_B⁻*, *m_B⁻*], *gal*, *dcm*, *lacY1*, DE3, pLysSRARE [Cm^R]), both from Novagen, was used to overexpressed recombinant protein and grown on LB media as DH5 α strain or M9 liquid media. When appropriate, ampicillin, and/or chloramphenicol, was added to solid or liquid media at respectively 100, and 34 μ g/ml.

Strains and clones are stored in their respective media, fresh and complemented with 20 % (v/v) glycerol at minus 80°C. Growth was monitored by measuring turbidity at 600nm.

Purification of *P. carrageenovora* λ -Carrageenase

All steps was performed at 4 °C. A 48 h-old *Pseudoalteromonas carrageenovora* cultures (5 l) was centrifuged (10000 g, 30 min) and the cell-free supernatant was concentrated by tangential ultrafiltration, using a 30 kDa cut-off polysulfone membrane, to a final volume of approximately 100 ml which was rinsed with buffer A (sodium phosphate 25 mM pH 7.2 containing NaCl 2 M) producing a final volume of 400 ml. This was achieved a pellicon system (Millipore). Fractional (NH₄)₂SO₄ precipitation was performed from this ultrafiltrate and the fraction obtained from 30 to 70 % was stirred 24 h. Centrifugation at 20000 g, 30 min yielded a precipitate which was suspended in 10 ml of buffer A. The suspension was dialyzed 12 h twice against 5 l of the same buffer. From these extract of 12 ml, to purify λ -carrageenase three chromatographies was carried out, using FPLC system (Pharmacia) to connect columns, and buffer A to equilibrate them and to elute samples. First, 10 ml was chromatographed on Sephacryl S-200 2.5/76 column, eluting at 1ml/min and harvesting 4 ml/tube. A pool of 26 ml from 7 fractions containing the peak of activity was then loaded on a Phenyl-Sepharose 6FF low sub column 2.4/14. The λ -carrageenase activity was eluted during washing step at 1 ml/min still harvesting 4 ml/tube. And with the aim to reduce the pool volume from 44 ml to 2 ml an ultrafiltration was done, taking care to preserve buffer properties and using a 10 kD cut-off YM membrane (Millipore). This was achieved with an Amicon cell 8050 (Millipore) under a nitrogen gas pressure of 1.5 bar. Finally 200 μ l of the (ultrafiltrate) retentat was chromatographed on a superdex 200HR 1.0/30, eluting at 0.5

ml/min and harvesting 0.5 ml/tube. Pure and active fractions are stored at 4 °C in buffer A several months without noticeable loss of activity.

Fractions were assayed for protein concentration (A_{280} or A_{220}) and λ -carrageenase activity. The purity of the enzyme was assessed by SDS-PAGE (Laemmli, 1973) with silver staining (kit PlusOne, Pharmacia Biotech). Each protein pool concentration from each main purification step was yielded by Bradford assay (Biorad; Bradford, 1976) compared to BSA (Bovine Serum Albumin, Sigma) range.

One N-terminal and two internal peptide sequences (respectively A, B and C, Fig. 3) were obtained by Edman degradation (service de Microséquence, Institut Pasteur, Paris) respectively from a PVDF membrane (Immobilon P-0.45 μ m, Millipore) where the enzyme was transferred, and the endolysine digest of acrylamide slices (stained with Amido Black 0.003 % w/v) purified by HPLC.

Assay for λ -Carrageenase Activity

λ -carrageenase activity was assayed by measurement of the initial rate of decrease of specific viscosity $(-d\eta/dt)_0$. A Brookfield LVDV-III+ cone/plate rheometer, monitored with the RHEOCALC 2.2. software and associated with a CP-52 cone spindle (1.2 cm diameter, and 3 ° angle), was used to measure viscosity at a speed of 100 rpm. The system was thermostated at 30 °C. Substrate was composed with 1 % (w/v) of λ -carrageenan (provided by CPKelco, GENU x7055, batch 98743 and extracted from *Gigartina skottsbergii*) diluted in buffer B (sodium phosphate 0.1 M pH 6.5 containing 0.1 M NaCl). It was prewarmed at 30 °C prior to use. Measures were made every 5 sec during 250 sec after the addition of 5 μ l of enzyme preparation to 495 μ l of substrate. In the determination of $(-d\eta/dt)_0$, the first 60-sec measurements were not considered as representative of the actual decrease of viscosity, to allow for stabilisation of the system. As an approximation, the rate of decrease of viscosity $(-d\eta/dt)_i$ was considered constant versus time between two successive measures and then equivalent to $(-\Delta\eta/\Delta t)_i$. Practically, at the beginning of the reaction, $(-\Delta\eta/\Delta t)_i$ was considered as the right initial rate of decrease $(-\Delta\eta/\Delta t)_0$, when $(-\Delta\eta/\Delta t)_i$ was constant and maximal during at least 30 sec after the stabilisation time. Thus, it was possible to plot $(-\Delta\eta/\Delta t)_0$ as a function of E, with E being the quantity of active enzymes, and $(-\Delta\eta/\Delta t)_0 \leq 0.25$ cP/sec. Then experiments (18 points) showed that the best description of relationship between those values are obtained with the parabolic regression curve:

$$(-\Delta\eta/\Delta t)_0 = K.E^{1.4379} \quad (1)$$

in which K is constant (correlation coefficient R=0.9910), instead of the linear regression curve $[(-\Delta\eta/\Delta t)_0 = K.E, R=0.9613]$. Since the number U of catalytic unity (UA) in a sample is directly proportional to E, equation (1) give:

$$U = K'.(-\Delta\eta/\Delta t)_0^{1/1.4379} \quad (2)$$

in wich K' is constant. In order to define K', one catalytic unity is arbitrary considered to be the quantity of enzyme which generate an initial rate of decrease of viscosity of 0.1 cP/sec. Finally, assuming a strict conservation of measures parameters between each samples, U of a sample was calculated from the empirical equation:

$$U = 4,96.(-\Delta\eta/\Delta t)_0^{0.695} \quad (3)$$

Analysis of λ -Carrageenase Products

The specificity of the enzyme was investigated on several λ -carrageenan mixtures extracted from different sources: *Gigartina pistillata* and *acicularis* (Sigma); *Gigartina skottsbergii* (CPKelco); *Gigartina acicularis* (Sanofi Bioindustries); *Chondrus crispus* (SKW biosystem); *Gigartina pistillata* (extracted according to Amimi et al., 2001). The κ -carrageenan (from *Euchema cottonnii*) and ι -carrageenan (from *Euchema spinosum*) both provided by CPKelco (respective batches GENU x6913 and GENU x6908), were also tested. Reactions were run in NaNO_3 0.1 M during 48 h with 3.8 UA/mg of polysaccharide. The enzyme came from a phenylsepharose 6FF step pool concentrated by ultrafiltration on a 10kDA cut-off membrane.

The kinetics of λ -carrageenan (from *G. skottsbergii*) hydrolysis was performed in buffer B using 1.4 UA/mg of polymer, adding 1 % (v/v) of enzyme from a sephacryl S200 step pool.

Whole hydrolysis were monitored by carbohydrate polyacrylamide gel electrophoresis (C-PAGE, Zablackis and Perez, 1990) using 6 % ((w/v)) polyacrylamide for the stacking gel, 27 % (w/v) polyacrylamide for the running gel (from a mix 36/4 of acrylamide/bisacrylamide). Gels were stained with 0.5 % (w/v) Alcyan Blue solution for 10 min and then, after washing, with silver nitrate. The main final products, the neo- λ -carratetraose (DP2) and neo- λ -carrhexaose (DP3), were HPLC purified and used as references (DP, expressed as the number of disaccharide repeating units).

Isolation and Analysis of λ -Carrageenase Clones

Total DNA from *P. carrageenovora* was prepared as previously described (Barbeyron et al ,1984), cut by the restriction endonuclease *Sau3AI*, and fractionated on a sucrose

gradient. DNA fragments of approximately 4-10 kb were cloned at the *Bam*HI site of plasmid vector pAT153 which were used to transform *E.coli* strain DH5 α . The genomic library contained approximately 6000 clones. Using the microsequences obtained from internal peptide C and N-terminal peptide A, degenerated DNA primers was designed, which allowed to get via polymerase chain reaction (PCR), a N-terminal DNA probes of 943 nt (Fig. 3). This probe was labelled radioactively (kit Megaprime DNA labelling systems, Amersham Lifescience), and used to screen the library according to Sambrook and Russell (2001). Clones of interest were shown consistent and independent by restriction mapping. The same labelled probe was used for a Southern Blot experiment (Sambrook and Russell, 2001) carried out on genomic *P. carrageenovora* DNA digested by several endonucleases combinations (*Sau*3AI, *Hin*DIII, *Hin*DIII/*Bam*HI, *Bam*HI, *Bam*HI/*Eco*RV, *Eco*RV, *Eco*RV/ *Hin*DIII) with the aim to determine the copy number of this gene.

Sequencing was carried out by gene walking using a 3100 Genetic Analyser (Applied Biosystem Instruments) with BigDye Terminator V3.0 chemistry (DNA sequencing kit, Applied Biosystem) and synthetic oligonucleotides as primers. Sequences were verified at least 5 times and most of them more than 10 times.

Sequence Analysis

Nucleotide sequence of the contig was searched for open reading frames which were translated according to the universal genetic code. They were investigated for putative Ribosomal Binding Sequences (RBS), putative promoters, and putative transcription terminators hairpins via *mfold* v.3.1 prediction server (Zuker et al., 1999). Signal peptides was predicted with PSORT algorithm (PSORT web server, Nakai et al, 1991) and SIGNALP v.1.1 and v.2.0 (SIGNALP web server, Nielsen et al, 1999). The BLASTp and PSI-BLAST (Altschul et al, 1997) searches were run on GENBANKnr and on PAT via the NCBI web server, but also on UNIPROT from the EBI server. Motif and Domain evidences were queried against the INTERPRO (via InterProSscan from the EBI web server. The InterPro Consortium, 2001; Zdobnov and Apweiler, 2001) and CDD (via RPS-BLAST, from NCBI web server) domain databases. Sequence was checked for repeats using the program RADAR (Heger and Holm, 2000) available from the EBI server.

To understand the domain architecture of the protein, further deep sequence analysis were carried out. Hydrophobic Cluster Analysis (HCA, Callebaut et al, 1997) were used to improve sequences comparisons and to decipher secondary structures elements. The Bioinfo.pl group's meta server was used, partially, to harvest also secondary structure

predictions from PSIPRED (McGuffin et al., 2000), SAM-t99-2D (Karplus et al., 1998) and PROFsec (improved version of PHDsec, Rost, 1996), but mainly to perform fold recognitions, since the meta predictor 3D-Jury system (Ginalski et al, 2003) was considered as the most specific tool during LiveBench 6 experiments (Rychlewski et al, 2003). The Structure Prediction Meta Server provides access to various fold recognition and local structure prediction methods. The Meta Server is also coupled to several consensus servers. They provide jury predictions based on the results collected from other services. Not all are used by the jury system. The 3D-Jury scoring was used with default settings. In such case, the significance cutoff of 50 provided a prediction accuracy of above 90% (Rychlewski et al, 2003). The 1.65 release from SCOP was used as fold libraries reference (Murzin et al., 1995).

Production of Recombinant λ -Carrageenase

The construction of the overexpression vector was based on GATEWAY Cloning Technology (Invitrogen) according to supplier instructions. The coding region of *P. carrageenovora* λ -carrageenase was amplified by PCR with high fidelity PLATINUM *Pfx* DNA polymerase (Invitrogen). PCR products were inserted into pDONR201 vector. The sequence was verified, and right clones were used to insert in-phase the coding region into the pDEST17 vector, which allows overexpression of N-terminal 6xHistidine tag protein fusion in *E. coli* from a T7 promoter. Such plasmid vectors were transformed in the DE3 *E. coli* strains BL21 and Rosetta (Novagen), with or without pLysS. Several conditions of cultures were then tested for intracellular overexpression in presence of various concentration of isopropyl-1-thio- β -D-galactopyranoside (IPTG) ranged from 0.1 mM to 1 mM. LB and M9 medium were tried at 15, 20, 25, 30 and 37 °C. The expression of the fusion protein was checked by SDS-PAGE analysis from induced culture cellular lysat supernatant (centrifuged 15000 g, 30 min) or full extract, compared to non induced culture. Inclusion bodies (IB) were purified using BugBuster Protein Extraction Reagent (Novagen) according to supplier protocol. IB were then made soluble in Tris buffer 50mM pH8 with 8M urea (Buffer C) and assayed for activity by C-PAGE.

RESULTS

Purification of *P. carrageenovora* λ -Carrageenase

P. carrageenovora was grown in presence of λ -carrageenan with the aim to induce the enzyme production according to Jonhston and McCandless results (1973). It was noticed that

the culture growth was faster and longer in presence of this polymer than without, indicating that the strains metabolised it (data not shown) as a carbon source. Moreover, Greer (1984) reported that in such culture conditions, the enzyme production in the cell free medium parallels the cellular growth with a maximum occurring at the early stationary phase. The production culture was then inoculated with exponential phase culture grew on the same media at 0.005 OD₆₀₀ and stopped after 40h (at 0.225 OD₆₀₀).

Besides, the reducing end of hydrolysed λ -carrageenan generated by this enzyme display a strong lack of reactivity compared to some other galactan as agarose, κ - or ι -carrageenan. The enzyme activity was thus measured by viscosimetry which appears much more sensitive. Nevertheless the decrease of viscosity is not directly reliable to a number of catalytic events. And in this case, these method was strictly considered as an useful tool to monitor the enzyme activity during the purification and to compare samples activity to each others.

The purification protocol exposed here (inspired from Greer, 1984) leads to obtain a pure extracellular protein which digests purified λ -carrageenan (Fig. 1). Based on SDS-PAGE analysis the size of this protein is estimated to 100 kDa (Fig. 1). This size was also investigated by gel permeation on superdex 200HR chromatography which indicated a much lower size of 78 kDa. The SDS-PAGE 100 kDa protein band was excised, and two internal and one N-terminal peptide sequences was determined by Edman degradation (underlined in Fig. 3).

The isolation process yielded a 17-fold purification with a final recovery of 3% by extrapolating the last step result achieved with only 10% to the previous pool (Table II). This low rate is mainly caused by a strong loose of activity during the 70% (NH₄)₂SO₄ precipitation. A measure of the 70% (NH₄)₂SO₄ supernatant activity didn't explain these fall which is then attributed to an enzyme denaturation. This observation illustrates the high sensibility of this protein which was also reported by Weigl and Yaphe (1966), Jonhston and McCandless (1973), and Greer (1984) specifically during desalting, freezing, freeze-drying, and even low heating (35°C during 30 min) which are described as abolishing whole activity. In presence of λ -carrageenan, the enzyme seems nonetheless to be slightly protected from freezing denaturation. In the purification buffer which has a high ionic strength (NaCl 2 M), this enzyme is stable several months at 4°C.

Characterization of *P. carrageenovora* λ -Carrageenase

The optimum pH and temperature were previously described as being respectively 6.5 and 30°C by Jonhston and McCandless (1973), and Greer (1984). The lack of reactivity of the reducing ends produced by the enzyme didn't allow a proper kinetic parameters studies. The biochemical work exposed latter was made with partially purified enzyme. But both carrageenan hydrolysis patterns (monitored by C-PAGE, Fig. 4B) from the partially purified and recombinant enzyme, are similar. It was therefore admitted the absence of “parasitic” activity about λ -carrageenan in the purification step pools used.

With the goal to observe the kinetics of enzyme products appearance, a time course of the degradation of λ -carrageenan from *Gigartina skottsbergii* was analysed by C-PAGE (Fig. 2A). Only the small DP (up to DP15 approximately) are well separated on the gel, while the longer products look like smeary. There are two major end products, which are neo- λ -carrhexaose and neo- λ -carratetraose as confirmed by ^1H NMR spectroscopy (data not yet published). Both appear very early during this kinetic and accumulate while longer product from the neo- λ -carrabiose serie appear latter in a transient way. Moreover, none disaccharides were observed while almost all the substrate is digest.

Some completion hydrolysis by this λ -carrageenase of λ -carrageenan mixture from different sources were then analysed by C-PAGE (Fig. 2B). They display complex patterns illustrating the high diversity of ends products yielded by this enzyme. The most homogenous source seems to be the λ -carrageenan extracted from *Gigartina skottsbergii* (lane 2) which yield two main bands corresponding to the idealized neo- λ -carrhexaose (DP3) and neo- λ -carratetraose (DP2). The hydrolysis of λ -carrageenan from *Gigartina acicularis* and *G. pistillata* (lane 3 and 5) display almost the same pattern which is quite different from those obtained from Sigma source (lane 1) provided as being an extract from a mix of these same seaweeds. The λ -carrageenan from *G. pistillata* has been previously described as very heterogeneous polysaccharide (Amimi et al., 2001) with four 3-linked (G) and four 4-linked galactose (D) derivatives identified, only 20 % of typical λ -carrageenan, and especially 50% of G residues 2- and 6- sulfated. This complexity is well confirmed by the digest pattern presenting much more bands than “ideal” end products. The same analysis is allowed for the completion hydrolysis pattern from *Chondrus crispus* which appears as the more complex of them. In the same conditions κ -carrageenan and ι -carrageenan were not attacked (data not shown).

The λ -carrageenase gene from *P. carrageenovora*

From the two more distant peptidic microsequences obtained (see above and underlined Fig. 3), a N-terminal probe of 906 nt was available after PCR from degenerated primers (underlined Fig. 3). Besides, a genomic library of total DNA from *P. carrageenovora* was made. On the 6000 screened clones with the probe, 14 were positives, ranging insert size from 20 kb to 4 kb. Three plasmids of them, named pAT153la38, pAT153la45 and pAT153la53 (with respective insert size of 3.8, 6.1 and 5 kb) was subjected to a physical mapping. And the three maps were consistent, sharing a common 3.7 kb region. Using the same probe, Southern Blot experiments were carried out on the pAT153la45 insert and on the whole genomic DNA from *P. carrageenovora*, digested by several endonucleases combinations. It was then proved that this gene has a single copy in this genome. Moreover this N-terminal probe was localized on the pAT153la45 insert, proving that this one contained the whole targeted gene.

The insert of plasmid pAT153la45 was then completely sequenced which gave a sequence 5665 nt in length. The C-terminal end of one partial open reading frame (ORF) was detected in 5' of this insert (*cgx1*). It was followed 602 nt downstream on the same strand by the targeted ORF referred to as *cglA* (2826 nt in length, Fig. 3). And a last complete ORF (*cgx2*, 891 nt in length) was found downstream, on the reverse strand, in 3' from *cglA*. In the 5' non coding region of *cglA*, two hexamers, TTGACg and, TAaAcT, separated by 17 nucleotides (in bold Fig. 3), are well consistent with *E. coli* promoters consensus -35 and -10 (Von heijne, 1987; Busby and Ebright, 1994). They are localized 68 nt upstream the start codon (ATG²¹⁷). The hexamer AGGAat located 5 nt upstream this start codon (in bold Fig. 3), constitutes a likely Ribosome-Binding site candidat, type Shine-Dalgarno (Kosak, 1999). In the 3' non coding region two inverted-repeats are found (arrows Fig. 3) with respectively a *mfold* predicted hairpin free energy of -16.0 and -16.9 kcal/mol. The first one starting 19 nt downstream of the stop codon, is followed by three thymidine and was maintained while various sequence length were submitted to prediction. The stem region show a high GC content. This stem-loop shares then the main characteristics of a *E. coli* rho-independent terminators (Washio et al., 1998).

The predicted product of λ -carrageenase gene from *P. carrageenovora* (*cglA*) is a preprotein of 942 amino acids giving a theoretical molecular mass of 105.6 kDa. It includes the three peptides determined by microsequencing (underlined Fig. 3). The peptide signal predictions converge on two contradictory results. First, PSORT and SIGNALP v.1.1 predicted a signal type prolipoprotein with a cut between G17 and C18, which is allowed by

the presence of the consensus of lipoprotein signal recognized by signal peptidase II (defined in PROSITE PS00013). Second, SIGNALP v.2.0 predicted a cut between A25 and S26, according to the microsequencing of the N-terminal peptide. Adopting the last prediction confirmed by experiments, the computed mass of the resulting mature protein is 103.0 kDa, similar to the SDS-PAGE estimation. This putative mature sequence of 917 residues was further used for sequence analysis.

The BLASTp and PSI-BLAST analysis of this sequence against GENBANKnr and SWISSPROT told indeed poor informations. Actually nearby 75% of the sequence has no significant similarities with any known protein sequence, and even the last 25% display quite low score. The best match (25% of sequence pairwise identities) was obtained about the first 200 residues. This part hits several sequences sharing a β -propeller structure. Majority of them are eukaryotic sequences which corresponds to the wider β -propeller occurrence in higher organisms. Nevertheless the λ -carrageenase don't share any of the repeat which are the main sequence characteristic of those structural domains (WD, FG-GAP, BNR and PQQ repeats). More sensible searches allows to extend β -propeller hits up to the first 400 residues. The next and last interesting match occurs on the 80 C-terminal residues which reveals consistent similarities (38% of sequence pairwise identities) with the C-terminal end of a cyclomaltodextrinase from *Thermotoga maritima* (AAD36898) belonging to family 13 of glycoside hydrolases (Henrissat and Bairoch, 1996). The match concerns the α -amylases C-terminal β -sheet (SCOP b.71.1; Interpro IPR006048) which follows the catalytic domain type $(\alpha/\beta)_8$ barrel of the GH 13 family. The hit was confirmed by PSI-BLAST analysis and by plot HCA comparison.

Crossed validated secondary structure predictions described two mainly beta areas in N-terminal (370 first residues) and C-terminal (100 last residues) corresponding both to the regions described above. Between them, a central segment of 470 amino acids seems to be composed of mixed structures alpha/beta. Concerning the N-terminal domain, the fold recognition methods, via the metaserver 3D-Jury, agree the β -propeller hypothesis, which are also mainly beta. In C-terminal, this fold recognition metaserver converge to some beta sandwich fold domain which follow the catalytic domain of several glycoside hydrolases (family GH 29 and 13).

The overexpression of λ -carrageenase gene from *P. carrageenovora*

Despite many tested conditions for expression in *E. coli* of the whole λ -carrageenase coding region, the protein was found systematically as insoluble inclusion bodies (Fig. 4A). Activity assays were then carried out from purified inclusion bodies, made soluble in urea buffer. It was then found that such treated proteins were able to fold properly in presence of its substrate by simple dilution of the urea in the reaction buffer. A low activity was thus restored.

Compared to the wild enzyme, the hydrolysis pattern of *G. skottsbergii* λ -carrageenan looks perfectly identical when monitored by C-PAGE (Fig. 4B).

DISCUSSION

A New Enzymatic Function Available

An extracellular enzymatic activity from *P. carrageenovora* degrading λ -carrageenan mixture was previously detected by Weigl and Yaphe (1966) and later described by Johnston and McCandless (1973), followed by Greer (1984). The enzyme was then termed “ λ -carrageenase”. The work exposed here clearly demonstrated that the described enzyme recognizes and hydrolyses canonical λ -carrageenan (Fig. 2A and B) yielding as end products neo- λ -carrhexaose (DP3) and neo- λ -carrtetraose (DP2). That firstly means that this enzyme behaves as a glycoside hydrolase splitting the $\beta(1 \rightarrow 4)$ glycosidic bond. Secondly, two disaccharide units are requested from both sides of the catalytic site to allow efficient catalysis. In other words, at least 8 recognition sub-sites must be filled in (with a sub-site corresponding to one galactose unit of the polysaccharide) to potentiate the catalysis. Moreover, the fact that this enzyme induces a fast decrease of the viscosity of the λ -carrageenan, points out an endo-attack of this polysaccharide, which is confirmed by the release during the kinetic of whole intermediates from the neo- λ -carrabiose series between initial substrate population and final products (Fig. 2A).

The early part of this kinetic shows the final products as the first oligosaccharides detected with endo behaviour on the high DP (>15) polysaccharide (lowering of the smear), which would suggest an endo-processive behaviour of the enzyme. But, because intermediary products are also substrate, their concentration depends on a production and a consumption velocity. Resulting from this equilibrium, the fact that low DP products were not accumulated at the beginning, could be interpreted differentially. On one hand, they are not produced, on the

other they are overconsumed. Therefore, respectively, either the enzyme yield final products without short intermediaries, which means an endo-processive behaviour, or it adopts an endo-random behaviour with a higher reactivity of the shorter intermediaries (DP<15, because for example a higher diffusion). Moreover, even if C-PAGE staining don't allow fine quantitative analysis, both end product bands show similar intensities. And, a low ratio DP2/DP3 (≈ 1) therefore disagrees a processive behavior because on one hand the DP3 product origin have to be likely associated with chain extremities by-products, and on the other, it would be very surprising that a processive "depolymerase" yield identically two different product sizes. Whatever it may be, the elucidation of DP3 product origin would be helpful to discriminate definitely both hypothesis.

The λ -carrageenan was originally defined by Smith et al. (1954) as KCl soluble carrageenan predominantly obtained from sporophytic red algae (Craigie, 1990). But it's well established now that such polysaccharide extracts cover a wide variety of carrageenan diads deviating from the idealized " λ " structure (Table I). This heterogeneity rest mainly on the sulfation pattern but also on 3-6-anhydrogalactose content, and pyruvate groups presence, which depend upon the sample history, e.g., the seaweeds species, the age of harvested tissues, the seasons of harvesting and even extraction process (Craigie, 1990; Falshaw and Furneaux, 1994, 1995, 1998; Amimi et al., 2001; Van de Velde et al., 2002). If some variant diads may present higher charge density than classical λ -carrageenan (especially the potential G2S6S-D2S6S), the others are equally (π -carrageenan) or lower charged (α -, δ -, ξ -, and θ -carrageenan). And actually, besides the two "ideal" end products, the λ -carrageenase generates from complex λ -carrageenan mixtures, several neighbouring bands of low DP not yet identified (Fig. 2B). This end-products diversity light on the presence of non-canonical diads in such oligosaccharides. Their different charge density, and may be also their position in the oligosaccharide, affect their electrophoretic properties. In addition, Greer (1984) separated by chromatography two peaks attributed to both smallest λ -carrageenase digested products from *C. crispus* λ -carrageenan. And the reported ^{13}C -NMR spectroscopy analysis of both pools (known now as being neo-tetracarrabiose and neo-hexacarrabiose), without providing clear structural features, shown that they contented a mixture of different disaccharide units. It arises from those considerations that some of the smaller "atypical" end-products might be irregular DP2 and DP3. Assuming this solid hypothesis, this λ -carrageenase should present a relative flexibility of recognition. And the idea that evolution selected a wide spectra enzyme instead of an arsenal of highly specific enzymes, seems coherent to degrade a cell wall polysaccharide containing many slightly different pattern.

Nevertheless this galactanase wasn't be able to attack almost pure κ -carrageenan and ι -carrageenan which belong to another carrageenan family (Table I).

To conclude, the λ -carrageenase studied here seems to be a new and suitable tool to investigate λ -carrageenan complexity and to obtain λ -carrageenan signature.

An Original Glycoside Hydrolase

The λ -carrageenase from *P. carrageenovora* is purified from the cell-free culture supernatant, and is inducible with its own substrate, as it was observed with κ -carrageenases from *P. carrageenovora* and *Zobelia galactanovorans* (Mclean and Williamson, 1979; Potin et al., 1991), and ι -carrageenases from *Alteromonas fortis* and *Zobelia galactanovorans* (Greer and Yaphe, 1984; Barbeyron et al., 2000). Nevertheless, it was suggested that, first ι -carrageenase from *Alteromonas fortis* was associated with membranes, and second the one from *Zobelia galactanovorans* harboured a lipoprotein signal-peptide which anchors the enzyme in membrane also. The N-terminal end sequence of the purified extracellular λ -carrageenase confirms a signal peptide prediction. But the presence of lipoprotein signal peptide consensus should be also considered, referring to both ι -carrageenases ambivalence.

The fact that substrate is metabolised by *P. carrageenovora*, suggests the implication of others enzymes (e.g. glycosulfatase). But the *cglA* gene is widely physically and functionally separated from neighbouring ORFs, and very likely transcription initiators and terminators were found. This gene shouldn't so participate to any operon.

The sequence analysis announced a modular protein, but none domain were clearly definable. Nevertheless, in spite of the lack of identified repeats, the N-terminal β -propeller hypothesis have to be strongly considered. Actually, β -propeller are circular arrangement of several 4-stranded repeated β -sheets (4 to 8), or blades, associated with a high functional diversity (Fülöp and Jones, 1999). Some insertions may occur between blades and even strands (often α helix, and more rarely separated domains) which introduce breaks in the sequence continuity of the β strands. In addition, these domains generally share no significant sequence similarity and frequently show no easily detectable sequence repeats. But their three-dimensional structures and their individual blades are very closely superimposable (Fülöp and Jones, 1999) which contribute to the suitability of this fold for reliable prediction. Moreover, amongst mainly- β folds with a suitable size (>250 residues) catalogued in the current release of CATH database, β -propellers are majority. Those considerations strengthen the fold recognition methods positive results whilst BLAST reply was ambiguous. But the size of predicted β -propellers may vary from one prediction to one other because threading

templates may contain various number of similar structural repeats, which are themselves composed by similar β -strands. In such conditions, the anchoring of request sequence on the template shouldn't be interpreted as certitude, since the risks of sliding is high without defined repeats. Catalytic β -propellers was already found in some Glycoside hydrolases families (GH-x): a six blades propeller in clan E (GH-33, 34 and 83; Crennell et al., 1996); a five blades propeller in clan F (GH-43 and 62; Nurizzo et al., 2002). A last one was predicted in family GH-32 (Pons et al., 1998) from clan J (GH-32 and GH-68) without also any detected repeats. None sequence homologies were found with them and this hypothetical propeller may be not carry catalytic function since this fold is often dedicated to interactions protein-protein, specially in membrane-associated proteins. Finally, the ι -carrageenases have also a wide N-terminal all- β domain arranged as a right-handed parallel β -helix, but again, none homologies were found with the λ -carrageenase.

To conclude, the sequence from λ -carrageenase is markedly different from whole known Glycoside Hydrolases and specially carrageenases all grouped in family 16 (κ -carrageenases) and 82 (ι -carrageenases). Therefore this enzyme should be considered as a new unclassified glycoside hydrolase with presumably a new carrageenase fold (Michel et al., 2001 A and B). This result substantiates the working hypothesis that, in the evolution of galactanases different structural features have been selected to accommodate the linear density and position of sulfate groups on galactans. At this stage, however, these results focus on the possibility that, beyond their chemical structure, the three-dimensional structures of the sulphated galactans in the solid state also are critical in interactions with the enzymes catalysing their hydrolysis, especially since the λ -carrageenan do not gelify as κ -carrageenan and ι -carrageenan.

REFERENCES

- Altschul, S.F., Madden, T.L., Schäffer, 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. *Nuc. Ac. Res.*, **25**, 3389-3402.
- Amimi, A., Mouradi, A., Givernaud, T., Chiadmi, N., and Lahaye, M. (2001) Structural analysis of *Gigartina pistillata* carrageenans (Gigartinaceae, Rhodophyta). *Carbohydr. Res.*, **333**, 271-279.
- Barbeyron, T., Gerard, A., Potin, P., Henrissat, B., and Kloareg B. (1998) The kappa-carrageenase of the marine bacterium *Cytophaga drobachiensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.*, **15**, 528-537.

- Barbeyron, T., Henrissat, B., and Kloareg, B. (1994) The gene encoding the kappa-carrageenase of *Alteromonas carrageenovora* is related to beta-1,3-1,4-glucanases. *Gene*, **139**, 105–109.
- Barbeyron, T., Kean, K., and Forterre, P. (1984) DNA adenine methylation of GATC sequences appeared recently in the *Escherichia coli* lineage. *J. Bacteriol.*, **160**, 586-590.
- 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.
- Busby, S., and Ebright, R.H. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell*, **79**, 743-746.
- Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J., Henrissat, B., and Moron, J.P. (1997) Deciphering protein sequence information through hydrophobic cluster analysis. Current status and perspectives. *Cell. Mol. Life Sci.*, **53**, 621-645.
- Craigie, J.S. (1990) Cell Walls. In : (K. M. Cole and R. G. Sheath, eds) *Biology of the Red Algae*. Cambridge University Press, Cambridge. pp. 221-257.
- Crennell, S.J., Garman, E.F., Philippon, C., Vasella, A., Laver, W.G., Vimr, E.R., and Taylor, G.L. (1996) The structures of *Salmonella typhimurium* LT2 neuraminidase and its complexes with three inhibitors at high resolution. *J. Mol. Biol.*, **259**, 264-280.
- Cuff, J.A. and Barton, G.J. (1999) Evaluation and Improvement of Multiple Sequence Methods for Protein Secondary Structure Prediction. *Proteins*, **34**, 508-519.
- Douguet, D. and Labesse, G. (2001) Easier threading through web-based comparisons and cross-validations. *Bioinformatics*, **17**, 752-753
- Falshaw, R., and Furneaux, R.H. (1994) Carrageenan from the tetrasporic stage of *Gigartina decipiens* (Gigartinales, Rhodophyta). *Carbohydr. Res.*, **252**, 171-182.
- Falshaw, R., and Furneaux, R.H. (1995) Carrageenans from the tetrasporic stages of *Gigartina clavifera* and *Gigartina alveata* (Gigartinales, Rhodophyta). *Carbohydr. Res.*, **276**, 155-165.
- Falshaw, R., and Furneaux, R.H. (1998) Structural analysis of carrageenans from the tetrasporic stages of the red algae, *Gigartina lanceata* and *Gigartina chapmanii* (Gigartinales, Rhodophyta). *Carbohydr. Res.*, **307**, 325-331.
- Fülöp, V., and Jones, D.T. (1999). β -propellers: structural rigidity and functional diversity. *Curr. Opin. Struct. Biol.*, **9**, 715-721.
- Gordon-Mills, E.M., and McCandless, E.L. (1977) Carrageenans in the cell walls of *Chondrus crispus* Stack. (Rhodophyceae, Gigartinales). I Birefringence. *Phycologia*, **16**, 169–176.
- Gordon-Mills, E.M., and McCandless, E.L. (1978) Carrageenans in the cell walls of *Chondrus crispus* Stack. (Rhodophyceae, Gigartinales). II Metachromasy and the optical rotation. *Phycologia*, **17**, 95–104.
- Greer, C.W. 1984. A study of carrageenases from marine bacteria. PhD thesis, McGill University. Montreal. 208 pp.
- Greer, C.W., and Yaphee, W.(1984) Purification and properties of ι -carrageenase from a marine bacterium. *Can. J. Microbiol.*, **30**, 1500-1506.
- Heger, A., Holm, L. (2000) Rapid automatic detection and alignment of repeats in protein sequences. *Proteins*, **41**, 224-237.

- Nielsen, H., Brunak, S., and von Heijne, G. (1999) Machine learning approaches to the prediction of signal peptides and other protein sorting signals. *Prot. Engineer.*, **12**, 3-9.
- Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.*, **316**, 695-696.
- Johnston, K.H., and McCandless, E.L. (1973) Enzymic hydrolysis of the potassium chloride soluble fraction of carrageenan: properties of “λ-carrageenases” from *Pseudomonas carrageenovora*. *Can. J. Microbiol.*, **19**, 779-788.
- Kelley, L.A., MacCallum, R.M., and Sternberg, M.J.E. (2000) Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM. *J. Mol. Biol.*, **299**, 499-520.
- Kloareg, B., and Quatrano, R.S. (1988) Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanogr. Mar. Biol. Annu. Rev.* **26**, 259–315.
- Kosak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187-208.
- McGuffin, L.J., Bryson, K., and Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404-405.
- Mclean, M.W., and Williamson, F.B. (1979) κ-carrageenase from *Pseudomonas carrageenovora*. *Eur. J. Biochem.*, **93**, 553-558.
- Michel, G., Helbert, W., Kahn, R., Dideberg, O., and Kloareg, B. (2003) The structural bases of the processive degradation of iota-carrageenan, a main cell wall polysaccharide of red algae. *J. Mol. Biol.*, **334**, 421-433.
- Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O., (2001A) The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure (Camb)*, **9**, 513-525.
- Michel, G., Chantalat, L., Fanchon, E., Henrissat, B., Kloareg, B., and Dideberg, O., (2001B). The iota-carrageenase of *Alteromonas fortis*. A beta-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J. Biol. Chem.*, **276**, 40202-40209.
- Morris, E.R., Rees, D.A., and Robinson, G. (1980) Cation-specific aggregation of carrageenan helices: domain model of polymer gel structure. *J. Mol. Biol.*, **138**, 349-362.
- Nakai, K., and Kanehisa, M. (1991) Expert system for predicting protein localization sites in Gram-negative bacteria. *Proteins*, **11**, 95-110.
- Nosedá, M.D., and Cerezo, A.S. (1993) Room temperature, low-field ¹³C-n.m.r. spectra of degraded carrageenans: Part III. Autohydrolysis of a lambda carrageenan and of its alkali-treated derivative. *Int. J. Biol. Macromol.*, **15**, 177-181.
- Nurizzo, D., Turkenburg, J.P., Charnock, S.J., Roberts, S.M., Dodson, E.J., McKie, V.A., Taylor, E.J., Gilbert, H.J., and Davies, G.J. (2002) *Cellvibrio japonicus* alpha-L-arabinanase 43A has a novel five-blade beta-propeller fold. *Nat. Struct. Biol.*, **9**, 665-668.
- Piculell, L., Svante, N., Viebke, C., and Zhang W. (1994) Gelation of (some) seaweed polysaccharides. In: *Food Hydrocolloids: Structures, Properties and Functions*, K. Nishinari and E. Doi Eds, Plenum Press, New York, pp. 35–44.
- Pons, T., Olmea, O., China, G., Beldarrain, A., Marquez, G., Acosta, N., Rodriguez, L., and Valencia, A. (1998) Structural model for family 32 of glycoside-hydrolase enzymes. *Proteins*, **33**, 383-395.

- Potin, P., Sanseau, A., Legall, Y., Rochas, C., and Kloareg, B. (1991) Purification and characterization of a new κ -carrageenases from a marine Cytophaga-like bacterium. *Eur. J. Biochem.*, **201**, 241-247.
- Rees, D.A. (1969) Structure, conformation, and mechanism in the formation of polysaccharide gels and networks. *Advan. Carbohydr. Chem. Biochem.*, **24**, 267–332.
- Sambrook, J., and Russel, D.W. (2001). Molecular cloning: a laboratory manual, third edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1, 1.126-1.128, 1.135-1.142
- Shi, J., Blundell, T.L., and Mizuguchi, K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure- dependent gap penalties. *J. Mol. Biol.*, **310**, 243-257.
- Smith, D.B., Cook, N.H., and Neal, J.L. (1954) Physical studies on carrageenin and carrageenins fractions. *Arch. Biochem. Biophys.*, **53**, 192-204.
- Van de Velde, F., Knutsen, S.H., Usov, A.I., Rollema, H.S., and Cerezo A.S. (2002). ^1H and ^{13}C high resolution NMR spectroscopy of carrageenans: application in research and industry. *Trends in food sc. and technol.*, **13**, 73-92.
- Viebke, C., Piculell, L., and Nilsson, S. (1994) On the mechanism of gelation of helix-forming biopolymers. *Macromolecules*, **27**, 4160–4166.
- Von Heijne, G. (1987). Sequence analysis in molecular biology, treasure trove or trivial pursuit. Academic Press Inc., London, UK, 23-29.
- Washio, T., Sasayama, J., Tomita, M. (1998). Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination. *Nuc. Ac. Res.*, **26**, 5456-5463.
- Weigl, J. and Yaphe, W. (1966) The enzymic hydrolysis of carrageenan by *Pseudomonas carrageenovora*: purification of a κ -carrageenase. *Can. J. Microbiol.*, **12**, 939-947.
- Zablackis, E., and Perez, J. (1990) A partially pyruvated carrageenan from Hawaiian Grateloupia filicina (Cryptonemiales, Rhodophyta). *Bot. Mar.*, **33**, 273- 276.
- Zobell, C.E. (1941) Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *J. Mar. Res.*, **4**, 41-75.
- Zuker, M., Mathews, D.H., and Turner D.H. (1999) Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide. In *RNA Biochemistry and Biotechnology*, 11-43, J. Barciszewski & B.F.C. Clark, eds., NATO ASI Series, Kluwer Academic Publishers.

TABLE AND FIGURES

Table I: Disaccharides repeating structures of some carrageenans ^a

(1→3)-linked ^b	(1→4)-linked ^b	Greek symbol	
G4S	DA	canonical κ	Kappa family
G4S	DA2S	canonical ι	
G	D2S,6S	δ	Lambda family
G	DA2S	α	
G2S	D2S,6S	canonical λ	
G2S	DA2S	θ	
G2S	D2S	ξ	
GP2S	D2S	π	
G2S,6S	D2S,6S (probable)	not attributed	
G2S,6S	D2S (probable)	not attributed	

^a Adapted from Amimi et al. (2001), Knutsen et al. (1994) and Craigie (1990).

^b Refer to (1→3)- and (1→4)-linked galactose residues, with **G**, β-D-galactopyranosyl; **D**, α-D-galactopyranosyl; **DA**, 3,6-anhydro-α-D-galactopyranosyl; **S**, ester sulfate; **P**, pyruvate acetal; and Arabic numerals corresponding to the carbon atom number carrying substitution.

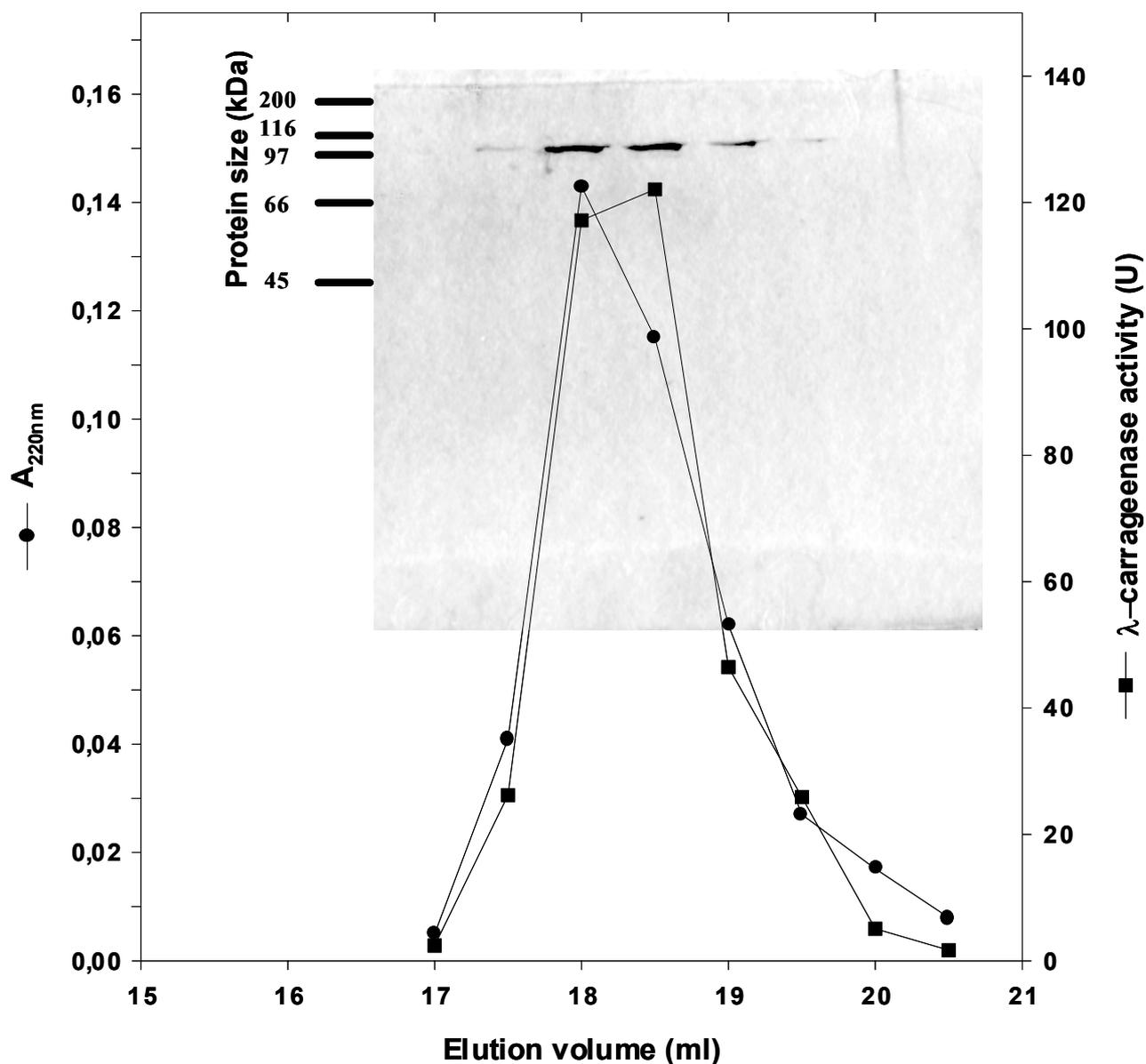


FIG. 1. Size exclusion chromatography of *P. carrageenovora* λ -carrageenase using a Superdex™ column 200HR 10/30. The λ -carrageenase-containing fractions pooled from affinity chromatography on Phenyl Sepharose 6FF low sub was reduced from 44 ml to 2 ml by ultrafiltrating on 10kD pore-size membrane using an Amicon cell with a 1 bar nitrogen pressure. And 200 μ l of these 2 ml was applied to the column and eluted with a 25mM phosphate buffer pH 7.5 containing 2M NaCl, using a 0.5ml.min⁻¹ flow rate. Fractions (0.5 ml) were assayed for protein at A₂₂₀ (plain circles) and λ -carrageenase activity (plain squares) by measuring the initial rate of decrease of the λ -carrageenane specific viscosity. Finally, homogeneity of the peak (the elution volume between 17 and 20.5 ml), was investigated by SDS-PAGE analysis on a 10% gel, with silver staining, each lane corresponding to one of the eluted fractions.

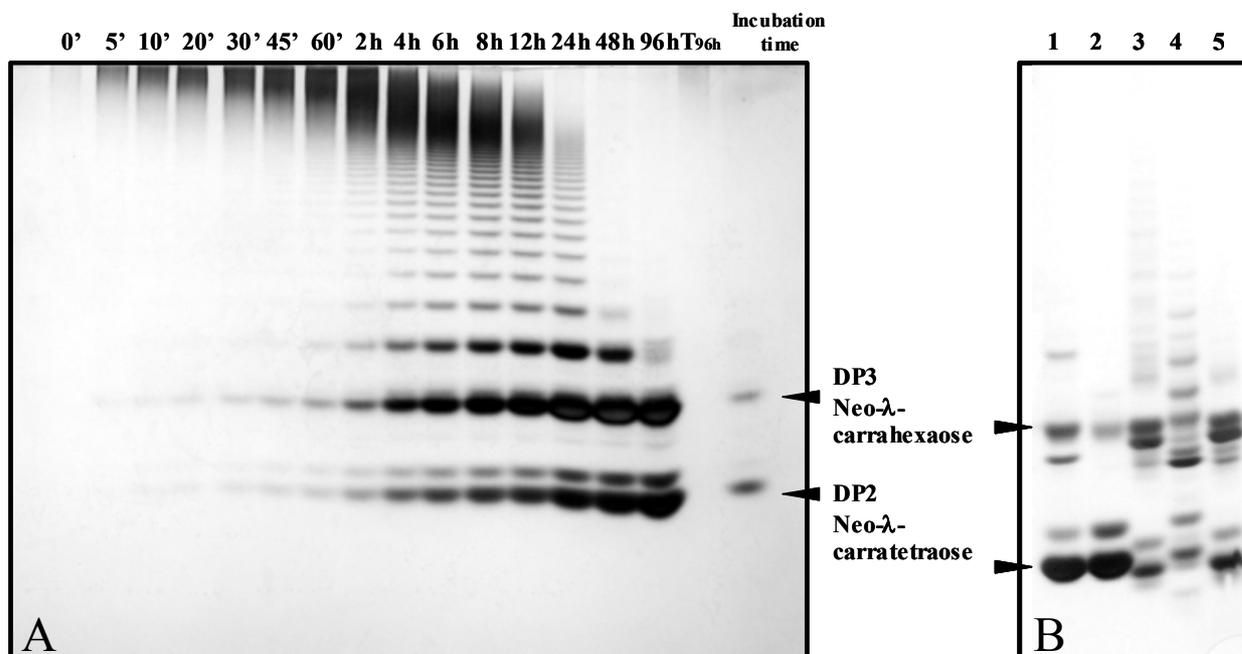


FIG. 2. Carbohydrate-PAGE analysis of *P. carrageenovora* λ -carrageenase products.

A, Time course of the hydrolysis of λ -carrageenan by λ -carrageenase (1.4 UA/mg of λ -carrageenan extracted from *Gigartina skottsbergii*). Reaction was run in buffer B. At intervals, aliquots from the reaction mixture were quickly frozen, in order to boil all of them at the same time. Same quantity of each sample were loaded on the gel. The lane marked T was loaded with a 96h old-reaction made with heat-inactivated enzyme. Both oligosaccharides used as size markers were HPLC purified. The enzyme came from a sephacryl S200 step pool.

B, Comparison of completion hydrolysis by λ -carrageenase of carrageenans from the family λ from different sources: 1, *Gigartina pistillata* and *acicularis* (Sigma); 2, *Gigartina skottsbergii* (CPkelco); 3, *Gigartina acicularis* (Sanofi Bioindustries); 4, *Chondrus crispus* (SKW biosystem); 5, *Gigartina pistillata* (extracted according to Amimi et al., 2001). Reactions were run in NaNO_3 0.1M during 48h with 3.8 UA/mg of polysaccharide. The enzyme came from a phenylsepharose 6FF step pool concentrated by ultrafiltration on a 10kDA cut-off membrane.

Fig. 3. Nucleotide and deduced amino acid sequences of *P. carrageenovora* λ -carrageenase (*cgIA*). The putative -35 and -10 boxes in the promoter region, and the potential ribosomal binding site, known as Shine-Dalgarno sequence (SD) are shown in bold. The potential hairpin loops are indicated by arrows with their own respective minimum free energy below. The open reading frame starts at nucleotide 217. A 25-aa signal peptide is proposed in bold. The underlined amino acids regions (sequence A, B, and C) were confirmed by protein sequencing which allowed us to design the PCR primers corresponding to the underlined nucleotide sequence. This open reading frame leads to a putative preprotein of 105.6 kDa.

```

GTATATGTATTATAAATAGAGGTTTCCTTATGATTTTAGGTTTAAATGTACTCAGTAAACCAATGAAAAAGGATTGTTTATGCTGTTTTTTCCTGAAAAAT 100
AATAATGCAGCTAATTTTTGACGCTCTTATTTTAAAAAGCATTAAACTCAATATTTAGGAAAATAGGTACAATCTAATTCCTCAACACGTTATAAATAACT 200
                                     -35          -10          +1
SD
TAATGTAAAGGAATGTTATGAAAAATAAAATTCATCTGCAATGGTAGCTAGCTCGTTATAAATGGCTGCGTTATCCCTACCCTTAAAGCTTCTCAATCG 300
      fM K I K I L S A M V A S S L L I G C V I P T V K A S Q S 28
Forward A
GCTATTAAAAGTATTGAAACAAACCGAACAATTAATAAGTTAGAACAGGAATGTTGAGTGGAGGCTCAATCAATCATAACTACAAGCTATGAAGGGACTG 400
A I K S I E T N R T I T K V R T G M L S G G S S I I T T S Y E G T 61
A
TAGCTGCATATAAGTTAATGGAGAAAACTGTGGAAAAATGAACCTCTCGGGTTTTATGAATCATGATATTTGGGTTCAAGATATTAATGGTGATGGACT 500
V A A Y K F N G E K L W E N E L S G F M N H D I W V Q D I N G D G L 95
TGTAGAGATATTTGCTGCGAATGCCGATGGCAATGTTTACTGTATTAATAGTGATGGTTCTTTAAAGTGGACGTTTGGTCTAAATGAAGTCCCTATGAAC 600
V E I F A A N A D G N V Y C I N S D G S L K W T F G L N E V P M N 128
TCTGTAACGTAAATCTCTGATGCAGATAAAAAGTATGTTGTGGCAGGTGGTTACGATAAAAACCTTGTATTACATATCGACTAATGGTGAACCTTTAAAAA 700
S V T V I S D A D K K Y V V A G G Y D K N L Y Y I S T N G E L L K 161
CAATGAATCAGGTACTTACTCAGAAGAAGGGGTGTTTGGGGATGGCGTCAAGCCTGAAGCTCGAACTCATACTGTAATAATTCGTCCTCCAGTAAAAATC 800
T I E S G T Y S E E G V F G D G V K P E A R T H T V N F V R P V K S 195
TAGTGACGGCACTGAAAACTAGTAGTTCTAGGGACTAATAATAGCCTTCAGAGTTCTGGTTCGGTCTATATTTTGAACCATTTGCTGATTTACCCAGT 900
S D G T E K L V V L G T N N S L Q S S G R F Y I F E P F A D L P S 228
GAAAAGAGTCGAATATCTATTAATAAAGGGGATAGGAGACCTTCGCCTGTAGATTTTGATAATGATGGTAAATGACGAACTTACGCTAGGTAATAGTGCAC 1000
E K S R I S I K K G I G D L R T V D F D N D G N D E L T L G N S A 261
AGATAGGTGATGCTGCAATTTCTGTTATGAATTTGGATGATCTGTACAGAGAAGAAAGCCAAATTAATGACATTGCACGTAGAAATAGATCGCTTTGGGTA 1100
Q I N D A I D A A I S V V M N L D D L S Q K K S Q I N D I A R R I A R R I A R R I A R R I A R R I A R R I A R R I A R R I 295
TAGGGTTGCTCAAAAGTGTGTTATGAATGAAGCACTCTACACTCTGACCTTTTTGGCTCTAGGATACTTCTAACACCAGAGTCAATTTGACGTG 1200
R V A Q T E V V M N E G T P T Y L T L F G S R I L L T P E S F D V 328
Reverse B
AACGATCTGAAATACTGCAATAAATACTCTTATTATGACATGTGGAAAGATAAAAAGTTCAAATAAGTTGGTTTTAGCTAGTGCACAAAGTGGTGCA 1300
N D S E I L A N K Y S Y Y D M W K D K S S N K L V L A S A Q S G G 361
B
GCCAAGTGCATATTATGATACTTCTAACCCAGTTGGAAGTCTGCCTACGAGGAATTAGAGCCTCAAGTAAATAGCTGCAATACAAGAAAACTAG 1400
S Q V H I I D T T S N P S W K S A Y E L E P Q G K L A A I Q E N T R 395
AGCAATGAAAGGCAACTATCTAACTTTCAAAAGCCTACGCGAGAGCGTCTCCTTTGCCCGTGTATTTTATTTTCAGAAAGTAGAAATGAAATCCAAACA 1500
A I E R Q L S N F Q K P T R E R A P L P V Y F I S E S R N E I P T 428
ACAATCGAGCGAAGTGAGTTTTTGTATGACTCACCAGTATTTTTAAATTACTCAACTCTACCAAATGTAGAGAAGTGGGATAGAAGTGAAGTCTTGCTG 1600
T I E R S E F L Y D S P V F L N Y S T L P N V E N W D R S E V L A 461
Reverse C
ATAATCCTAAATACAGAGATAAGAGACCGTCGTA AAAACTATACACTATCATCTGAAGAAATGTTAATAAGCTTTCGGCTGGATATGATAATAGCGA 1700
D N P K Y R D K R D R R K N Y T L S S E E M F N K L S A G Y D N S D 495
C
TGGGATATCACAATGGCCAGGGCATGGAATGATCCGTATATGATTAGTTTAGCTACTATGAAAAGAATTATTAGTAGTGGCGATGGGAAAAAACGGTC 1800
G I S Q W A G H G N D P Y M I S L A T M K R I I S S G D G K K T V 528
AATATATATCCAGAGATAGAAGGACATGGAGATGCATTTAATAAAGTACTAAGTGATCATTTCTACCCCTTGGCTGAGTTTAGTAGTGAGAATAATGCTA 1900
N I Y P E I E G H G D A F N K V L S D H F Y P L A E F S S E N N A 561
ATTTGTTTATGCGCAATAAACATACATTTTGGCAATCAACTATATATGCTCCAGAGTGGTCGGAGTTACGTTCTGGACGTTTAGTGATGCGTTGTACC 2000
N L F M R N K H T F W Q S T I Y A P E W S E L R S G R L A D A F V P 595
TGCAATGGAAGAAACAACAGATAAATCAATGGAAATGAGTGTGCCCGGAGAATGGGGTTATGGGCTGCAGGCTCTGTAGATAATGGGGTGAAGATAC 2100
A M E E T T D K S M E M S V A G R M G L W A A G S V D N W G E R Y 628
GCAAGAGACAATCTAGTTTTGACGGTTTACGCCAACACTCTCATCAAATGGTCCCAACCATGCCTAAGGCAATAATATATAAGATAGCGTCTGGTG 2200
A R D N P S F D R L R Q H S H Q M V P N H A L R Q I I Y K I A S G 661
CCAGATATATAAATAACTTCGGCTTCAATCAGGAGTATATGAGCCTTGCTGGGAGTTAATGGTAAAGGCGCTTTATGTACCAAAGCGTGAAGAATT 2300
A R Y I N N F G F N Q E Y M S L A W E L I G K G A L Y V P K R E E L 695

```

ATTAAGCTTGTACCAGTTCACATAAGTATGAAAGAACCAGATCCAATCTACAGAGAAACATCAAATAATGTGAAGTGGACTACGTTCTATGATGAAGAA 2400
 L S L S P V H I S M K E P D P I Y R E T S N N V K W T T F Y D E E 728
 AAAGACTCTATTCCATATGTGTTTAGTCGCTTAAATGGTACTTGGCCAGGTGCAAAAACCTTTGCCTTGGGATTACTCTAATTATGCTGCCGATACTAAAG 2500
 K D S I P Y V F S R L N G T W P G A K T L P W D Y S N Y A A D T K 761
 AAAGGAGATTAGATTTTATACCAAAATTTCCCTAAAGGATTAGTTTTAATTACGCCTGTTTCAGCAAGGTAATTTAAAGATGAAGGCACCGTTAGAGGCAC 2600
 E R R L D F I P K F P K G L V L I T P V Q Q G K F K D E G T V R G T 795
 ATTAGCTGATAACATGCACCCTATCTATAAAGATATAATGAAAGAGTATATTACTGACGGTAAAACTACTATAACCCTAATGGTGAGCAAGTAATGGCT 2700
 L A D N M H P I Y K D I M K E Y I T D G K N Y Y N P N G E Q V M A 828
 GCTGATAGTGTAGATACAGACAAATTAATAAAGATGAAGAAAAATCGAATCTTTTACCAATGACTGTATCAGGAGAAGCCGCTTGGGTTGTTGCTC 2800
 A D S V R Y R Q I K N K I E E K S N L L P M T V S G E A A W V V A 861
 AATCTGCTGAAAAGCATTGCGACTGACGCTGGTTGATAGTGGTTATTTAAATCCGAGTAACAAAGTTGCTAAGGTTAAATTTAACTCTGTAACACCTGT 2900
 Q S A E K H L R L T L V D S G Y L N P S N K V A K V K F N S V T P V 895
 AGCGATAGTAGATGTATTGTCAGGTGAGACATTTTCGCCAGATTCAAATGGGGTTGTAGAAATACCGGTCTTAGCTGGTGCTTTTAGGTTCAATTGATGTG 3000
 A I V D V L S G E T F S P D S N G V V E I P V L A G A F R F I D V 928
 AAAATTACTGAAGACCTTAGAAAACATGCAAAGTTCAACATTGTAATTACGGCGTTTAAATTTATAGTGCGGGATTGAAGTTTATTAGTCCTGCCTATTTA 3100
 K I T E D L R N M Q S S T L Ter -16.0 kcal/mol -16.9 kcal/mol 942
 GTCTAAAGTAAGTTGCTCACAAAAGCAGAAATTGCAAGGCGGCTACGAGCGAGAAGCAGACGGTGAGTAAACTCGTTTTTGGACAAAAGCTCGTCAGAG 3200
-16.9 kcal/mol
 TAAATAACCTAGAACCCGCTCTCAGATAACTTCTCGTATCTCAAACGTCAAAGACCACCGTTGTCAGCGGTGGTCGGTTTGC CGGAACTGGACAGTT 3300

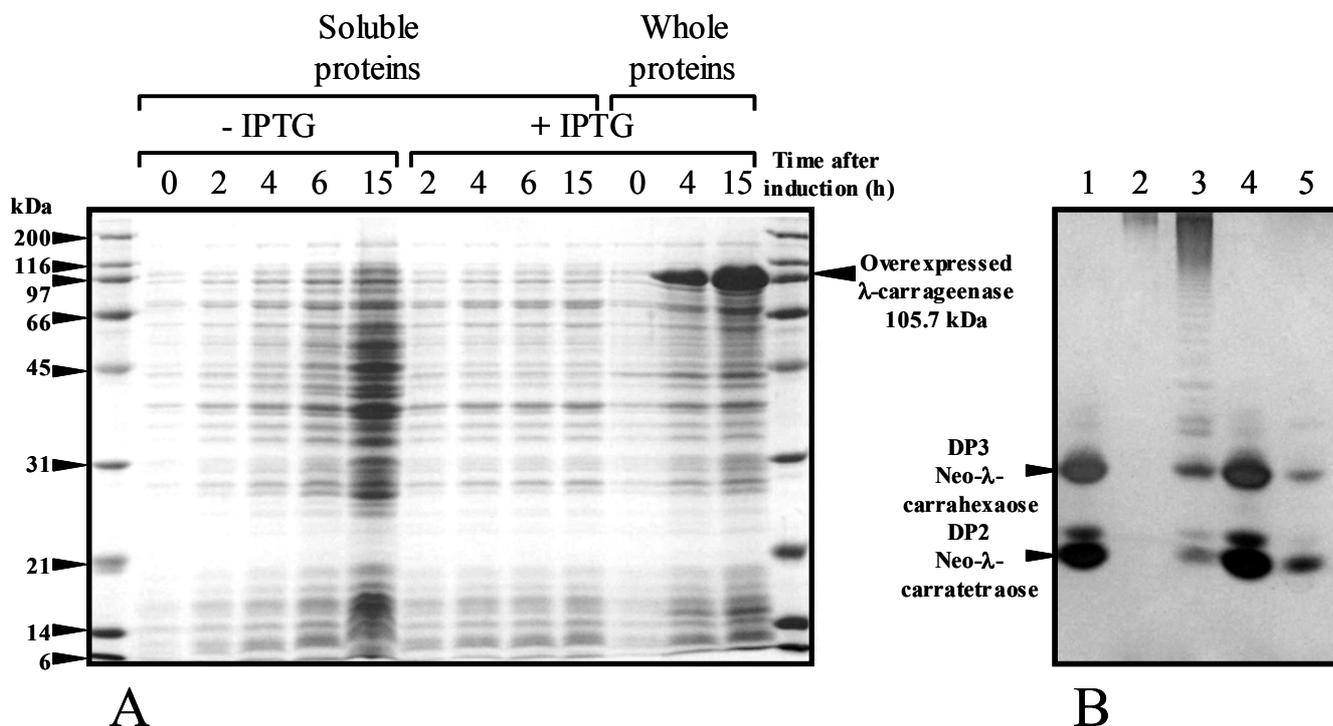


FIG. 4. Analysis of the recombinant *P. carrageenovora* λ -carrageenase.

A, SDS-PAGE analysis of the overexpression. The gene *cglA* of *P. carrageenovora* λ -carrageenase was inserted using Gateway technology (Invitrogen) in the pDEST17 vector without the putative signal peptide in order to produce a fusion protein of 105.7 kDa with a N-terminal 6xHis tag. Here the overexpression was carried out in *E. coli* Rosetta (DE3) pLysS strain (Novagen). Culture was grown on LB at 37°C up to 0.6 OD₆₀₀ and then separated in two batches with fresh media. One of them was induced with 1mM of IPTG whereas the second one was not. Both of them were then grown at 20°C during 15 h and aliquoted regularly. Samples were centrifuged (5000xg, 15 min) and the cellular pellets were lysed with BugBuster reagent (Novagen). “Soluble proteins” lanes correspond to the supernatant (1500xg, 30 min) of such lysat, while “Whole proteins” lanes correspond to the full lysat which contains insoluble proteins as inclusion bodies. Same quantity of each sample were loaded on the gel.

B, Comparison by Carbohydrate-PAGE analysis of λ -carrageenase activity obtained from the recombinant and the purified enzyme. The recombinant enzymes was obtained from purified inclusion bodies (which were produced by overexpression in BL21 (DE3) at 37°C in LB, induced by 1 mM IPTG during 6h), solubilized in Tris buffer 50mM, pH8 containing 8 M urea (Buffer C). Hydrolysis of λ -carrageenan were carried out during 96 h at 30°C (lanes 1 to 3) and several weeks at room temperature (lane 4), with, wild partially purified enzyme (lane 1) from the sephacryl S200 step pool, buffer C as control (lane 2), recombinant enzyme solubilized in buffer C (lanes 3 and 4). The main final products, the neo- λ -carratetraose and neo- λ -carrahexaose, HPLC purified, were loaded on lane 5 as references.

CHAPITRE 3

LA FUÇOÏDANASE FcnA :
CLONAGE & SÉQUENÇAGE DU GÈNE,
ANALYSE DE LA SÉQUENCE,
ANALYSE DU CONTIG,
SUREXPRESSON DE FcnA ET DE SulA.

Concernant le catabolisme des parois de Phaeophyceae (algues brunes) seule celui des alginates (épimérasés et lyases) a été vraiment approfondi (Draget et al., 1997 et 2000 ; Wong et al., 2000) du fait de la très large utilisation industrielle de l'alginate comme gélifiant. Le fucoïdane, second polysaccharide matriciel des parois de Phaeophyceae, très complexe et dépourvu d'applications comme texturant, n'a été quant à lui que très superficiellement abordé (cf. Chapitre 1). Cependant, depuis une quinzaine d'années, l'intérêt porté à la bioactivité des polysaccharides sulfatés a conduit à une intensification des recherches concernant sa composition et sa structure par des méthodes de fractionnement et d'analyses physico-chimiques. Ces méthodes ont été suffisantes pour caractériser les fucanes très homogènes d'échinodermes. Mais appréhender la complexité des fucoïdanes d'algues brunes nécessite désormais l'emploi de biocatalyseurs issus du catabolisme de ce polymère dans le but d'isoler des fractions spécifiques et homogènes. En outre, il s'avère que les produits de tels enzymes pourraient présenter des activités biologiques très intéressantes tant du point de vue de la phytoprotection (Descamps-Lepagnol et al., 1998) que de la pharmacie médicale. A ce jour très peu de ces enzymes ont fait l'objet d'investigations, mais les recherches les concernant s'intensifient depuis une dizaine d'années (cf. Chapitre 1). Plusieurs souches bactériennes ont été isolées. Des lyases, des hydrolases et des sulfatases ont été complètement ou partiellement purifiées. Mais seules deux hydrolases sœurs ont été clonées (à partir d'une γ -protéobactérie), séquencées et surexprimées (Takayama et al., 2002). Dans le brevet les concernant, l'activité précise et les produits de ces dernières n'ont toutefois pas été divulgués, ou simplement élucidés. En revanche, la même société (Takara Bio Inc.) a publié récemment des travaux concernant une fucoïdanase issue de la même souche (Sakai et al., 2004) et tout porte à croire qu'elle corresponde à l'un des enzymes du brevet. Cette compagnie japonaise a également publié des données très précises sur l'action et les produits de deux lyases de Flavobacteriaceae agissant sur des fucoïdanes riches en acides uroniques (Sakai et al., 2003).

L'UMR 7139 s'est attachée depuis maintenant une quinzaine d'années à la description d'enzymes de modification des polysaccharides pariétaux de macroalgues. Un intérêt particulier a été porté sur des enzymes procaryotiques de dégradation des galactanes neutres (agarose) ou sulfatés (κ -, ι -, et λ -carraghénane) provenant des rhodophytaeae (algues rouges). Plusieurs d'entre eux ont été finement caractérisés ainsi que leurs produits (cf. Chapitre 1 et 2). Les travaux de thèse de Descamps-Lepagnol (1998), réalisés en collaboration avec le CNRS et les Laboratoires GOËMAR), ont initié quant à eux l'étude d'une fucoïdanase bactérienne hydrolysant le fucoïdane extrait de la paroi de diverses fucales. C'est dans ce cadre qu'ont été réalisés l'isolement de la souche bactérienne SW5 contenant l'enzyme et la purification à l'homogénéité de ce dernier. Les structures de deux oligosaccharides majoritaires, purifiés à partir de la dégradation à complétion du fucoïdane de *Pelvetia canaliculata* par un pool protéique non pure (fractionnement au sulfate d'ammonium) contenant cette fucoïdanase, ont également été caractérisées. Il s'agit du DP2 et du DP3 de motif $[\rightarrow 4)\text{-}\alpha\text{L-Fucp-2,3-(diOSO}_3\text{)-(1}\rightarrow 3)\text{-}\alpha\text{L-Fucp-2-(OSO}_3\text{)-(1}\rightarrow]_n$. Mais surtout, le pool d'oligosaccharides obtenus par une dégradation identique, quoique moins poussées, s'est avéré déclencher une très forte élicitation des réactions de défense du tabac vis à vis du virus de la mosaïque (Klarzinski et al., 2003 ; cf. Chapitre 1) lorsqu'il était mis au contact des feuilles de cette plante. Ces résultats prometteurs ont poussé les laboratoires GOËMAR à vouloir maîtriser le comportement ainsi que la production de ce biocatalyseur et évidemment confirmer l'activité de ses produits. Le pendant académique de ces recherches repose sur le gain que représente cette activité enzymatique dans l'exploration du fucoïdane, tout autant que sur l'intérêt à disposer d'un nouveau model d'étude glycosidase/polysaccharide.

Mes travaux, succédant à ceux de Descamps-Lepagnol, avaient pour but d'isoler le gène codant pour cette protéine, de le séquencer, pour aboutir à sa surexpression. Ils sont récapitulés dans l'article 3, placé en fin de chapitre, que j'invite le lecteur à parcourir avant d'aborder certains développements et résultats complémentaires présentés et discutés ci-dessous.

1. Clonage et séquençage du gène de la fucoïdanase (FcnA) de SW5

L'obtention de la sonde nucléique permettant le criblage de la banque génomique de SW5 est décrite intégralement dans l'article 1. La synthèse de cette sonde de 203 pb ainsi que son séquençage ont été réalisés antérieurement à ce travail de thèse par Thuault (1999) au cours de son DUT. Cette synthèse se base sur un principe classique également utilisé pour le clonage de la λ -carraghénase (cf. Chapitre 2). Contrairement à cet enzyme, les microséquences de la fucoïdanase ont été obtenues à partir de la protéine pure. Par contre le peptide N-terminale n'étant pas accessible au microséquençage la sonde obtenue est interne au gène de la protéine cible.

Elle a été employée également au cours d'une expérience de Southern blot (Thuault, 1999) sur diverses digestions (par *Hind*III, *Bam*HI, *Eco*RV et *Sau*3AI) de l'ADN génomique de la souche SW5, celui de *Pseudoalteromonas carrageenovora* servant de control négatif. Le résultat de l'hybridation de la sonde fucoïdanase (non montré) sur les profils électrophorétiques de ces digestions révèle la présence d'une seule copie du gène de cette fucoïdanase dans la bactérie SW5.

La banque génomique a été criblée à l'aide de la sonde marquée radioactivement. Trois clones ont répondu positivement pFU11, pFU33 et pFU14, ce qui représente 0,06% de la banque. Après réisolement sur boîte, un second criblage sur 8 colonies isolées de chacun de ces 3 clones a confirmé le premier. Les deux premiers ont tous les deux des inserts de 4 kb. Le dernier a posé des problèmes de culture et de purification du plasmide. Il semblerait que son développement ne soit pas compatible avec la conservation du plasmide. Néanmoins j'ai produit suffisamment de ce dernier pour pouvoir l'analyser. Son insert est estimé à 13 kb.

Etant donné le faible nombre de clones positifs, je me suis attaché à les décrire tous les trois par une cartographie physique. Les inserts des plasmides pFU11 et pFU33, hors mis le fait d'être ligés en sens inverse dans pAT153, sont quasiment identiques comme le montre la figure 34. Cependant, les tentatives d'alignement des cartes de pFU11 et pFU14 (Fig. 35) posent problème. Elles ne se correspondent en effet que partiellement. Ayant démontré la présence d'un seul gène fucoïdanase sur le génome de SW5, j'ai émis l'hypothèse d'un insert composite pour l'un des deux

plasmides. Autrement dit, au moins deux fragments génomiques de SW5 auraient été ligués l'un à l'autre lors de l'insertion dans le vecteur pAT153 de la banque. Or si l'insert de pFU11 est bipartite cela implique que pFU33 le soit aussi et de surcroît au même endroit, ce qui, en terme de probabilité, est défavorable à cette hypothèse. Donc tout laisse présager qu'il s'agisse de pFU14. En outre, cela pourrait expliquer les problèmes de croissance des souches transformées par ce plasmide (par création d'un ORF fonctionnel mais défavorable par exemple).

Conformément aux travaux réalisés pour la λ -carraghénase, j'ai procédé à la localisation de la région de pFU11 et pFU14 où s'hybride la sonde par Southern blot sur divers profils de digestion. Les résultats obtenus (Fig. 36) positionnent la sonde à l'extrémité avale de l'insert de pFU11 (*PvuII*(4200)/*BamHI*(4400)) et sur le fragment de restriction lui correspondant sur pFU14 (*PvuII*(9100)/*ClaI*(10000)). Ils sont donc cohérents.

Afin d'affiner le résultat évoqué ci-dessus, cette extrémité de pFU11 (*PstI*(4000)/*SalI*(4700)) a été sous-clonée dans pAT153 (sous le nom de pAFU11) puis séquencée. Seule la moitié de la séquence de la sonde y est présente puisqu'elle s'interrompt au site *BamHI*(400) après lequel débute la séquence du vecteur (Fig. 36). La conséquence directe de ce résultat est l'impossibilité d'avoir le gène en entier sur pFU11 et donc sur celui de pFU33, ce qui pouvait être préjudiciable au clonage ultérieur du gène. Néanmoins je conservais la possibilité de l'avoir dans son intégralité sur pFU14. Mais contrairement au gène de la λ -carraghénase où la sonde est N-terminale, ici je n'avais aucune idée de l'endroit du gène à laquelle elle correspond. Ceci signifie que les 3000 bases (la longueur putative du gène de la fucoïdanase) en amont et en aval de la sonde était susceptible de contenir ce gène. Or cette zone comprend le point de divergence entre pFU11 et pFU14 (Fig. 36). A ce stade, la question était donc : le gène de la fucoïdanase débute-t-il avant ou après ce point critique *HindIII*(2900) ?

Cette interrogation a donc guidé les étapes de sous-clonage/séquençage, la démarche étant de remonter jusqu'à l'extrémité N-terminale du gène sur pFU11. Les parties de l'insert de pFU11 sous-clonées et séquencées sont représentées sur la figure 37. Il s'agit des clones pAFU11 contenant le fragment *PstI*(4000)/*SalI*(4750),

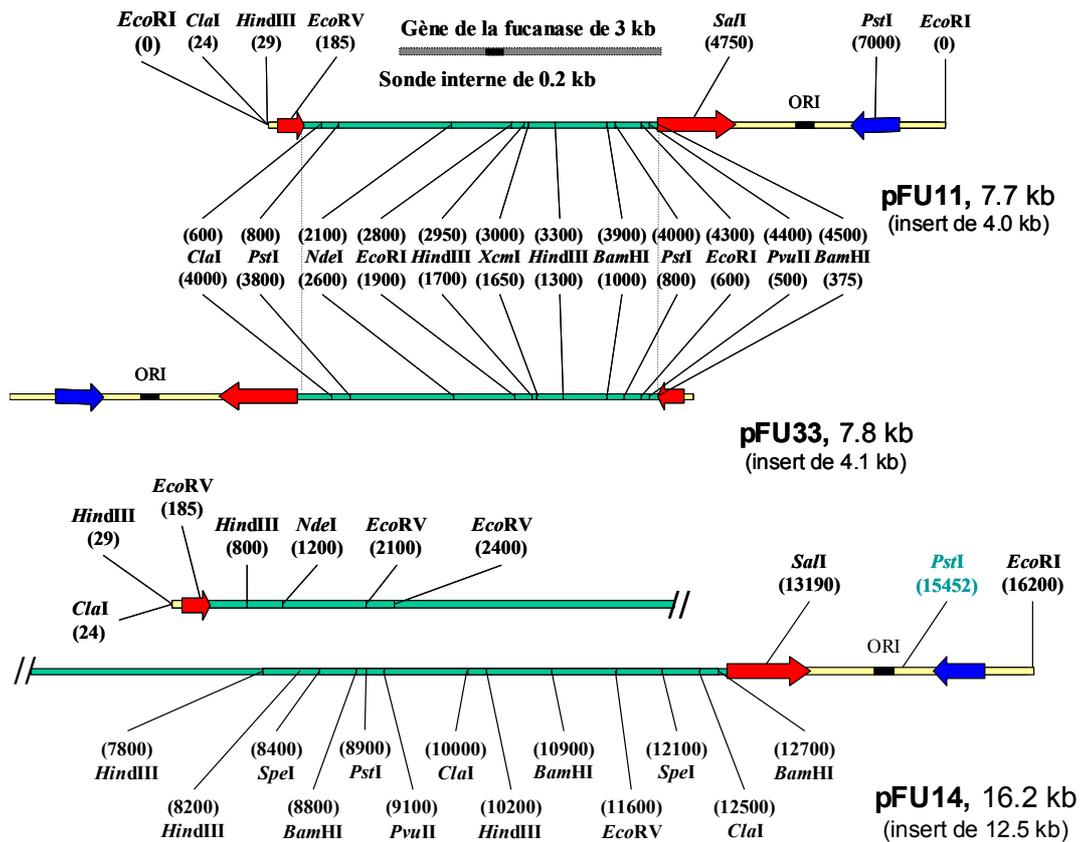


Figure 34 : Cartes de restrictions des plasmides pFU11, pFU33 et pFU14.

Les inserts sont représentés en vert, le jaune correspondant au vecteur pAT153 dont en bleu le gène de la β -lactamase et en rouge celui de la résistance à la tétracycline.

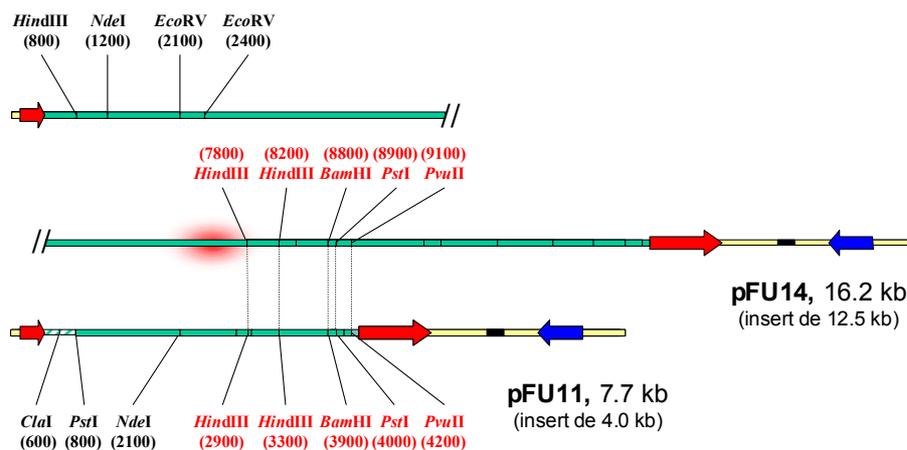


Figure 35 : Mise en évidence de la correspondance partielle des inserts de pFU11 et pFU14. L'hypothèse d'un insert composite chez pFU14 avec une jonction dans les 800 bases amont au site *HindIII*(7800).

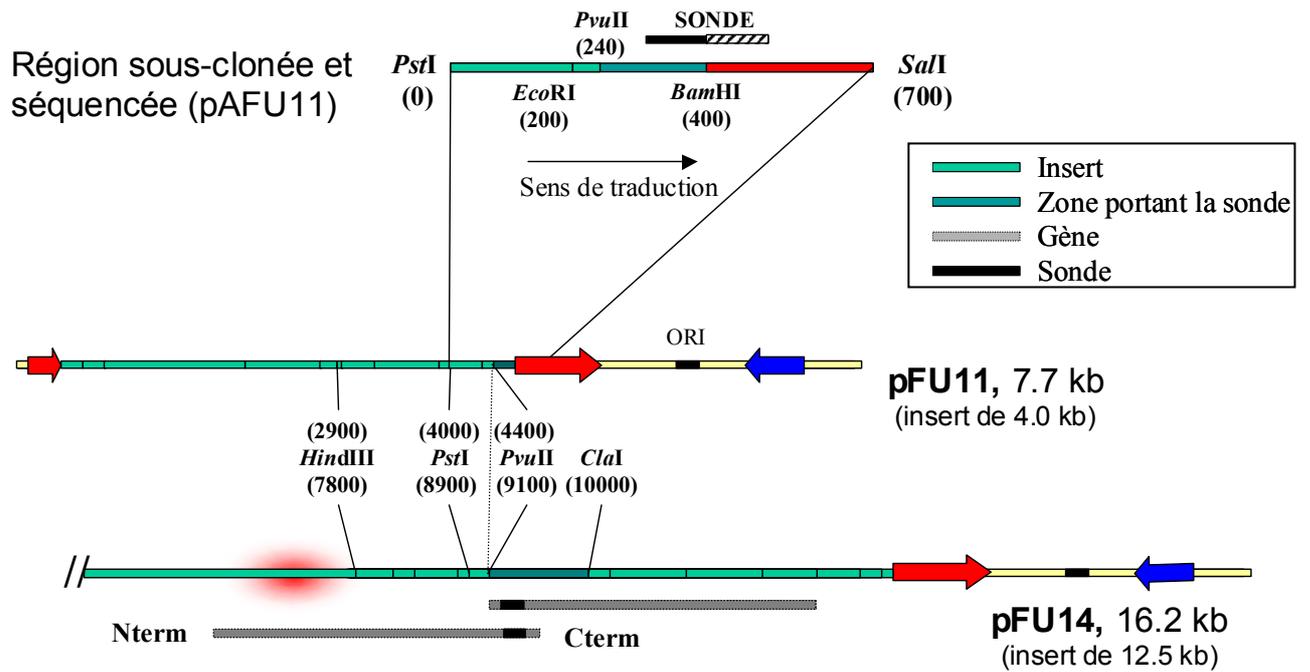


Figure 36 : Localisation et orientation de la sonde sur les inserts de pFU11 et pFU14 : conséquences sur les chances d'avoir le gène de la fucoïdanase d'un seul tenant.

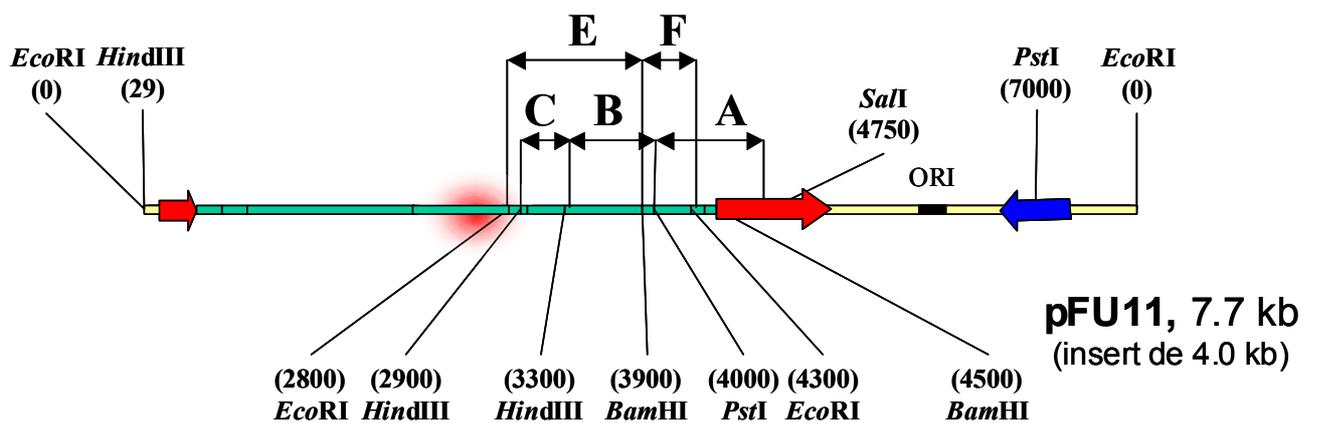


Figure 37 : Représentation des sous-clonages réalisés dans le cadre de la recherche du gène de la fucoïdanase.

pBFU11 contenant le fragment *Hind*III(3300)/*Pst*I(4000), pCFU11 contenant le fragment *Hind*III(2900)/*Hind*III (3300), de pFFU11 contenant le fragment *Bam*HI(3900)/ *Eco*RI(4300) et enfin pEFU11 contenant le fragment *Eco*RI(2800)/*Bam*HI(3900).

L'arrivée d'un séquenceur capillaire au laboratoire a grandement facilité ce travail, me permettant de séquencer par gène « walking » les inserts de pFU11 entièrement et pFU14 seulement partiellement. En effet, en remontant en amont de *Eco*RI(2800) sur pFU14 et pFU11, j'ai identifié le point de divergence entre les deux inserts. Le cadre de lecture de la fucoïdanase restant malgré tout ouvert au delà de ce point sur les deux inserts, il allongeait le gène respectivement de 342 bases ou 9 bases selon que je suive pFU11 ou pFU14. Dès lors, rigoureusement, deux hypothèses étaient toujours envisageables : soit pFU11 était composite soit c'était pFU14. Par ailleurs, la séquence obtenue à partir de pFU14 en aval de *Eco*RI(2800) correspond exactement à celle issue de pFU11. Au delà du site *Sau*3A1(4500), limite aval de ce dernier, le cadre ouvert de lecture de la fucoïdanase se poursuit sur pFU14 le long de 1068 bases. Autrement dit, la zone commune aux deux inserts faisant 1611 bases, selon que l'on soit dans l'hypothèse longue ou la courte, nous avons un gène de 3021 ou 2674 bases (1007 ou 896 résidus). Or la masse de la protéine sauvage purifiée a été estimée entre 105kDa et 143 kDa (Descamp-Lepagnol, 1998). Donc là encore pFU14 apparaissait être constitué d'un insert mixte et signifiant ainsi que le début du gène ainsi que toute la partie amont se trouvait sur pFU11. Sur la base de cette hypothèse, j'ai conçu des amorces de part et d'autre du gène long putatif (3021 pb). Et j'ai obtenue par PCR, sur l'ADN génomique de SW5, le fragment de taille attendue, confirmant d'une part l'intégrité du gène dans ce génome ainsi que sa longueur. Ce dernier résultat assura par ailleurs de pouvoir obtenir le gène entier par PCR en vue de son clonage. L'intégralité de l'insert de pFU11 a donc été séquencée ainsi que l'ensemble de la partie en aval disponible sur pFU14. J'ai ainsi obtenu un contig de 7485 nucléotides, contenant trois cadres de lecture dans leur entier et deux partiels. La figure 38 récapitule l'ensemble du travail de séquençage réalisé et les cinq cadres de lectures (ORF) mis en évidence dans ce contig. L'ensemble des séquences est disponible en annexe 2.

J'ai également procédé à un essai de séquençage directement à partir de l'ADN génomique de SW5. Les longueurs de lecture se sont avérées assez courtes traduisant la sensibilité de cette technologie de séquençage à la qualité de la matrice employée (diversité et nombre des sites d'initiations). Néanmoins j'ai réussi à gagner 237 pb (79 résidus) à l'extrémité 3' du contig (ORF5) alors que les essais de séquençage en 3' n'ont pas aboutit. Je n'ai pas prolongé ces essais et les ORF1 et 5 restent donc partiels.

2. Analyse de séquence

2.1. Analyse du contig

Les cinq cadres de lectures obtenus sont colinéaires (Fig. 38, Annexe 2). Les caractéristiques générales de ces gènes putatifs sont rassemblées dans le tableau 7. Il apparaît que le gène de la fucoïdanase est précédé d'une sulfatase putative partielle, d'une carboxylestérase putative, et suivi d'une sulfatase putative elle même suivie d'une fucosidase putative (voir ci-dessous). La fucoïdanase dégradant un polymère de L-fucose sulfaté pouvant porté des substituant acétyles estérifiés sur les fonction hydroxyles du sucre, il semble que l'on ait un lien métabolique très solide entre ces cinq gènes. Il parait donc intéressant de considérer l'hypothèse d'un cluster fonctionnel voire d'un opéron. Les distances séparant les cadres de lecture les uns des autres, bien que n'étant pas un critère très significatif, sont suffisamment courtes pour ne pouvoir exclure l'un ou l'autre de ces gènes d'un opéron putatif. La recherche de séquences promotrices de transcription en amont de la fucoïdanase n'a pas été réalisée. Aucune étude ne nous permet d'affirmer que les séquences connues chez *E. coli* soient suffisamment conservées avec celles présentes chez les Flavobacteriaceae. De surcroît, la distance qui les sépare du codon initiateur est également inconnue.

Disposant de quatre débuts de gène, j'ai également procédé à l'étude des séquences de fixation au ribosome (RBS) assurant l'initiation de la traduction, ceci dans le but de mettre en évidence des conservations et de confirmer les ATG initiateurs (Fig. 39). Paradoxalement, aucune de ces séquences n'affiche le motif Shine-Dalgarno, alors qu'il est bel et bien présent sur l'extrémité 3' du 16S des Flavo-

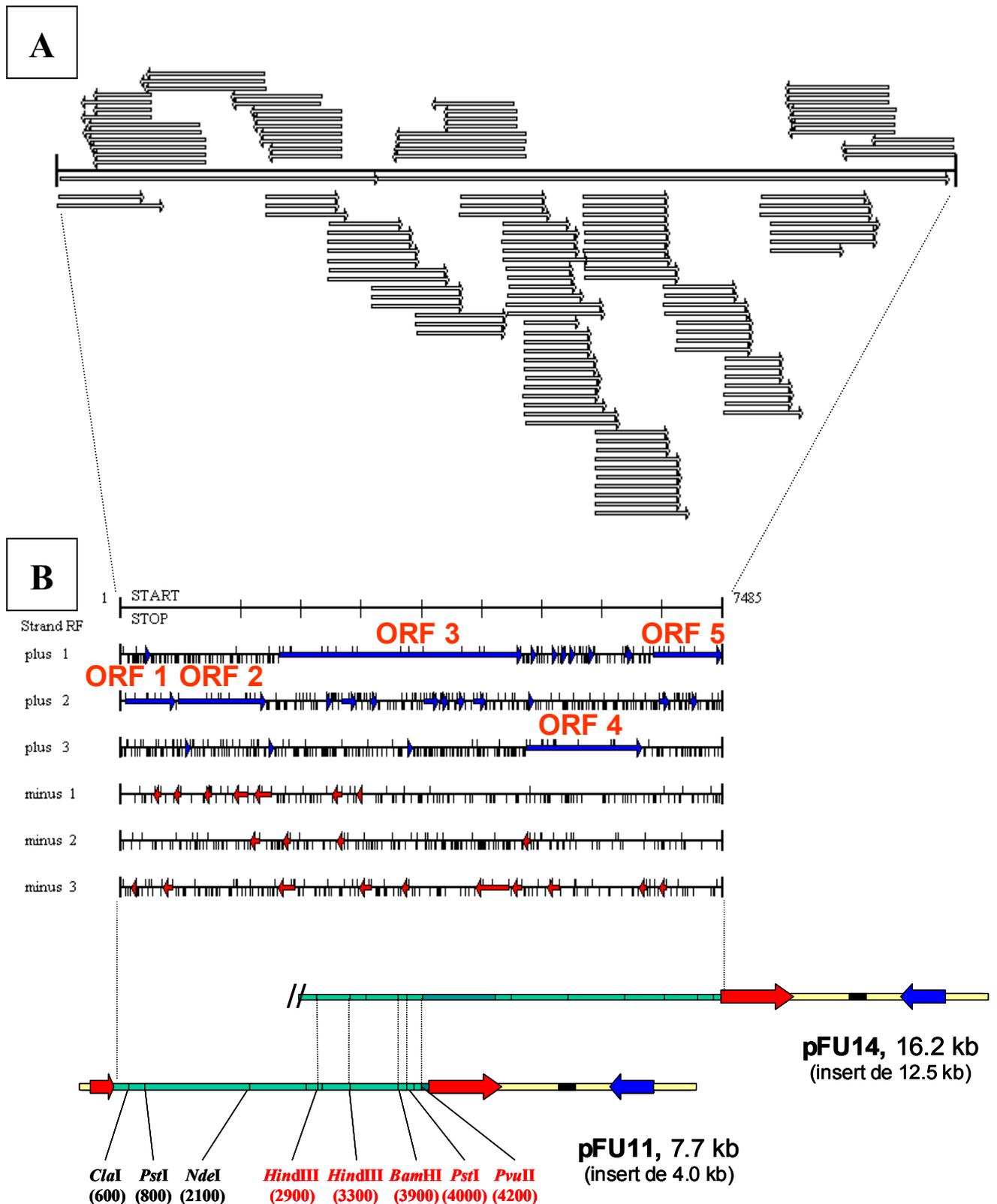


Figure 38 : Séquençage des inserts de pFU11 entièrement et de pFU14 partiellement. L'ensemble des séquençages réalisés est récapitulé dans la partie A. Le contig obtenu traduit selon le code génétique universel révèle 5 cadres ouverts de lecture colinéaires dont 2 sont partiels aux extrémités du contig (B).

A

```
FUCANase ATG4      ATCCAAACCAAGGATTACGTGCCGAAATGGATGCGAGGCGCATTGGGCATG
FUCANase ATG3      GGCACAAGTACCAGATCCAAACCAAGGATTTACGTGCCGAATGGATG----
FUCANase ATG2      AAAAACTACCAACTAAAAAGTCTAATTTATGGTTCCTTATG-----
FUCANase ATG1      GAATATCAAACCTAACTATATACTAACTAAATAAATCTTTTTTATTATG-
SULFATase          TAAATTATCTAATAAATAGATAAATAACTGTATTTATAAATAAAAATGATG
FUCOSIDase         GAAGTTACATGTAACCAATAAAAAATAAATAGTATATATAGACAGATG
ESTERase           AAAGCAAAGAATAGCATAGAGTTAAAGTAAATGGTAACATTTAAATATG
EPSILON (458 to 466 sur 16S Ecoli) TTAACCTTA
```

B

```
FUCANase ATG4      ATCCAAACCAAGGATTACGTGCCGAAATGGATGCGAGGCGCATTGGGCATG--
FUCANase ATG3      TACCAGATCCAAACCAAGGATTACGTGCCGAATGGATG-----
FUCANase ATG2      AAAAACTACCAACTAAAAAGTCTAATTTATGGTTCCTTATG-----
FUCANase ATG1      TATCAAACCTAACTATATACTAACTAAATAAATCTTTTTTATTATG-----
SULFATase          ATTATCTAATAAATAGATAAATAACTGTATTTATAAATAAAAATGATG-----
FUCOSIDase         CAGAAGTTACATGTAACCAATAAAAAATAAATAGTATATATAGACAGATG
ESTERase           AAGCAAAGAATAGCATAGAGTTAAAGTAAATGGTAACATTTAAATATG---
OMEGA (1343 to 1355 sur 16S Ecoli) ACAATTACTATTT
```

C

```
FUCANase ATG4      CAAACCAAGGATTACGTGCCGAATGGATGCGAGGCGCATTGGGCATG---
FUCANase ATG3      AGGCACAAGTACCAGATCCAAACCAAGGATTACGTGCCGAATGGATG---
FUCANase ATG2      GAAAAAACACTACCAACTAAAAAGTCTAATTTATGGTTCCTTATG----
FUCANase ATG1      AGAATATCAAACCTAACTATATACTAACTAAATAAATCTTTTTTATTATG
SULFATase          AATTATCTAATAAATAGATAAATAACTGTATTTATAAATAAAAATGATG--
FUCOSIDase         AAGTTACATGTAACCAATAAAAAATAAATAGTATATATAGACAGATG-
ESTERase           AAAGCAAAGAATAGCATAGAGTTAAAGTAAATGGTAACATTTAAATATG
Shine-Dalgarno (16s end From G. algens) AGAAAGGAGGTGTTCCAGCCGCA
```

D

```
FUCANase ATG4      ATCCAAACCAAGGATTACGTGCCGAATGGATGCGAGGCGCATTGGGCATG
FUCANase ATG3      TAGAGGCACAAGTACCAGATCCAAACCAAGGATTACGTGCCGAATGGATG
FUCANase ATG2      TTATGAAAAAACACTACCAACTAAAAAGTCTAATTTATGGTTCCTTATG
FUCANase ATG1      ATCAAACCTAACTATATACTAACTAAATAA-----ATCTTTTTTATTATG
FUCOSIDase         TTACATGTAACCAATAAAAAATAAATAGT----ATATATAGACAGATG
SULFATase          TAAATTATCTAATAAATAGATAAATAACTGTATTTATAAATAAAAATGATG
ESTERase           CAAAGAATAGCATAGAGTTAAAGTAAATGGT----AACATTAATATGATG
kappase            CCATAGCCTCTACTTTAGTTCGGCTAACAAAA----ACAATTATAGTAATG
Iotase cgiA        TACACTATATATGCGACATTATTAACCT-----TAAATCTTACAATG
Orf2 suivant cgiA  AATAAAAAGCAGCTGTCCCTTATTAACGG-----GCGGCTGCTTTATG
β-agaraseB agaB    CCCTTTTTATCCITTTTTCATTTTAAACAAA----TTACGTATAAACATG
β-agaraseA agaA    TTGTTTGTGAAATACATTTTTATTAACCCCT----AAAATACATATCATG
dnaJ               ATGCTGAGATTTACGTCTTCGTTTAAAT-----AAACATCAGATG
```

Figure 39 : Etude des séquences directement amonts au codon initiateur des ORF 2 (estérase), 3 (fucoidanase), 4 (sulfatase) et 5 (fucosidase) et comparaison aux ORFs identifiées dans *Zobellia galactanivorans*. A : alignement avec la séquence RBS epsilon ; B : alignement avec la séquence RBS oméga ; C : alignement avec la séquence RBS Shine-Dalgarno ; D : alignement des séquences entre elles (rouge >80% de conservation, vert >60% et jaune <40%).

bacteriaceae, groupe taxonomique auquel appartient la souche SW5 (Fig. 17). Seul la séquence amont à la fucosidase présente le minimum requis de quatre identités. J'ai donc poussé l'investigation en cherchant les motifs de type epsilon et oméga bien que leur réalité fonctionnelle soit toujours controversée (Kosak, 1999). Concernant le motif oméga, des homologies intéressantes sont présentes mais à des distances variables de l'ATG, ce qui atténue leur intérêt. C'est le motif epsilon qui paraît le plus consistant avec de fortes correspondances pour la fucoïdanase et la sulfatase, toutes à distance égale de l'ATG. J'ai enfin aligné ces séquences entre elles en minimisant la création de gap de façon à renforcer le rôle de la distance au codon initiateur. Cet alignement révèle des motifs communs forts mais ne ressemblant à rien de connu. En outre les mêmes recherches ont été effectuées sur les séquences amont au deuxième ATG de chaque gène sans apporter de résultats plus convaincants.

Tableau 7 : Récapitulatif des analyses réalisées sur les ORFs du contig de SW5.

* recherche par le programme BLAST (paramètres par défaut) appliqué à GENBANK (septembre 2004).

** recherche sur Interpro, via le programme Interproscan.

*** recherche selon alignement et le programme de prédiction SignalP v2.2.

NOM	TAILLE (résidus)	POSITION (bases)	MW (kDa)	PLUS PROCHE HOMOLOGUE DANS GENBANK*	DOMAINES IDENTIFIES**	PROBABILITE DE PEPTIDE SIGNAL***
RF1	230	2 à 691, partiel	-	31 % d'identité arylsulfatase (NP_863771, 608 résidus) <i>Rhodopirellula baltica</i>	arylsulfatase (IPR000917)	-
ORF 2	364	728 à 1819, entier	41.4	48 % d'identité estérase (NP_870897, 393 résidus) <i>Rhodopirellula baltica</i>	motif EF-hand (IPR002048) estérase (IPR000379)	très probable
ORF 3	1007	1987 à 5007, entier	110.3	-	Cadherin-like (IPR006644)	très fortement probable
ORF 4	486	5052 à 6509, entier	54.6	42 % d'identité arylsulfatase (NP_864421, 495 résidus) <i>Rhodopirellula baltica</i>	arylsulfatase (IPR000917)	très fortement probable
ORF 5	287 (366)	6625 à 7485 (7722), partiel	-	33 % d'identité α -L-fucosidase (NP_905969, 674 résidus) <i>Porphyromonas gingivalis</i>	α-L-fucosidase (IPR000933)	probable

2.2. Analyse de la séquence protéique de la fucoïdanase FcnA

La traduction selon le code génétique universel de ce cadre de lecture donne la séquence d'une préprotéine de 110425.5 Da dont le pH isolélectrique serait a priori acide (variation selon les méthodes de calcul: MWCALC donne 4,7 , ProtParamTool donne 4,4 , et ABI donne 4,2; les liens pour ces outils en ligne sont disponibles sur le serveur INFOBIOGEN). La protéine a été segmentée en quatre régions (cf. ci-dessous). L'ensemble de ces régions est conforme à ce caractère acide à l'exception du domaine C-terminal. Tous ont un pHi putatif légèrement supérieur à 4 alors que le pHi de ce dernier est estimé à 9,2. Ceci s'illustre par la partition des résidus acides (A) et basiques (B) suivante : A/B de 2 sur le domaine N-terminal, A/B de 3 sur les trois répétitions, A/B de 2 sur la zone inconnue, et $A/B < 1$ pour la région C-terminale.

La protéine sauvage étant purifiée à partir du surnageant de culture, il est fort probable qu'elle soit excrétée dans le milieu extracellulaire. Conformément à l'étude réalisée pour la λ -carraghénase, j'ai donc cherché l'existence d'un peptide signal via les programmes de prédictions PSORT, SIGNALP v1.0 et v2.0. PSORT ne prédit aucun peptide signal alors que la première version de SIGNALP suggère une coupure entre le résidu A28 et Q29. Elle est confirmée par la deuxième version qui utilise deux méthodes de calcul indépendantes. Je ne dispose pas de l'extrémité N-terminale microséquencée de la protéine sauvage. L'expérience ne peut donc pas corroborer cette prédiction d'autant que la protéine sauvage n'a pas non plus été analysée en spectrométrie de masse. J'ai néanmoins poursuivi l'analyse sur la base d'une protéine mature de 979 résidus (sans le peptide signal).

La recherche de séquences homologues par BLAST sur GENBANKnr ne donne là encore aucun résultat consistant sur la séquence entière. De même que pour la λ -carraghénase, aucune protéine homologue pleine longueur n'y est à ce jour répertoriée. Cette séquence est relativement longue, elle est donc probablement organisée en plusieurs domaines. Mais, la recherche par RPS-BLAST de domaines/motifs répertoriés dans CDD (serveur du NCBI) n'a donné aucun résultat probant. Cependant, le programme RADAR (serveur de l'EBI) a conduit à l'identification de trois répétitions contiguës situées au centre de la séquence. La deuxième répétition a été

reconnue par la composante SMART de Interproscan (dans la « méta-banque » de domaine/profil Interpro, serveur de l'EBI). J'ai donc procédé à des recherches d'homologie plus sensibles par PSI-BLAST, sur ces répétitions directement, puis en amont et en aval de cette région. La région C-terminale est partagée par de nombreuses protéines appartenant toutes au phylum bactérien des Bacteroïdètes. La recherche d'homologie sur la banque PAT (serveur NCBI, banque de séquences brevetées) a révélé deux protéines homologues partageant la région amont à ces trois répétitions. Il se trouve que ces protéines sont les deux fucoïdanases identifiées par la société TAKARA (Takayama et al., 2002). L'ensemble de ces résultats est développé ci-dessous. Mais la première remarque concernant ces recherches est que, contrairement à une idée reçue, GENBANKnr n'est pas exhaustive concernant les séquences disponibles publiquement.

2.2.1. Trois répétitions contiguës

Le programme RADAR disponible sur le serveur de l'EBI a permis la détection de trois répétitions disposées en tandem de 105 résidus chacune. La première commence à G₄₁₉, la seconde à G₅₂₆, et la dernière à G₆₃₂ pour finir à G₇₃₆, la fin de l'une étant exactement le début de l'autre. Elles partagent 31 % (rep. 1/2), 27 % (rep. 1/3), et 48 % (rep. 2/3) d'identité de séquence (Fig. 40). Le comparateur de profil SMART (via Interproscan du serveur de l'EBI) reconnaît sur la seconde répétition un domaine répertorié sous le nom de CDAG pour « Dystroglycan-type cadherin-like domains » (SM00736, IPR006644). Cette famille présente des homologues chez les Métazoaires, les levures et les Bactéries et serait impliquée dans des reconnaissances homophiles calcium-dépendantes (Dickens et al., 2002). Les résultats de PSI-BLAST confirment la proximité de ces trois répétitions avec les domaines eucaryotiques de type cadhérine qui sont fonctionnellement et structuralement caractérisés. Les scores sont d'autant plus significatifs que la séquence soumise contient plus de répétitions. Les domaines cadhérines sont en effet toujours disposés en tandem illustrant ainsi la notion de supra-domaine bien que le nombre de répétitions soit variable.

Les meilleurs scores de similitude obtenus dans GENBANKnr proviennent exclusivement de protéines bactériennes présentant une organisation en tandem similaire. Les 50 meilleurs scores appartiennent tous à des séquences de Protéobactéries, le meilleur provenant de *Microbulbifer degradans* (Fig. 40, 45 et 48). Les répétitions de FcnA partagent entre 9 % et 28 % d'identité de séquence avec les autres domaines cadherin-like alignés dans la figure 40 provenant des Protéobactéries (*Shewanella oneidensis*, NP_715831; *Microbulbifer degradans*, ZP_00067001; *Geobacter sulfurreducens*, NP_953122; *Magnetococcus sp.* ZP_00045566; *Vibrio parahaemolyticus*, NP_798012), Cyanobactéries (*Nostoc sp.*, NP_490234), Planctomyces (*Pirellula sp.*, NP_864105). Si on retrouve ces domaines dans de nombreux taxons bactériens, curieusement, je n'ai identifié qu'une seule protéine homologue et relativement distante dans le phylum des Bacteroïdètes (*Cytophaga hutchinsonii*, ZP_00116997) dont plusieurs espèces ont cependant leur génome séquencé. A ce jour, les domaines CDAG ne sont pas encore répertoriés dans ce phylum bactérien (Fig. 45).

Les prédictions de structures secondaires réalisées via les programmes PSIPRED (McGuffin et al., 2000), PROFSEC (version améliorée de PHDsec, Rost, 1996), et SAM-t99-2D (Karplus et al., 1998) (accessibles notamment via le metaserveur BIOINFO.PL) s'accordent à reconnaître dans toutes ses répétitions une organisation en brin β (7 à 8) quasi exclusive. De plus, dans une large proportion, les modélisations tridimensionnelles par homologie ou par threading, proposées par le metaserveur BIOINFO.PL reconnaissent dans ces domaines un repliement de type β -sandwich Ig-like (ref. SCOP superfamille b.1). Soumises séparément à ces programmes de reconnaissance de repliement, chacune de ces trois répétitions se voit proposer pour meilleur support de modélisation des domaines cadhérines dont quelques structures sont connues chez les eucaryotes (ref. SCOP b.1.6 ; alignés Fig. 40). Le programme de hiérarchisation 3D-JURY du metaserveur (outil de modélisation 3D reconnu comme le plus performant du moment par le projet Livebench d'évaluation en continu ; Ginalski et al., 2003; Rychlewski et al., 2003), utilisé avec les paramètres par défaut, attribue cependant des scores inférieurs (47) au seuil de con-

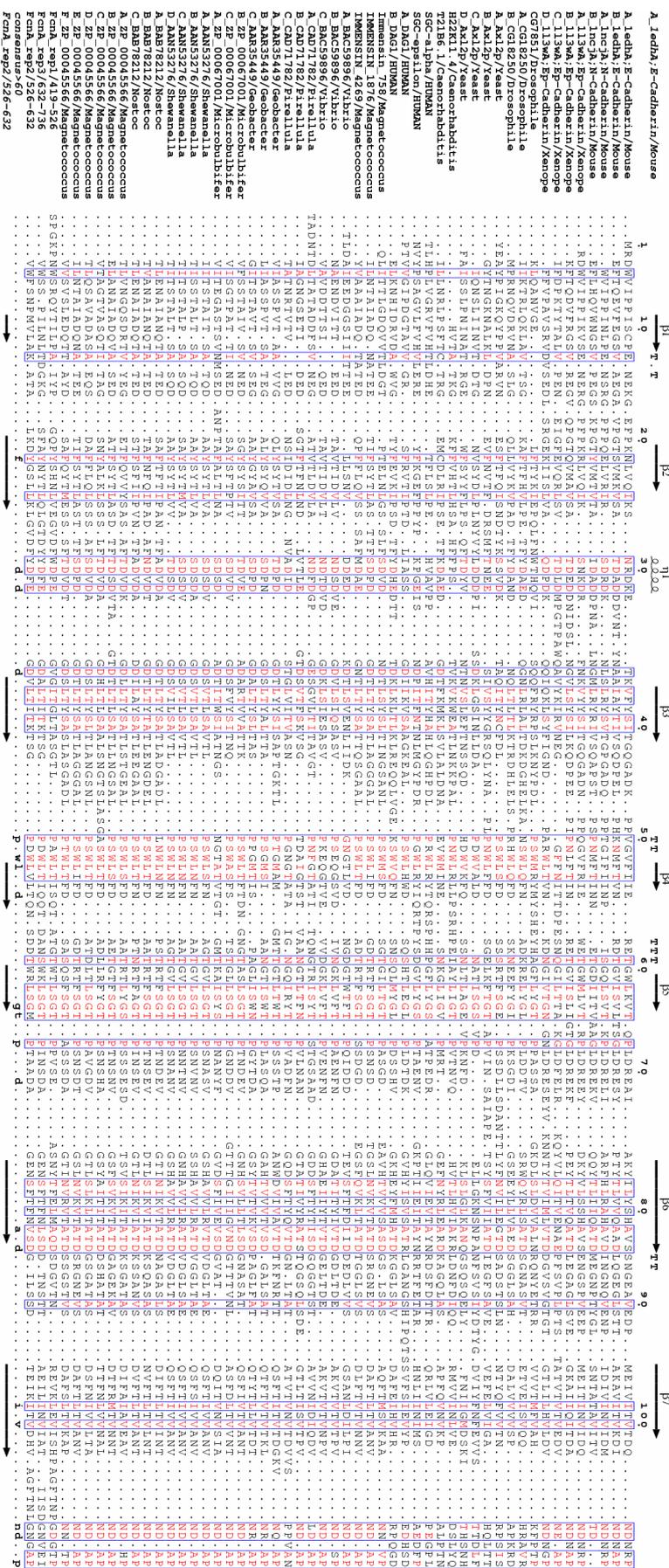


Figure 40 : Evidence de la présence de domaines cadherin-like dans FcnA. Alignement multiple de domaines cadherine homologues :

- de structures connues dont le code PDB est ledha, Incija, 113wA (8 premières séquences),
- de protéines représentatives du domaine CDAG SM00736 (16 séquences suivantes)
- et d'une sélection des meilleurs hits de GENBANKnr produits par Blast (25 séquences suivantes) avec la requête S417-1740 (FcnA)
 - Les répétitions rep1, rep2 et rep3 (3 dernières séquences).

Pour un même numéro d'accession, toutes les séquences de cadhérines présentées ici forment des répétitions contiguës ce qui souligne les conservations des charnières (placées à droite). Les structures secondaires de ledha sont alignées au dessus. Les prédictions de structures secondaires produites par PSIPRED sur rep2 sont alignées ci-dessous. Les triangles rouges indiquent les résidus impliqués dans la chélation d'ions Ca^{2+} . Cette figure a été réalisée à l'aide du serveur ESPRIP (Gouet et al., 1999), l'encadrement des similarités repose sur un taux d'équivalence dont le score global est supérieure à 0,6, et la séquence consensus montre les conservations supérieures ou égales à 60 %.

fiance (50). Mais lors d'une soumission groupée, les trois répétitions ont engendré un score largement significatif (86) à partir de la structure 1q55_A. J'ai observé un effet de groupe identique lors des analyses PSI-BLAST. Si cette consolidation des scores est en partie due à un simple effet d'allongement de la séquence, il est très probablement lié également au rôle structural prépondérant joué par la charnière courte, très conservée, reliant chacun de ces domaines (Fig. 40). On sait en effet que, à l'interface des domaines cadhérines, la charnière participe à la formation de 3 sites de fixation de calcium, permettant de figer cette dernière dans une position qui donne au chapelet une forme de tige rigide. La figure 40 montre une conservation des résidus impliqués dans la chélation de trois ions Ca^{2+} , entre les domaines cadhérines eucaryotes, CDAG, les homologues bactériens proches de FcnA et les répétitions de FcnA.

L'analyse fonctionnelle de ces domaines est moins claire. De toutes évidences il s'agit de domaines impliqués dans des phénomènes d'adhésion. La fonction des cadhérines et des CDAG, établie chez les Métazoaires, est la reconnaissance protéines-protéines éventuellement homophile permettant l'adhésion entre cellules (domaine absent du règne végétal). Mais leur fonction chez les bactéries n'est pas décrite. Cependant, il est intéressant de souligner que la majorité des domaines bactériens homologues font parties de très larges protéines (de 2000 à 12000 résidus) où ces successions de répétitions peuvent être multiples, mêlées à d'autres domaines d'adhésion. Ces protéines sont annotées comme des toxines de type RTX suggérant une localisation à la surface des cellules les exprimant (lien avec la S-layer). On trouve également des séquences homologues dans des protéines annotées comme étant des protéases et des glycosidases (chitinases notamment). Enfin, la proximité de ce repliement avec ceux de certains CBMs (Carbohydrate Binding Modules) attire l'attention. Toutefois aucune homologie de séquence n'existe avec les 42 familles de CBMs répertoriées à ce jour sur le serveur CAZy (Coutinho & Henrissat, 1999).

Pour conclure, si l'on considère l'hypothèse d'une fonction d'adhésion à d'autres protéines, cela pourrait servir à retenir la protéine en périphérie de la bactérie, à la formation de complexe multienzymatique, ou encore à reconnaître une phase protéique liée au fucoïdane. Mais on ne peut rigoureusement pas exclure une fonction de type CBM.

2.2.2. Le domaine C-terminal

Lorsque la séquence de FcnA est analysée par BLAST sur GENBANKnr, les meilleurs scores d'identité de séquence sont obtenus sur une courte région C-terminale de 75 résidus. Les protéines homologues appartiennent toutes à trois classes et sept genres du phylum bactérien Bacteroidetes (Fig. 45). En outre, ce domaine est systématiquement localisé à l'extrémité C-terminale de ces protéines. Des signatures très fortes le caractérisent (Fig. 41). Les génomes de *Porphyromonas gingivalis* et surtout *Cytophaga hutchinsonii* fournissent la plupart des homologues, mais les genres *Zobellia*, *Flavobacterium*, *Microscilla* et *Rhodothermus* sont également bien représentés. Les prédictions de structures secondaires (PSIPRED, PROFSEC, SAM-t99-2D) proposent une structure principalement β de six à sept brins. A ce jour, ce domaine n'est toujours pas répertorié dans Interpro et aucune fonction ne lui est assignée. Cependant il faut ici souligner la présence de ce domaine dans de nombreux enzymes du catabolisme des carbohydrates et des polypeptides. Ainsi on le retrouve dans la κ -carraghénase (GH16) et la β -agaraseA (GH16) de *Zobellia*, dans quatre β -agarases (GH16, 50 et 86) de *Microscilla*, deux fucoïdane lyases de *Flavobacterium* et diverses xylanases et endoglucanase (Fig. 41). Ce domaine ne montre aucune similarité avec les 42 familles de CBM répertoriées dans la banque de séquence du serveur CAZy.

Etant donnée la position terminale de ce domaine, il assume vraisemblablement un rôle d'adressage, soit comme signal de transport, soit par ancrage à une protéine ou à un carbohydrate par exemple pariétal. D'après la taille de la protéine sauvage excrétée par SW5 (estimée à 105 kDa par SDS-PAGE, et à 143 kDa par SEC), ce domaine n'est pas clivé. L'exclusivité du phylum bactérien Bacteroidetes pour ce domaine montre que sa fonction est spécifique à ce taxon, ce qui à mon avis va dans le sens de la deuxième hypothèse, si l'on considère que les machineries d'excrétion des espèces Gram-négatives seraient plus conservées qu'un système de localisation par adhésion. D'autant que ce domaine est sous-représenté dans les génomes de *Bacteroides thetaiotaomicron* et *B. fragilis*, deux bactéries des Bacteroidetes du tractus intestinal. Mais ces considérations sont évidemment à envisager avec la plus extrême prudence et méritent d'être approfondies.

```

.....10
unkn_Cytophaga hutchinsonii.
unkn_Porphyrromonas gingivalis.
unkn_Porphyrromonas gingivalis.
unkn_Porphyrromonas gingival
antigen_Porphyrromonas gingival
antigen_Porphyrromonas gingival
antigen_Porphyrromonas gingival
antigen_Porphyrromonas gingiva
collagenase_Cytophaga.
A-tubulin-suppres_Cytophaga hu
protease_Cytophaga hutchinsoni
protease_Cytophaga hutchinsoni
Rtx-1toxins_Cytophaga hutchinso
xylanase_Rhodothermus marinus.
mannosidase_Rhodothermus marin
Glc-dehydrogenases_Cytophaga h
Endoglucanase_Cytophaga hutchi
Endoglucanase_Cytophaga hutchi
xylanase_Cytophaga hutchinsoni
Glycosyltransferases_Cytophaga
B-agarase_Microscilla.
B-agarase_Microscilla.
B-agarase_Microscilla.
B-agarase_Microscilla.
B-agarase_Microscilla.
B-agarase_Zobellia galactanti
B-agarase_Zobellia galact
Fucooidan lyase2_Flavobacterium
Fucooidan lyase1_Flavobacterium
FcnA_Cter.
.....20
.....30
.....40
.....50
.....60
.....70
.....80
.....90
.....100
.....110
.....120
.....130
.....140
.....150
.....160
.....170
.....180
.....190
.....200
.....210
.....220
.....230
.....240
.....250
.....260
.....270
.....280
.....290
.....300
.....310
.....320
.....330
.....340
.....350
.....360
.....370
.....380
.....390
.....400
.....410
.....420
.....430
.....440
.....450
.....460
.....470
.....480
.....490
.....500
.....510
.....520
.....530
.....540
.....550
.....560
.....570
.....580
.....590
.....600
.....610
.....620
.....630
.....640
.....650
.....660
.....670
.....680
.....690
.....700
.....710
.....720
.....730
.....740
.....750
.....760
.....770
.....780
.....790
.....800
.....810
.....820
.....830
.....840
.....850
.....860
.....870
.....880
.....890
.....900
.....910
.....920
.....930
.....940
.....950
.....960
.....970
.....980
.....990
.....1000

```

Figure 41 : Evidance de la conservation d'un domaine C-terminal chez FcnA est diverses protéines issues de bactéries du phylum des Bacteroidetes. Alignement multiple des domaines homologues sélectionnés à partir des meilleurs hits de GENBANKnr produits par BLAST (25 séquences suivantes) avec la requête F₉₂₀-K₁₀₀₇ (FcnA). La figure est réalisée à l'aide du programme libre BioEdit ; les conservations (colorées sur la base de leurs propriétés physicochimiques) et identités (encadrées) sont mises en évidence au delà de 60 % (conservations calculées en employant la matrice BLOSUM62).

La κ -carraghénase et la β -agaraseA (GH16) de *Zobellia galactanivorans* sont également purifiées du surnageant de culture. La taille de leur séquence est voisine mais elles sont purifiées tronquées respectivement de 229 et 245 résidus en C-terminale. Le tronquage a lieu dans une charnière suivant le domaine catalytique et la partie manquante contient donc le peptide C-terminal commun, avec en plus intercalé, un module de fonction inconnue pour la β -agaraseA (environ 130 résidus) et un CBM16 pour la κ -carraghénase (environ 140 résidus). La première explication fournie pour cette amputation, l'attribuait à un mécanisme d'excrétion (Barbeyron et al., 1998 ; Jam et al., 2004). L'intégrité présumée de FcnA suggère que ce n'est pas le cas, d'autant que les parties tronquées représentent près de 50 % de ces protéines et contiennent des modules fonctionnels. Ceci vient appuyer l'hypothèse de l'ancrage périphérique. On sait à travers plusieurs exemples que la membrane externe est un compartiment cellulaire très favorable pour disposer des activités de dégradation. Cette localisation des enzymes permet d'assurer le contact cellule/substrat et de faciliter l'incorporation des produits intermédiaires de dégradation. Si on pense immédiatement au cellulosome avec le binôme dockerine/cohésine (Bayer et al., 2004), le chitinosome de *Microbulbifer degradans* (γ -protéobactérie) illustre aussi ce concept (Howard et al., 2003, 2004). De plus, la κ -carraghénase (GH16) de *P. carrageenovora* (γ -protéobactérie) possède également un domaine C-terminal de taille similaire (repliement Ig-like ; ref. Interpro IPR003343) localisant très vraisemblablement cet enzyme à la surface de la cellule via une interaction protéine-protéine. Cette protéine a également été purifiée sous forme tronquée après le domaine catalytique (Potin et al., 1995).

La récurrence de ce clivage au même endroit dans trois glycosidases extracellulaires de la famille GH16 est troublante. On est en droit de se demander si cette protéolyse est subie (charnière interdomaine trop exposée) ou si elle est organisée. Les régions clivées contiennent des modules fonctionnels. De plus, l'une d'entre elle ne présente avec les deux autres aucune similitude qui puisse servir de signal. En outre, selon toute vraisemblance, la fucoïdanase FcnA n'est pas amputée de 200 résidus en C-terminal. Il est donc légitime de penser que cette protéolyse est subie et qu'elle est liée à une sensibilité aux protéases de la charnière suivant le domaine

catalytique de ces κ -carraghénases (fragilité non partagée par FcnA). Si l'on accepte l'idée d'un ancrage de surface il est possible qu'une forte expression de l'enzyme (lors d'induction au substrat) s'accompagne, dans les conditions expérimentales, d'une libération d'enzymes dans le milieu extracellulaire, expliquant les fortes quantités d'enzymes libres et la fragilité de ces dernières vis à vis des protéases.

2.2.3. Le domaine catalytique N-terminal

La partie amont aux trois répétitions est longue de 390 résidus, peptide signal exclu. La seule séquence ressortant de l'analyse BLAST avec les paramètres de calculs les plus sensibles dans GENBANKnr est une α -amylase (GH13) de *Thermoactinomyces vulgaris* (Q60053, α -amylase I). Malgré un score non significatif, cet alignement correspondant à la région 160-320 de la fucoïdanase est cependant intéressant. La région concernée de l' α -amylase porte en effet les résidus impliqués dans la catalyse et dans la stabilisation du substrat. Affiné par plotHCA (Fig. 42), cet alignement révèle une conservation de l'un d'entre eux (un acide aspartique) ainsi que de son environnement. Concernant le second résidu catalytique (un acide glutamique) les correspondances sont moins nettes, et le seul candidat envisageable chez la fucoïdanase est un aspartate. Avec plus de sensibilité, la recherche de similitudes par WU-Blast2 (serveur de l'EBI) sur la banque UNIPROT apporte de nombreux hits renforçant cette hypothèse d'un lien évolutif avec la famille 13. A supposer que ces ressemblances reposent bien sur une réalité structurale, ces deux enzymes n'appartiennent donc pas à la même famille de glycoside hydrolase. Les α -amylases ont un domaine catalytique de type TIM barrel (tonneau $[\beta/\alpha]_8$). Ces domaines correspondent en taille (environ 400 résidus) à cette région de la fucoïdanase. Enfin, curieusement, sur les six cystéines de FcnA, cinq se situent dans ce segment.

Les résultats de BLAST sur la banque de séquences brevetées PAT (serveur du NCBI) ont révélé l'existence de deux protéines homologues sur l'ensemble de cette région N-terminale. Ces deux enzymes, Fda1 et Fda2 (respectivement 814 et 881 résidus), sont deux fucoïdane hydrolases isolées de la souche *Alteromonas sp.* SN-1009 (γ -protéobactérie ; Takayama et al., 2002). Leurs gènes respectifs sont contigus, ce qui,

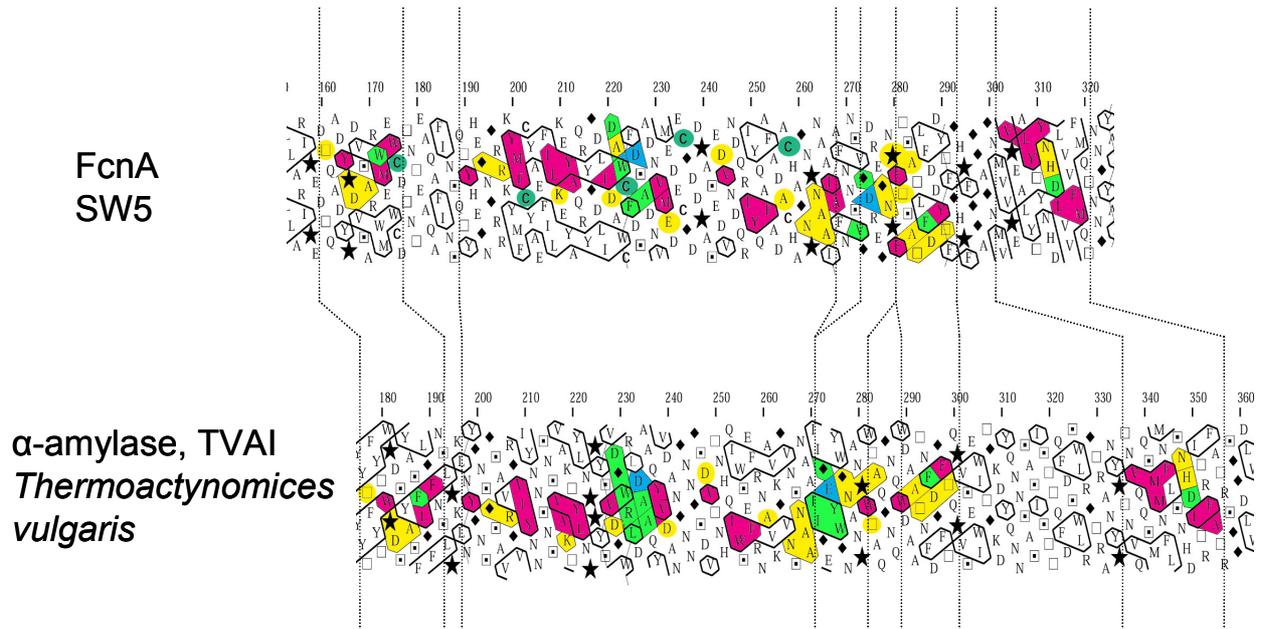


Figure 42 : Alignement des plots HCA de FcnA et d'une α -amylase (TVAI, *Thermoactinomyces vulgaris*, Q60053). Les acides aminés catalytiques de l'amylase et les candidats à correspondance de la fucanase sont colorés en bleu clair ; en vert sont colorés les résidus consensus des α -amylases et leurs vis à vis chez FcnA ; le rouge et le jaune illustrant les conservations respectivement hydrophobes et hydrophiles.

compte tenu de leur forte similarité (70 % d'identité de séquence), suggère un phénomène de duplication récent. La région homologue à FcnA partage avec ces dernières respectivement 20 et 18 % d'identité sur approximativement 380 résidus, elle est également située en N-terminale. De tel niveau d'identité se situe dans la « twilight zone », mais ici la conservation claire de plusieurs motifs permet de bien asseoir l'alignement et de montrer le lien de parenté évolutive entre ces trois séquences (Fig. 43). J'ai réalisé un profil PSSM à partir de ces trois séquences de façon à utiliser PSI-BLAST contre GENBANKnr et rendre les recherches d'homologies plus sensibles, mais aucun résultat supplémentaire ne s'en est dégagé (de même avec un profil HMM via HMMER). Ce domaine est vraisemblablement responsable de l'activité fucoïdanase. C'est la première fois qu'il est précisément borné puisque les seuls homologues connus Fda1 et Fda2 sont modulaires et quasi jumelles, empêchant tout recoupement par intersection. Je propose donc la création d'une nouvelle famille de glycoside hydrolase.

Fda2 possède en amont une région de basse complexité (motif PEPEVE répété environ neuf fois) de 62 résidus qui s'apparente à un bras espaceur. Fda2 se voit prédire un peptide signal amont de 25 résidus par SignaP v3, mais Psort reconnaît la présence des consensus type lipoprotéine, curieusement le cas est analogue à celui de la λ -carraghénase (cf. Chapitre 2) issue d'une γ -protéobactérie également. Fda1 a un peptide signal de 22 résidus prédit par les deux serveurs. En aval de la région d'intérêt, Fda1 et Fda2 possèdent deux domaines Laminin-G. Ces derniers sont impliqués chez les métazoaires dans l'adhésion à l'héparine et aux α -dystroglycanes. Ces domaines sont agencés en tandem dans les laminines et on retrouve des homologues dans certaines des larges protéines bactériennes contenant des domaines cadhérine-like (cf. ci dessus ; Fig. 48). Chez Fda1 et Fda2, il est probable qu'ils servent de module d'adhésion aux fucoïdanes sulfatés.

Sachant que la catalyse des glycosidases est assurée par un couple de résidus acides, j'ai recherché ce d'entre eux conservés dans l'alignement réalisé. Onze positions sont ainsi ressorties ; mais seulement quatre d'entre elles appartiennent à des motifs bien conservés (D221, D226, D275 et D289 d'après la numérotation de FcnA entière). Afin d'affiner cette recherche, j'ai tenté de reconnaître les structures

secondaires et le repliement de ce domaine. De telles prédictions gagnent en vraisemblance lorsqu'elles se basent sur un alignement multiple. Cependant le nombre restreint de séquences homologues disponibles ajouté à l'incertitude de certaines régions de l'alignement (faible identité) n'a pas permis d'employer une telle stratégie. Ces investigations sont néanmoins consolidées ici par la comparaison des résultats issus de trois programmes performants (PSIPRED, PROFSEC, et SAM-t99-2D) appliqués séparément à ces trois séquences homologues. Les résultats ont même permis d'ajuster certaines parties floues de l'alignement des structures primaires. Les structures secondaires annoncées par les programmes sont partiellement convergentes (Fig. 43). Néanmoins elles laissent apparaître de multiples blocs composés d'une large hélice α suivi après une courte boucle par un brin β . Ces blocs sont mieux définis dans les deux premiers tiers de l'alignement. Les prévisions de repliement sont récoltées sur le métaserveur BIOINFO.PL. La hiérarchisation des modèles fournie par le programme 3D-JURY (avec les paramètres par défaut) montre que les meilleures prédictions sont obtenues par les programmes de threading ce qui reflète simplement l'absence d'homologie de séquence suffisante avec les protéines de la PDB. En outre, pour les trois protéines FcnA, Fda1, et Fda2, les meilleurs scores sont obtenus par les modélisations se reposant sur des TIM barrels de glycosidases (ref. SCOP c.1.8). Malgré la forte divergence de séquence on observe donc une cohérence des résultats. Les meilleurs modèles obtiennent tous des scores supérieurs au seuil de confiance (meilleur score de 92 avec un seuil de 50). La structure 1gqi de la α -glucuronidase (GH 67) est classée deux fois première et une fois troisième. Les alignements de structures secondaires fournis par le serveur calent les blocs α/β évoqués ci-dessus sur ceux des TIM barrels sélectionnés. Ils sont beaucoup moins convaincants dans le dernier tiers du domaine.

Les tonneaux $(\beta/\alpha)_8$ canoniques sont des structures en solénoïde circulaire où 8 brins β parallèles constituent l'intérieur du tonneau alors que les 8 hélices α se placent à l'extérieur. Leur séquence comprend donc huit répétitions β - α (numérotées de 1 à 8), puisque la séquence débute par un brin β (Fig. 44A). Ce sont des structures très représentées au sein des glycosidases. Ce repliement concerne en effet la moitié de leur séquences, un tiers des familles et deux tiers des réactions catalysées au sein de la clas-

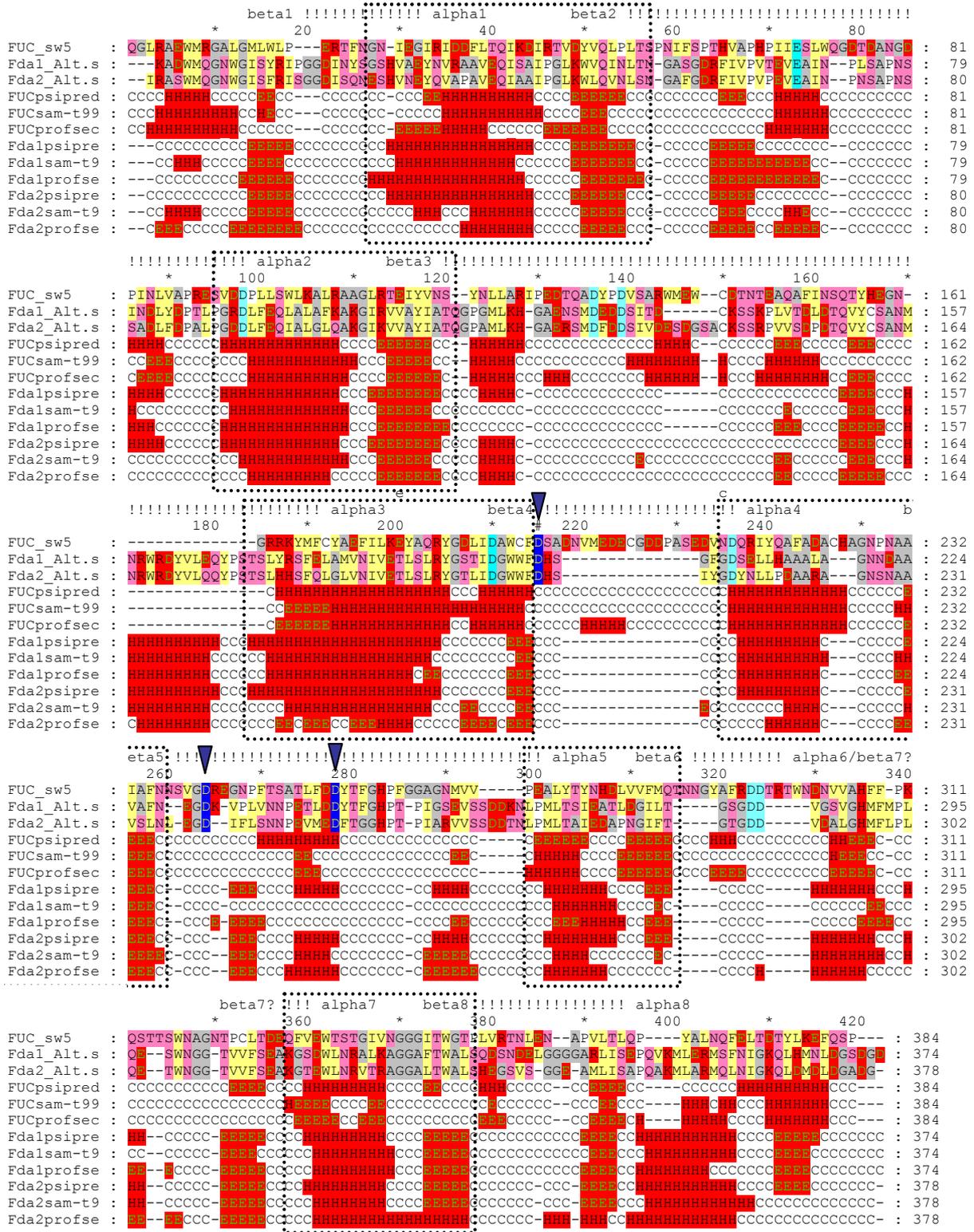


Figure 43 : Alignement des domaines catalytiques des trois fucoïdane hydrolases. Les résidus sont colorés selon leurs propriétés physico-chimiques (petit/gris, polaire/rose, chargé/rouge, et hydrophobe ou aromatique/jaune). Le bleu clair désigne les résidus acides conservés, alors que le bleu foncé indique parmi ces derniers les candidats catalytiques (flèches bleues sur D226, D275 et D289). Les structures secondaires de chacune des trois séquences FcnA (29-399), Fda1 (25-398), Fda2 (87-464) ont été prédites par Psipred, Samt99-2D et ProfSec. Ces prédictions sont alignées sous les séquences avec H, E, C signifiant respectivement, hélice α , brin β , et boucle. Les rectangles désignent les blocs $[\alpha_i\text{-}l_i\text{-}\beta_{i+1}]$ pouvant participer à un TIM barrel hypothétique. Ils séparent les boucles L_i pouvant participer au site actif.

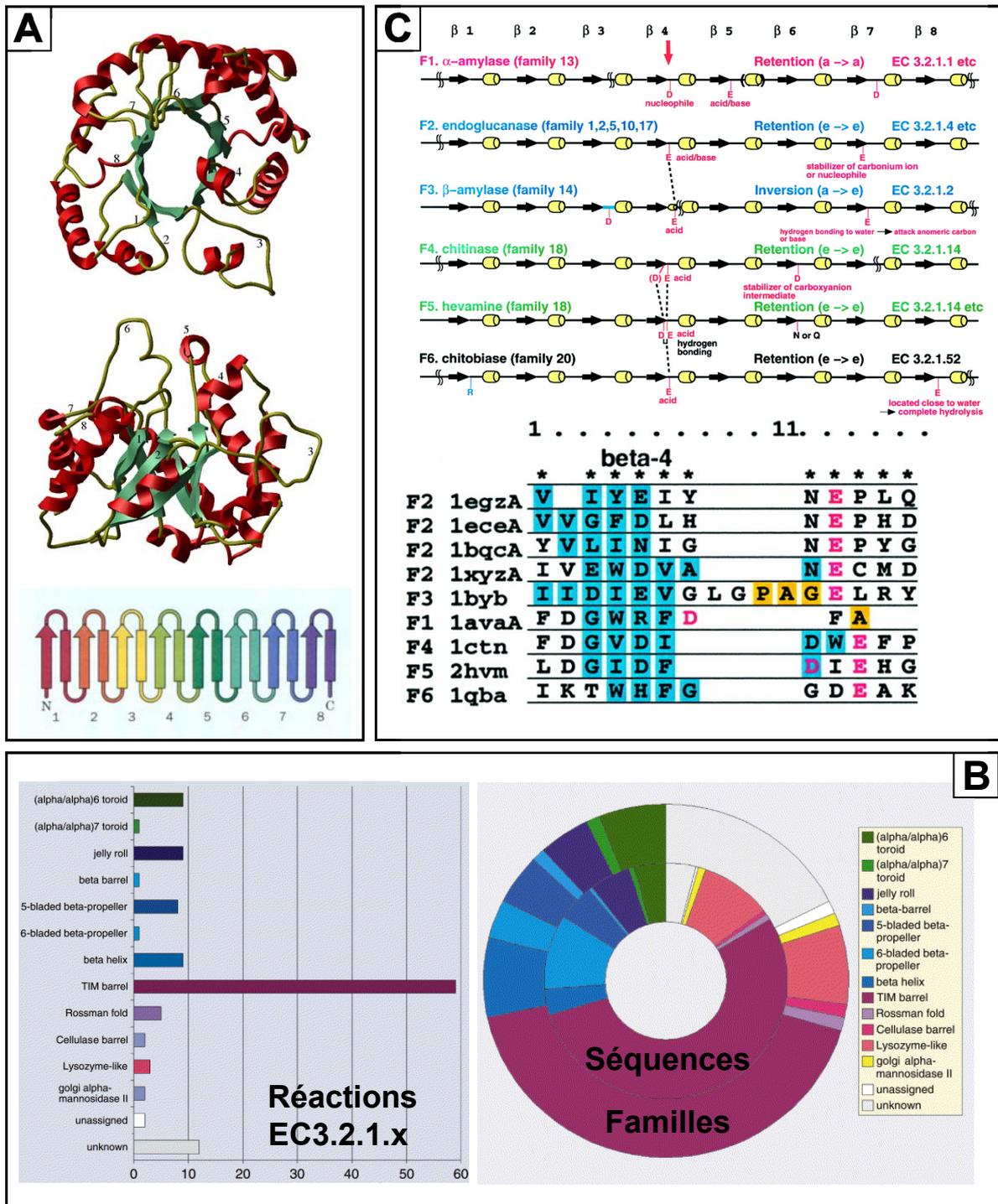


Figure 44 : Récapitulatif des caractéristiques des tonneaux (β/α)₈, ou TIM barrel, chez les glycosidases. A : un exemple de TIM barrel permettant de reconnaître les 8 répétitions β/α propre à ce tonneau, et la face catalytique C-terminale de ce dernier (extrait de Wierenga, 2001). B : contribution de ce repliement chez les glycosidases parmi leurs séquences, leurs familles, et leurs réactions EC3.2.1.x (d'après Rigden et al., 2003) comparée à celle des autres repliements identifiés. C : Répartition selon des critères structuraux des familles GH1, 2, 5, 10, 13, 14, 17, 18 et 20 en six classes F1 à 6, elles-mêmes regroupées en 4 groupes (code couleur) apparemment indépendants. Leur point commun est de tous conserver un résidu catalytique (en rouge dans l'alignement) à l'extrémité C-terminale du brin β_4 (en bleu). La position du deuxième résidu catalytique est variable (d'après Nagano et al. (2001).

se EC3.2.1.x (Rigden et al., 2003 ; Fig. 44B). La contribution exceptionnelle de ce repliement à cette classe d'enzymes soulève évidemment le problème de leur évolution convergente ou divergente qui a suscité plusieurs études comparatives (Nagano et al., 2001, 2002). Chez les glycosidases, ce repliement possède majoritairement huit répétitions (excepté la famille GH6 qui en a sept) et affiche une taille moyenne de 380 résidus (variant entre 270 et 500). Mais le point le plus frappant est la présence quasiment systématique dans les glycosidases d'un résidu catalytique proche de l'extrémité C-terminale du brin β_4 (Fig. 44C). C'est en effet toujours la face C-terminale du tonneau qui porte le site actif, ce dernier étant formé par les boucles (L_i) reliant les brins β (β_i) à l'hélice α (α_i) qui suit (Fig. 44A). Ces boucles peuvent être assez larges et porter des éléments de structures secondaires voire même des sous-domaines. A l'inverse, les boucles N-terminales (L_i entre α_i et β_{i+1}) sont généralement courtes et participeraient à la stabilisation du repliement. L'étude de ces dernières montre des préférences conformationnelles distinctes selon qu'elles appartiennent à une répétition numérotée paire ou impaire, illustrant la théorie de l'évolution des TIM barrels par la duplication d'entités $(\beta/\alpha)_2$. Ceci se traduit sur la structure primaire par la succession de sept blocs α_i - β_{i+1} plus ou moins espacés.

En termes de probabilité, étant donné la représentation de ce repliement chez les glycosidases, toute nouvelle famille a de forte chance de l'adopter. A partir de l'alignement des prédictions de structures secondaires, ainsi que de l'alignement aux structures sélectionnées pour la modélisation 3D, je propose sur la figure 43 cinq blocs α_i - β_{i+1} pouvant satisfaire à la formation d'un TIM barrel. Le manque de convergence des prédictions sur le dernier tiers de l'alignement rend cette identification très difficile et pourrait expliquer les deux blocs manquants. De plus, la structure d'une α -L-fucosidase (*T. maritima* ; 1hl8 ; Sulzenbacher et al., 2004) récemment résolue, montre l'absence des hélices α_5 et α_6 . Les deux premiers tiers de l'alignement paraissent plus consistants et les résidus conservés D226, D275 et D289 se place dans les boucles putatives L_4 et L_5 (Fig. 43). D221 appartenant à un motif très conservé se place dans la boucle hypothétique L_3 ce qui empêcherait sa participation à la catalyse. La position de D226 satisfait la conservation d'un résidu catalytique en L_4 chez les glycosidases adoptant ce repliement. La position putative d'un second résidu catalytique dans L_5

n'est observé que dans la famille GH13 et le clan H en général où il s'agit d'un glutamate. Cette observation rappelle le résultat de BLAST concernant l' α -amylase TVAI de *Thermoactinomyces vulgaris* (Fig. 42).

Il existe donc un faisceau de présomptions convergeant vers l'hypothèse d'un tonneau $(\beta/\alpha)_8$ qui désigne trois résidus candidats pour la catalyse. Afin de résumer l'ensemble de cette analyse, la modularité de FcnA est schématisée dans la figure 49.

2.3. Analyse des autres séquences protéiques du contig

2.3.1. La sulfatase partielle SulB

En amont du contig se trouve un cadre de lecture non clos. Cette séquence partielle de 230 résidus contient 88 résidus de la fin d'un domaine catalytique type arylsulfatase, suivi d'une queue C-terminale de 142 résidus comprenant aucune signature de domaine connu. Les deux protéines homologues les plus proches appartiennent à la bactérie *Rhodopirellula baltica* (groupe des Planctomycetes Fig. 45). Ce sont les deux seules sulfatases de GENBANKnr partageant cette queue C-terminale en plus du domaine catalytique. RB198 (NP_863771, 608 résidus) et RB1736 (NP_864562, 622 résidus) ont respectivement 31 et 30 % d'identité de séquence avec SulB. RB198 est annotée comme étant une N-acétylgalactosamine-4-sulfatase putative. La longueur de ces deux protéines indique qu'il manque probablement environ 400 résidus à la séquence de SulB. Malgré la séquence largement incomplète de sulB, la proximité de cette dernière avec ces deux sulfatases de *R. baltica* est relativement fiable du fait du partage exclusif de cette queue C-terminale.

2.3.2. La carboxylestérase EstA

Suivant SulB, le cadre de lecture de EstA contient 363 résidus. Il y a une bonne probabilité de peptide signal (SignalP v3) avec une coupure soit après le résidu 19 soit après le 21 qui est appuyée par l'alignement avec les protéines homologues. La séquence homologue la plus proche dans GENBANKnr est une protéine de la bactérie

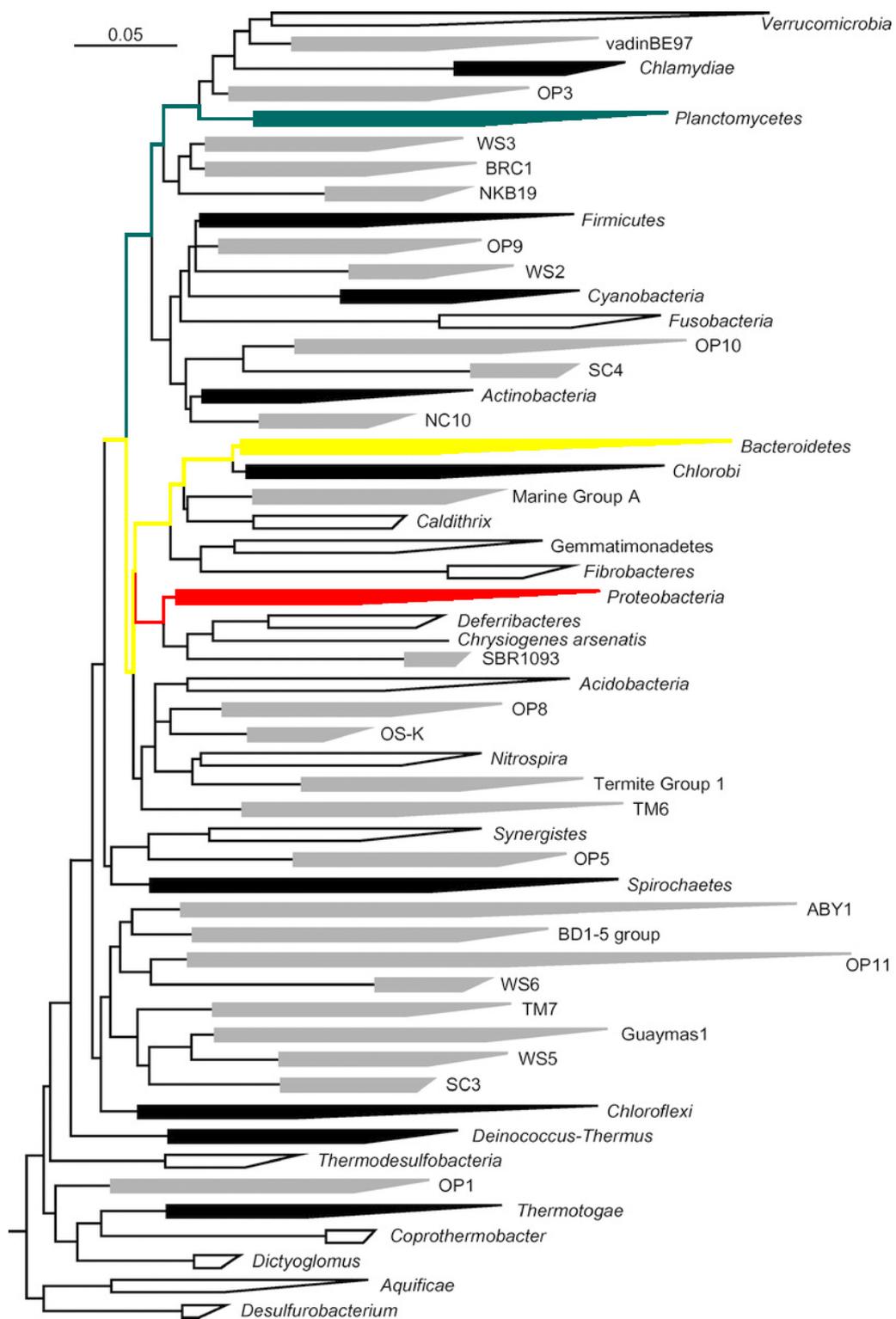


Figure 45 : Arbre phylogénétique illustrant les phyla majeurs du groupe des Bactéries (modifié à partir de Rappé et Giovannoni, 2003). Il est basé sur l'alignement de 12000 gènes pleine longueur codant pour l'ARN16S des espèces non cultivables (gris) et cultivables (noir). L'échelle correspond à 0,05 mutations par nucléotides. Il met en évidence l'ancienneté de la divergence des phylums Planctomycetes, Bacteroidetes, et Protéobactéries. Le code couleur est conservé entre les figures 45, 46, 47 et 48).

R. baltica là encore. RB13156 (NP_870897, 393 résidus) est une homologue pleine longueur, partageant 48 % d'identité de séquence avec EstA ; elle est annotée lipase/estérase putative. Parmi les dix hits suivants la moitié provient de bactéries Firmicutes (Gram-positives), alors que trois proviennent de *R. baltica* et une de *Microbulbifer degradans* (γ -Protéobactéries ; ZP_00317670). Les 20 hits suivant (BLAST) sont également des séquences bactériennes qui présentent de 27 à 24% d'identité avec EstA. Le premier hit se détache donc largement des suivants. EstA et RB13156 partage très vraisemblablement le même substrat (identité > 45 %). Dans le but de récapituler ces observations, j'ai réalisé un arbre phylogénétique (Fig. 46) comprenant EstA, les meilleurs hits identifiés par BLAST dans la PDB, SwissProt et parmi les séquences bactériennes de GENBANKnr. L'arbre obtenu discrimine fortement les séquences eucaryotes (groupe I et II) des procaryotes. Cependant parmi les séquences procaryotes, il est clair que l'arbre obtenu ne satisfait l'arbre phylogénétique des espèces bactériennes représentées (Fig. 45). Ainsi, les nombreuses estérases de *R. baltica* ne se regroupent pas spécialement. Enfin, EstA et RB13156 forment un groupe très isolé parmi les autres homologues bactériens. Et d'un point de vue fonctionnel cet arbre est peu informatif étant donné le groupement des protéines expertisées.

Interproscan reconnaît un motif EF-hand (IPR002048) isolé en N-terminal (60 résidus) débutant à la position 25 et se terminant à la 75. Ce domaine composé de deux larges hélices α (12 résidus chacune) reliées par une boucle de 12 résidus servant à chélater un ion calcium. Généralement il est souvent répété mais ce n'est pas le cas ici. C'est à sa suite que commence le domaine catalytique de type estérase (IPR000379, ref SCOP C.69.1.14). Il adopte un repliement de type α/β -hydrolases (trois couches $\alpha/\beta/\alpha$) comprenant un feuillet principalement parallèle de huit brins. La catalyse est assurée par une triade S, H et D.

La meilleure modélisation par homologie désignée par 3D-Jury est basée sur un alignement avec la carboxylestérase d'archéobactérie 1jji (*Archaeoglobus fulgidus*). Elle ne présente en effet seulement quelques petites délétions (< dix résidus) positionnées dans des boucles. D'après les alignements obtenus, je propose la triade catalytique suivante : S₂₀₂, H₃₃₆, et D₃₀₄. SwissProt ne produit pas d'homologue

suffisamment proche pour déterminer la spécificité de substrat de cet enzyme (meilleur hit à 27 % d'identité). Par ailleurs, EstA ne présente aucune homologie de séquence détectable avec les 14 familles de carbohydrates estérases répertoriées dans le serveur CAZy.

2.3.3. La sulfatase entière Sula

Après le cadre de lecture de EstA se trouve celui de FcnA qui précède à son tour celui de Sula. Il code pour une protéine de 486 résidus. La recherche d'un peptide signal (SignalP v3) montre la très forte probabilité de son existence avec une coupure entre les résidus 22 et 23 (confirmation par alignement avec les séquences homologues). Interproscan reconnaît un domaine catalytique type sulfatase (IPR000917). L'analyse par BLAST révèle l'existence de nombreux homologues pleine longueur dans GENBANKnr. Le plus proche, en partageant 42 % d'identité de séquence (cinq indels inférieurs à cinq résidus), provient encore de *R. baltica*. Il s'agit de RB1456 (NP_864421, 495 résidus) annotée comme étant une N-acetylgalactosamine-6-sulfate sulfatase putative. Le dix meilleurs scores sont tous issus de *R. baltica* à l'exception du deuxième provenant de *Bacteroides thetaiotaomicron* (BT3486, 453 résidus, annotée N-acetylgalactosamine-6-sulfatase putative, NP_812398). Ils présentent entre 33 et 28 % d'identité soulignant le fossé qui existe entre le premier hit et les suivants. La majorité des hits sont issus des programmes de génomiques et les annotations sont donc peu fiables. Toutefois la majeure partie d'entre elles est relative à la désulfatation de polysaccharides type glycosaminoglycanes (GAGs) sulfatés (dermatane, héparane, chondroïtine sulfatés). Le plus proche homologue dans SwissProt est la N-acetylgalactosamine-6-sulfatase humaine (P34059, 522 résidus) qui partage 28 % d'identité avec Sula malgré l'existence de larges indels dans l'alignement. La spécificité de substrat n'est donc pas prédictible même si il apparaît clairement que les meilleurs hits disponibles dans SwissProt sont également reliés aux métabolismes de polysaccharides sulfatés. Sula et RB1456 partage très vraisemblablement la même spécificité de substrat et si il s'agit d'un polysaccharide sulfaté rien n'indique que ce soit nécessairement un GAG. En effet

Figure 46 : Arbre phylogénétique des homologues à la protéine EstA. L'arbre est réalisé à partir des meilleurs homologues à EstA issues de la PDB (●), de SwissProt (■), ainsi que d'une sélection des meilleurs hits de GENBANKnr (bactéries) identifiés par BLASTp (▲). L'arbre est construit selon la méthode de neighbor-joining en employant la matrice JTT, en acceptant les gaps. Il est testé par 100 pas de Bootstrap (valeur indiquée sur les branches). L'alignement utilisé contient 203 positions alignables. La couleur jaune se réfère aux Bacteroidetes, le bleu-gris aux Planctomycetes, le rouge aux Protéobactéries, le rouge plein aux γ -Protéobactéries, Le bleu clair aux Firmicutes, le verts aux Cyanobactéries, le bleu foncé aux Actinobactéries, le rose aux Thermotogales et le noir aux eucaryotes. (réalisé avec le programme MEGA v3)

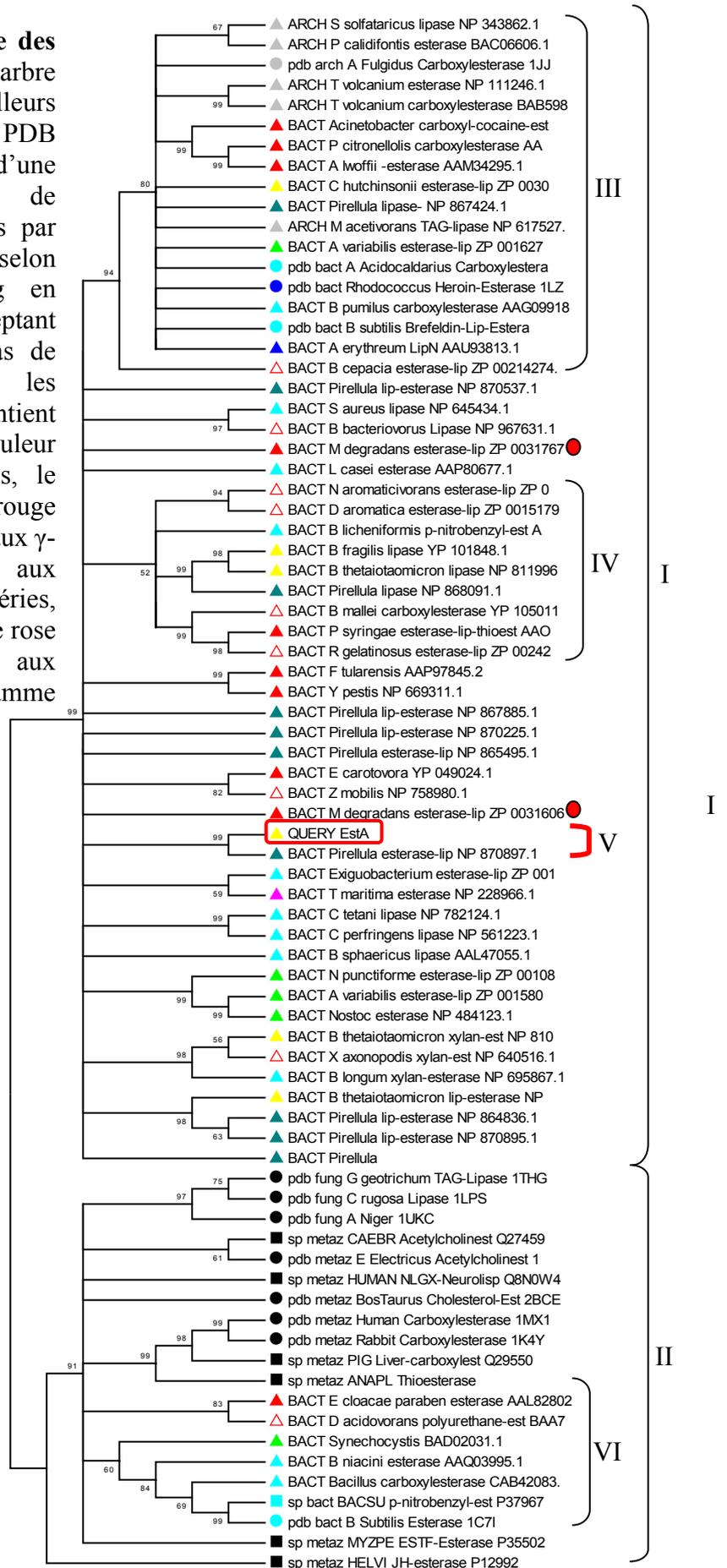
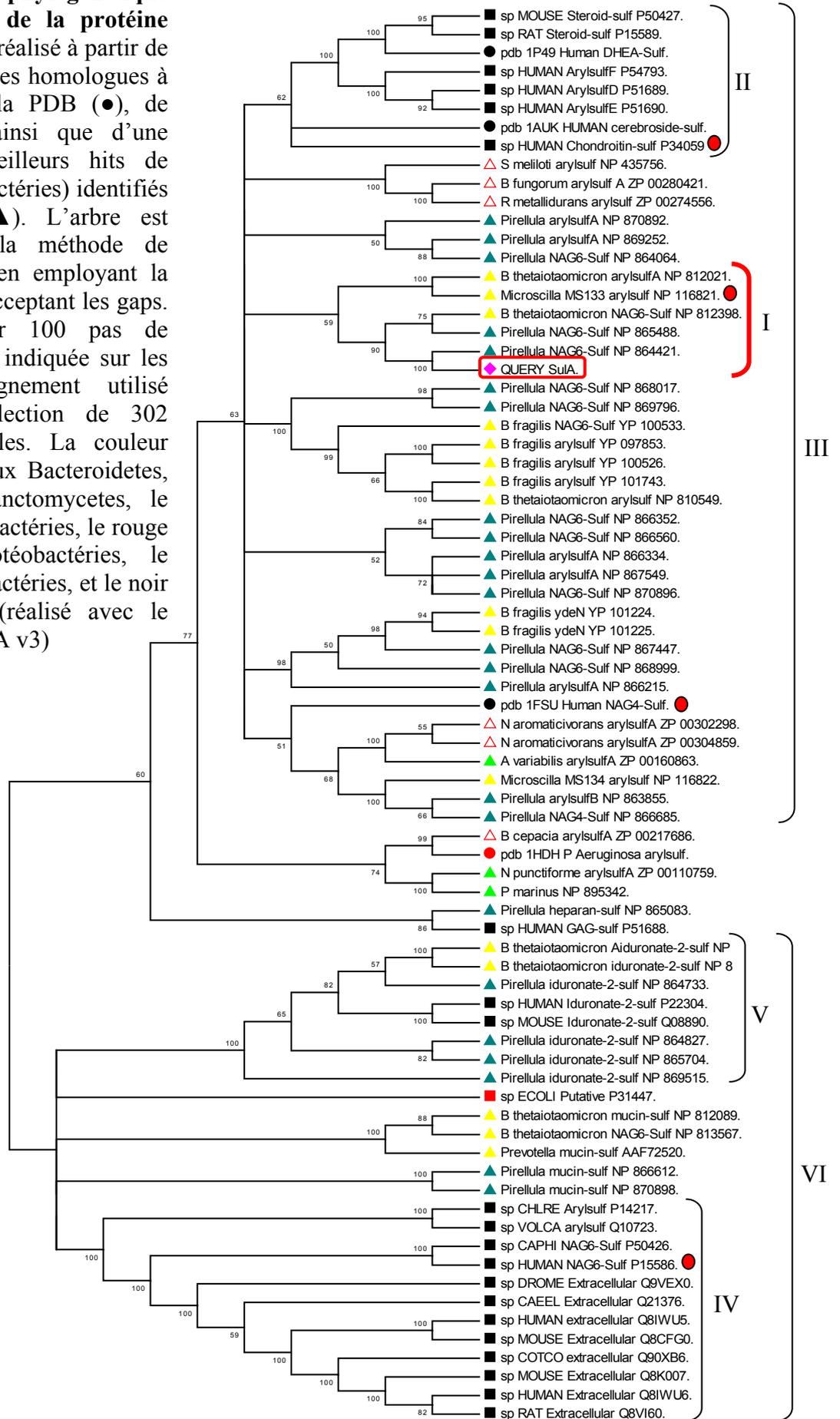


Figure 47: Arbre phylogénétique des homologues de la protéine Sula. L'arbre est réalisé à partir de toutes les séquences homologues à Sula issues de la PDB (●), de SwissProt (■), ainsi que d'une sélection des meilleurs hits de GENBANKnr (bactéries) identifiés par BLASTp (▲). L'arbre est construit selon la méthode de neighbor-joining en employant la matrice JTT, en acceptant les gaps. Il est testé par 100 pas de Bootstrap (valeur indiquée sur les branches). L'alignement utilisé contient une sélection de 302 positions alignables. La couleur jaune se réfère aux Bacteroidetes, le bleu aux Planctomycetes, le rouge aux Protéobactéries, le rouge plein aux γ -Protéobactéries, le verts aux Cyanobactéries, et le noir aux eucaryotes. (réalisé avec le programme MEGA v3)



malgré son annotation, RB1456 ne présente que 30% d'identité avec N-acetylgalactosamine-6-sulfatase (NAG6-sulfatase) humaine (P34059, 522 résidus) et avec de plus l'existence de nombreux indels dans l'alignement.

Avec l'espoir de raffiner ces comparaisons de séquences, j'ai réalisé une analyse phylogénétique (Fig. 47) de Sula et des sulfatases contenues dans la PDB et SwissProt. J'ai également incorporé à l'alignement les meilleurs hits identifiés par BLAST parmi les séquences bactériennes. L'arbre obtenu discrimine fortement les séquences eucaryotes (groupe II et IV) des procaryotes. Cependant parmi les séquences procaryotes, l'arbre ne satisfait pas l'arbre phylogénétique des espèces bactériennes représentées. La fonctionnalité de ces enzymes étant élucidées presque uniquement chez les eucaryotes (entrées SwissProt et PDB) cette séparation nuit à la prédiction de fonction. Il apparaît cependant que l'annotation « NAG6-sulfatase » portée par de nombreuses protéines de *R. baltica* et *B. thetaiotaomicron* semble abusive dès lors que les annotations sont basées sur la ressemblance aux protéines humaines émergeant dans des groupes distincts (P34059, P15586). On peut néanmoins avancer que la fonction « désulfatation des polysaccharides » est dispersée dans les deux grands groupes émergeant III et VI (mucin, iduronate, N-acétylgalactosamine, cérébroside). Enfin, Sula forme un cluster relativement isolé avec les meilleurs hits trouvés par BLAST (I). Curieusement ce groupe semble lié à une sulfatase (MS133) portée par le plasmide de *Microscilla* sp. nécessaire à la dégradation de l'agar.

Pour mémoire, les sulfatases ont un résidu catalytique modifié post-traductionnellement en une formyl-glycine à partir soit d'une cystéine soit d'une sérine. Sula est une sulfatase à cystéine (C₇₉). Les arylsulfatases adoptent un repliement proche de celui de EstA puisqu'il est de type alcaline phosphatase like (ref SCOP C.76.1.2) avec trois couches $\alpha/\beta/\alpha$ mais dont le feuillet principalement parallèle de huit brins est différemment agencé que dans un repliement de type α/β -hydrolases. Les arylsulfatases ont en plus un domaine C-terminal d'environ 75 résidus et constitué d'un feuillet β (quatre brins) décoré d'une à trois hélice α (le domaine C-terminale de Sula semble contenir trois hélices α). La meilleure modélisation par homologie désignée par 3D-Jury est basée sur un alignement avec l'arylsulfatase 1hdh de la bactérie *Pseudomonas aeruginosa*. Il présente quelques indels mineurs (< cinq résidus)

positionnées dans des boucles, à l'exception d'un segment de 50 résidus dans 1hdh qui comprend une hélice α excédentaire. C'est également le cas pour 1e3c (ou 1auk ; cérébroside sulfatase humaine) et 1p49 (stéroïde sulfatase humaine) qui malgré de bons scores contiennent au moins une forte insertion (50 à 100 résidus) lors de l'alignement avec SulA. A l'inverse, l'alignement avec 1fsu (N-acetylgalactosamine-4-sulfatase humaine) présente une protubérance plus courte (33 résidus), et c'est SulA qui contient cette fois une insertion de 27 résidus. Le taux d'identité sur les segments homologues des structures de ces sulfatases tourne autour de 35 % d'identité. C'est la limite basse pour la réalisation d'un modèle structural par homologie qui puisse être exploitable.

2.3.4. La fucosidase FcsA

Le dernier cadre de lecture est partiel et donne accès à 366 résidus. La prédiction de peptide signal est négative. Toutefois l'alignement réalisé avec les protéines homologues semble montrer le contraire. Interproscan reconnaît un domaine catalytique de type α -L-fucosidase (IPR000933), glycosidases de la famille GH29 dont la structure a été récemment résolue (1hl8, *Thermotoga maritima*, Sulzenbacher et al., 2004). Les résultats de BLAST révèlent que parmi les cinq premiers homologues à cette séquence partielle se trouvent 4 séquences bactériennes du groupe CFB. Le meilleur hit provient de PG1884 (NP_905969, 674 résidus), α -L-fucosidase putative de *Porphyromonas gingivalis*, avec 33 % d'identité. FcsA affiche entre 29 et 22% d'identité avec les fucosidases eucaryotes de SwissProt. Ces analyses sont évidemment à prendre avec recul dans la mesure où la séquence de FcsA est partielle. La longueur des protéines homologues à cette séquence indique une partie manquante de 80 résidus au minimum pour clore le domaine catalytique.

Les fucosidases de la famille GH29 adopte un repliement de type tonneau $(\beta/\alpha)_8$ avec une absence des hélices α_5 et α_6 (Sulzenbacher et al., 2004). FcsA montre la conservation du résidu catalytique nucléophile D₂₄₀ à l'extrémité C-terminale du brin β_4 . Par contre le glutamate proposé comme second résidu catalytique ne semble pas conservé. Suite aux alignements réalisés sur cette structure et sur les homologues de

SwissProt, je propose D₂₈₂ pour substituer ce glutamate. Et les fucosidases de SwissProt (toutes eucaryotes), alignent bien un aspartate à cette position. Mais un alignement plus général des protéines rassemblées sous l'étiquette GH29 confirme la dualité de cette position. La cohabitation de ces deux couples de résidus catalytiques au sein d'une même famille est un fait unique à ce jour, mais les fortes homologues de séquence observées satisfont la définition de la famille proposée par Henrissat en 1991 (Henrissat et Bairoch, 1993 ; Henrissat et Davies en 1997).

2.4. L'hypothèse du cluster fonctionnel

Le fucoïdane est un polysaccharide complexe. Sa dégradation par des bactéries à des fins énergétiques nécessite donc de faire appel à plusieurs activités enzymatiques distinctes. On sait le polysaccharide sulfaté et parfois acétylé. En outre sa métabolisation par les bactéries passe vraisemblablement par l'obtention en bout de chaîne de fucose. Sous cet éclairage, les gènes découverts autour de celui de FcnA deviennent très intéressants puisqu'il s'agit justement de sulfatases, d'estérase et de fucosidase (Fig. 48). Néanmoins leur analyse de séquence n'a apporté aucune certitude quant à leur spécificité de substrat. Elle montre en outre l'absence de domaine commun. La démonstration expérimentale de la spécificité de substrat est lourde et risquée puisqu'il s'agit dans un premier temps d'exprimer ces enzymes putatifs puis de chercher presque en aveugle leur substrat biologique. Ce dernier point n'est pas une mince affaire si l'on considère que, par exemple, une activité sulfatase sur un polymère peut être taille et motif dépendant, très spécifique ou au contraire de reconnaissance très souple, et finalement son action peut n'être observable qu'après l'action d'un autre enzyme (une glycosidase par exemple). Donc en préambule de ce type d'investigations expérimentales je souhaiterais développer certains arguments qui vont dans le sens d'une participation de ces différents enzymes putatifs à une même voie catabolique que ce soit dans le cadre d'un opéron ou d'un simple cluster fonctionnel.

2.4.1. Clusters, opérons et transferts horizontaux

La concentration physique de gènes participant à une même voie métabolique sous forme d'opéron, de clusters ou d'éléments mobiles tels que les plasmides a très tôt été sujet de questionnement. Plusieurs modèles tentent d'expliquer l'apparition de la proximité physique de gènes liés à des fonctions voisines ou identique (Lawrence et Roth, 1996).

Le premier (modèle de Natal) propose une origine due à la duplication suivie de divergence. Mais les clusters contiennent des familles protéiques très différentes et ce modèle n'explique pas la persistance du regroupement. Le second modèle (modèle de Fisher) avance que la proximité physique de gènes coadaptés limite la perte de l'un d'entre eux suite à la recombinaison. Mais ce raisonnement s'applique plus aux eucaryotes qu'aux bactéries où la reproduction est essentiellement asexuée. Le troisième (modèle de la corégulation) suggère que le regroupement facilite l'expression et la régulation ce qui fournit un bénéfice adaptatif à l'hôte. C'est le cas de l'opéron où le bénéfice vient de la cotranscription de plusieurs gènes sous la domination d'un couple promoteur/opérateur. Le problème de ce modèle repose sur la coalescence de l'opéron qui doit être rapide pour que le bénéfice se fasse sentir et que sa sélection ait lieu. Or juxtaposer par réarrangement plusieurs gènes les uns derrière les autres colinéairement est très improbable et requiert donc un temps évolutif très important. En outre la corégulation existe sans regroupement physique. Lawrence et Roth (1996) suggèrent que la corégulation dans le cadre d'opéron soit un bénéfice dérivant secondairement de la mise en place d'un cluster (un raffinement supplémentaire en quelque sorte) mais qui conforte la conservation et la maintenance de ce dernier. Et ces auteurs suggèrent en dernier modèle, celui de l'opéron égoïste.

Il est bien établie à l'heure actuel que les transferts de gènes latéraux (TLG) contribuent pour une part significative aux génomes procaryotes (Ochman et al., 2000; Boucher et al., 2003). D'une part, les TLG observés (et donc sélectionnés) concernent essentiellement non pas des fonctions vitales mais l'acquisition de résistances, de virulences ou de voies métaboliques (ex. catabolisme d'un polysaccharide), bref des fonctions dont l'intérêt peut n'être qu'occasionnel. Ces fonctions sont évidemment sélectionnées par l'avantage évolutif qu'elles procurent à la population les possédant,

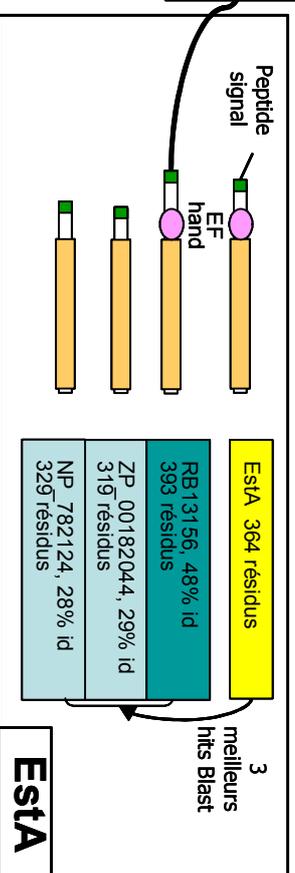
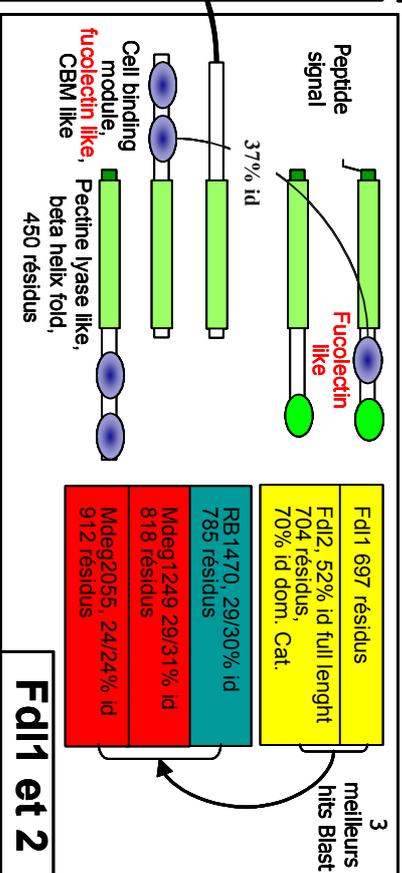
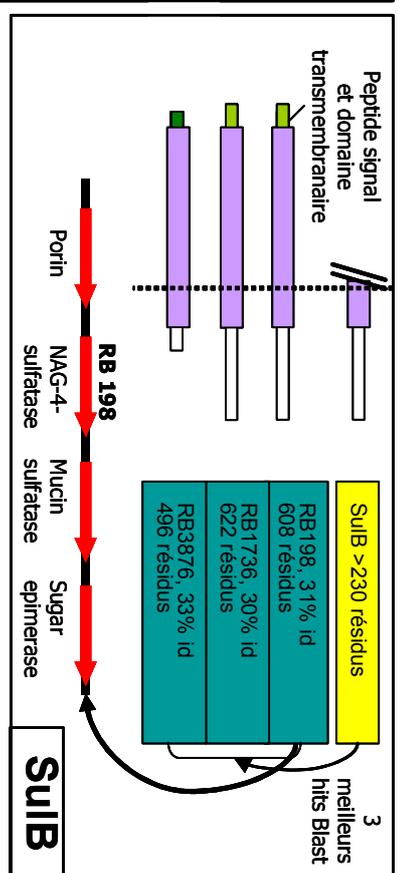
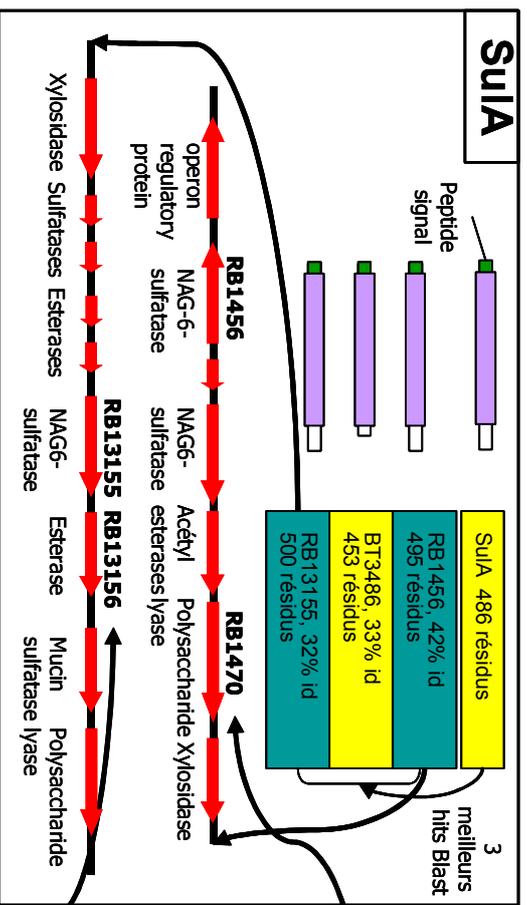
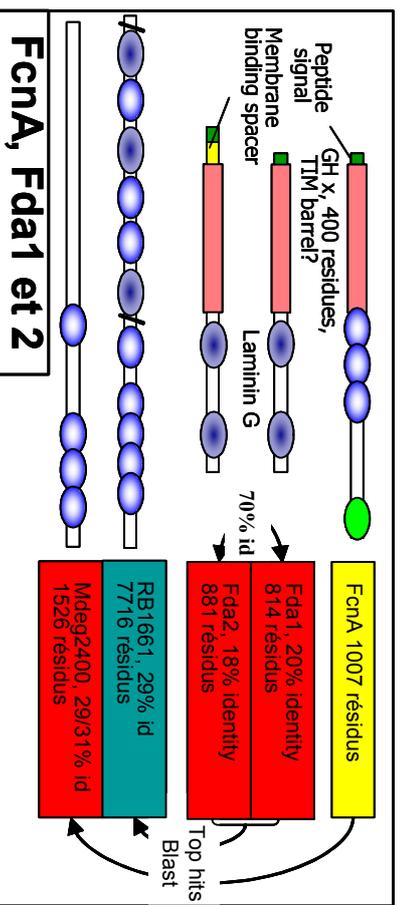


Figure 48 : Récapitulatif des analyses d'homologie (BLAST menées sur GENBANKnr) concernant : les fucoidane hydrolases (Fcna, Fdal et 2), les autres gènes du contig de Fcna (SuiA, SuiB et EstA) et les deux fucoidane lyases (Fdi1 et 2). Les homologues de *R. batlica* sont remplacés dans leur contexte génomique.

dans un contexte écologique donné. La plupart d'entre elles nécessitent l'intervention de plusieurs gènes. D'autre part, un TLG est un événement historique concernant un individu receveur physiquement proche d'un individu émetteur consentant ou non. Or on comprend bien que le maintien du matériel génétique exogène dans la descendance du receveur est dépendant de la pression de sélection qui s'y rapporte. Concrètement, il est par exemple inutile pour un organisme de recevoir une cellulase unique si tout le reste de la machinerie de dégradation et d'importation est absente, auquel cas la dérive génétique éliminera petit à petit ce matériel dans la descendance du receveur initial. La probabilité de transferts latéraux efficaces n'étant pas très élevée (taux de transfert estimé chez *E. coli* à 16 kb/million d'années, Ochman et al., 2000), le transfert groupé de tous les gènes liés à une fonction paraît donc indispensable à la sélection positive du matériel transmis. On explique ainsi la corrélation qui apparaît entre le matériel génétique exogène détecté dans les génomes, et l'organisation en cluster fonctionnel des gènes qui les composent. Le regroupement physique de gènes dans un cluster n'apporte alors pas de bénéfice direct au receveur, il ne favorise que la survie génétique du groupe.

Par ailleurs, les fonctions biologiques portées par les TLG sélectionnés peuvent être soumises à une pression de sélection irrégulière dépendant de la fréquence de leur sollicitation par l'environnement de l'hôte. On peut donc concevoir que les gènes concernés, malgré et à cause d'une utilité épisodique, aient du mal à se maintenir uniquement par transmission verticale. L'aptitude aux TLG apparaît alors comme une solution de maintenance de pools de gènes utiles à travers une communauté bactérienne partageant un même environnement (ou du moins des contraintes identiques). Ces TLG révèlent une plasticité des génomes procaryotiques longtemps sous-estimée (Doolittle, 1999). Cette plasticité inter-génomique est maintenant considérée comme un moteur d'innovation bactérienne participant à une modularité efficace des protéines et à une dispersion des solutions fonctionnelles avantageuses (Ochman et al., 2000 ; Boucher et al., 2003). En outre, la pression de sélection sur ces clusters pouvant se relâcher, cela laisse une marge de manœuvre évolutive favorable à l'apparition de nouvelle fonction.

Pour conclure le raisonnement, l'observation de clusters de gènes liés fonctionnellement dans les génomes et à fortiori dans des plasmides suggèrent donc fortement des événements de transferts horizontaux. Et réciproquement, montrer le transfert horizontal de gène contigus permet donc de suggérer un lien fonctionnel entre ces gènes.

2.4.2. Hypothèse des transferts horizontaux

Tentons maintenant de transposer ces considérations aux observations décrites précédemment (Fig. 48).

2.4.2.1. Des homologues forts entre espèces distantes phylogénétiquement

Le premier fait marquant qui ressort des analyses de séquence reportées ci-dessus est l'existence chez *Rhodopirellula baltica* (Planctomyces saprophyte isolé de « marine snow ») d'homologues forts à Sula et B, et EstA. Les taux d'identité observés (>40%) suggèrent fortement l'existence de substrats communs. Pour les sulfatases cela n'a rien de surprenant puisque cette bactérie hors du commun contient plus de 100 gènes de sulfatases qui à eux seuls représentent presque la moitié des séquences de sulfatases disponibles dans GENBANKnr. Mais lorsque que la même observation se reproduit avec EstA, le fait est troublant.

Des deux arbres phylogénétiques réalisés pour EstA et Sula (Fig. 46 et 47), il ressort que les séquences procaryotiques considérées ne branchent pas conformément à la phylogénie des espèces bactériennes, ce qui tend à montrer des événements d'échange horizontaux de ces gènes. De surcroît, ces deux arbres révèlent un certain isolement des deux petits clusters auxquels appartiennent les deux enzymes Sula et EstA. Ces clusters contiennent des protéines de Planctomyces (*R. baltica*) et de Bacteroïdetes (*Microscilla* sp., *B. thetaiotaomicron*, SW5).

R. baltica contient également une fucosidase bien qu'elle ne fasse pas partie des homologues les plus proches de FcsA (avec la séquence partielle dont je dispose). Elle y est voisine d'une L-fucose perméase et d'une L-fucose dehydrogénase ce qui suggère la capacité de *R. baltica* à métaboliser le L-fucose exogène.

Si FcnA n'a en tout état de cause aucun homologue identifiable dans cet organisme, on y trouve cependant le plus proche homologue connu des fucoïdane lyases (Fig. 48 ; RB1470, 27% d'identité de séquence) isolées par la société Takara (*Flavobacterium* sp. ; Takayama et al., 2002 ; cf Chapitre 1), leur deuxième et dernier homologue (Mdeg1249, 26% d'identité de séquence) étant identifié chez *Microbulbifer degradans* (γ -Protéobactéries saprophyte; Howard et al., 2003).

Toujours concernant FcnA, il me paraît nécessaire de rappeler ici que les répétitions cadhérines-like contenues par cette protéine trouvent leurs 50 meilleurs homologues chez des protéines de Protéobactéries, ceci malgré que neuf génomes de Bacteroidetes soient disponibles, cinq partiels et quatre complet (*Cytophaga hutchinsonii*, *B. fragilis*, *B. thetaiotaomicron*, et *Porphyromonas gingivalis*). Et le plus proche homologue de ces domaines cadhérines-like est de surcroît trouvé chez *M. degradans* (Fig. 48) saprophyte notoire. En outre, Fda1 et 2, qui partage leur domaine catalytique avec FcnA, sont également isolée d'une γ -Protéobactérie.

Les Planctomyces, les γ -Protéobactéries et les Bacteroïdetes sont trois phyla bactériens très distants (Fig. 47) qui ont divergé avant l'apparition des eucaryotes et donc des algues brunes et rouges et de leurs polysaccharides.

2.4.2.2. Des homologues forts entre espèces proches écologiquement (ou proximité écologique et TLG)

Le deuxième point intéressant est que *R. baltica* et *M. degradans* sont deux bactéries marines saprophytes isolées de matière organique en décomposition (respectivement « marine snow » et *Spartina alterniflora*) de même que *Microscilla* sp. PRE1 (sédiments marins de surface), *Zobellia galactanivorans* (*Delesseria sanguinea*, Rhododophytes), SW5 ou *Pseudoalteromonas galactanivorans*. A ce jour, ce sont les deux seuls organismes exploitant ce type de niche écologique marine dont le génome soit disponible (Glöckner et al., 2003 ; Howard et al., 2003). Mais un long plasmide (101 kb) nécessaire à la dégradation de l'agar isolé d'une souche de *Microscilla* a également été séquencé (Zhong et al., 2001). L'analyse de ces génomes et de ce plasmide ainsi que divers résultats expérimentaux montrent la forte aptitude de ces bactéries à métaboliser une large variété de polysaccharides présents dans le milieu

marin dont les polysaccharides d'algues (cellulose, agar, carraghénane, alginate, pullulanane, laminarine). Enfin, la figure 45 révèle la proximité de SulA avec deux sulfatases de *B. thetaiotaomicron* (Bacteroïdetes). Cette bactérie exploite le bol alimentaire et le mucus (polysaccharides sulfatés) de l'intestin humain.

Le plasmide de *Microscilla* partage avec d'autres bactéries des Flavobacteriaceae (*Z. galactanivorans*) et des γ -Protéobactéries (*Pseudoalteromonas atlantica*) des homologues forts de β -agarases impliquées dans la dégradation de l'agar d'algue rouge. *R. baltica*, *Z. galactanivorans* et *P. carragenovora* possède chacune un homologue fort de κ -carraghénase. *Z. galactanivorans* et *A. fortis* contiennent de même deux ι -carraghénases très proches l'une de l'autre. La présence d'homologue fonctionnel dans des espèces bactériennes distantes (Planctomyces, les γ -Protéobactéries, et Bacteroïdetes) implique soit une coévolution, soit des phénomènes de transferts de gènes horizontaux.

La dégradation microbienne des polysaccharides marins prend place dans des macro-agrégats de débris tissulaires, cellulaires et moléculaire de tailles plus ou moins importantes (Azam et Long, 2001 ; Engel et al., 2004 ; Thornton, 2004). L'agrégation de ces débris donne lieu à une colonisation microbienne sous forme de biofilms (ex. « marine snow », et « Transparent Exopolymeric Particules » ou TEP) qui participe à la cohésion de l'ensemble (Passow, 2002). La concentration de souches saprophytes au contact de cette matière organique est propice aux TLG (Molin et Tolker-Nielsen, 2003).

2.4.2.3. Répartitions en clusters des sulfatases de *R. baltica*

Chez *R. baltica* l'ensemble des sulfatases sont réparties en 22 clusters ; 20% d'entre elles sont analogues à des enzymes hydrolysant les liaisons ester-sulfate d'hétéropolysaccharides, alors que 29% appartiennent à un contexte génomique (cluster) impliquant le métabolisme des polysaccharides (ex. carraghénane). Pour *M. degradans*, surtout étudié pour sa capacité à dégrader la chitine, ces études n'ont pas encore été publiées.

J'ai donc procédé à la recherche et à la comparaison des clusters de *R. baltica* contenant les homologues à SulA, SulB, et EstA. Force est de constater qu'ils ne

participent pas aux mêmes clusters (Fig. 48). Cependant RB1456 (homologue de SulA) est proche du locus RB1470 meilleur homologue des deux fucoïdanes lyases de *Flavobacterium* sp., ces deux locus appartenant apparemment à un cluster lié à la dégradation du xylane (par annotation). De plus, RB13156 (homologue de EstA) est le voisin immédiat de RB13155 second meilleur homologue de SulA.

Le fait que le cluster de SW5 ne soit pas conservé à l'identique dans les organismes suscités n'est pas rédhibitoire avec l'hypothèse de TLG. Ces derniers en effet ne seraient pas nécessairement récents et concerneraient probablement des intermédiaires ce qui laisse de la place aux remaniements. Avec seulement deux génomes disponibles dans ces niches écologiques, il est légitime de penser que quelques morceaux du puzzle soient manquants.

2.4.3. Conclusions

Récapitulons. Aucun argument à l'heure actuelle ne permet de clairement montrer le lien fonctionnel entre les différents gènes que j'ai identifiés dans ce contig de SW5. Il apparaît toutefois que les meilleurs homologues de ces gènes appartiennent à des bactéries partageant des niches écologiques identiques ou proches (dégradation de matières organiques en milieu marin) ou du moins la capacité à dégrader des polysaccharides sulfatés (bactéries commensales ou pathogènes de l'intestin humain, *B. thetaiotaomicron*, *B. fragilis*, *Prevotella* sp., pouvant dégrader la mucine du mucus intestinal ou d'autres polysaccharides du bol alimentaire). Il ressort aussi que SulA a probablement pour substrat un polysaccharide. Il apparaît également que trois des meilleurs homologues appartiennent à *R. baltica* où ils participent à des clusters bien que séparés.

Il apparaît enfin que *R. baltica*, *M. degradans*, *Microscilla*, partagent, entre elles ou avec d'autres bactéries des Flavobacteriaceae (*Z. galactanivorans*) et γ -protéobactéries (*P. carragenovora*, *A. fortis*), des homologues forts de glycosidases impliquées dans la dégradation de polysaccharides d'algues. Or ces trois taxons bactériens sont très distants (Fig. 45). Ainsi leur point de divergence est nécessairement antérieur à l'apparition des polysaccharides d'algues. Donc la présence

d'homologue fonctionnel implique soit une coévolution de tous les gènes impliqués dans une voie de dégradation complexe d'hétéropolysaccharides, soit des phénomènes de transferts de gènes horizontaux entre espèces partageant de mêmes niches écologiques. C'est cette dernière hypothèse qui me sied le mieux et que j'entends soutenir. Pour ce faire j'ai donc choisi de surexprimer SulA en plus de FcnA, malheureusement l'intérêt de EstA ne m'est pas apparu suffisamment tôt, et je n'ai donc pas abordé sa surexpression.

3. Surexpression de FcnA et SulA

3.1. Surexpression de FcnA1, FcnA2, et SulA1

Les modalités de surexpression intracellulaire de FcnA1 et 2 (Fig. 49) sont détaillées dans l'article 3 ; elles ont été menées dans la souche d'*E. coli* BL21pLysS (Novagen), les constructions étant réalisées dans le vecteur pDEST17 (technologie Gateway™, Invitrogen). Toutefois deux constructions supplémentaires n'y sont pas reportées, celle de FcnA3 (Fig. 49) et SulA1. L'activité fucoïdanase recombinante s'obtient relativement facilement entre 15 et 30°C (Fig. 50). Toutefois, FcnA1, 2 et 3 ne sont pas clairement visible au sein du profil SDS-PAGE des protéines solubles totales alors que leur bande respectives apparaissent clairement. Parmi les protéines insolubles, illustrant l'importance quantitative des corps d'inclusion au cours de ces surexpressions. De plus, FcnA3 est inactive. Cette construction a malheureusement été réalisée avant que ne soient disponibles les séquences Fda1 et 2 qui ont permis de border précisément le domaine catalytique. Son but était de n'exprimer que ce dernier, mais il s'avère que souhaitant tronquer la séquence au niveau d'une charnière amont à la première répétition, j'ai amputé le domaine catalytique de ces 20 derniers résidus (Fig. 49). Faute de n'avoir recommencer cette construction, c'est cette raison que je propose pour expliquer l'inactivité de FcnA3. Le segment manquant contient en effet une hélice α prédite qui pourrait clore le TIM-barrel putatif. Sans cette dernière, il est probable que FcnA3 ne puisse pas se replier correctement et soit systématiquement insoluble ou du moins inactive. Ce résultat négatif ainsi que le succès de FcnA1 et 2

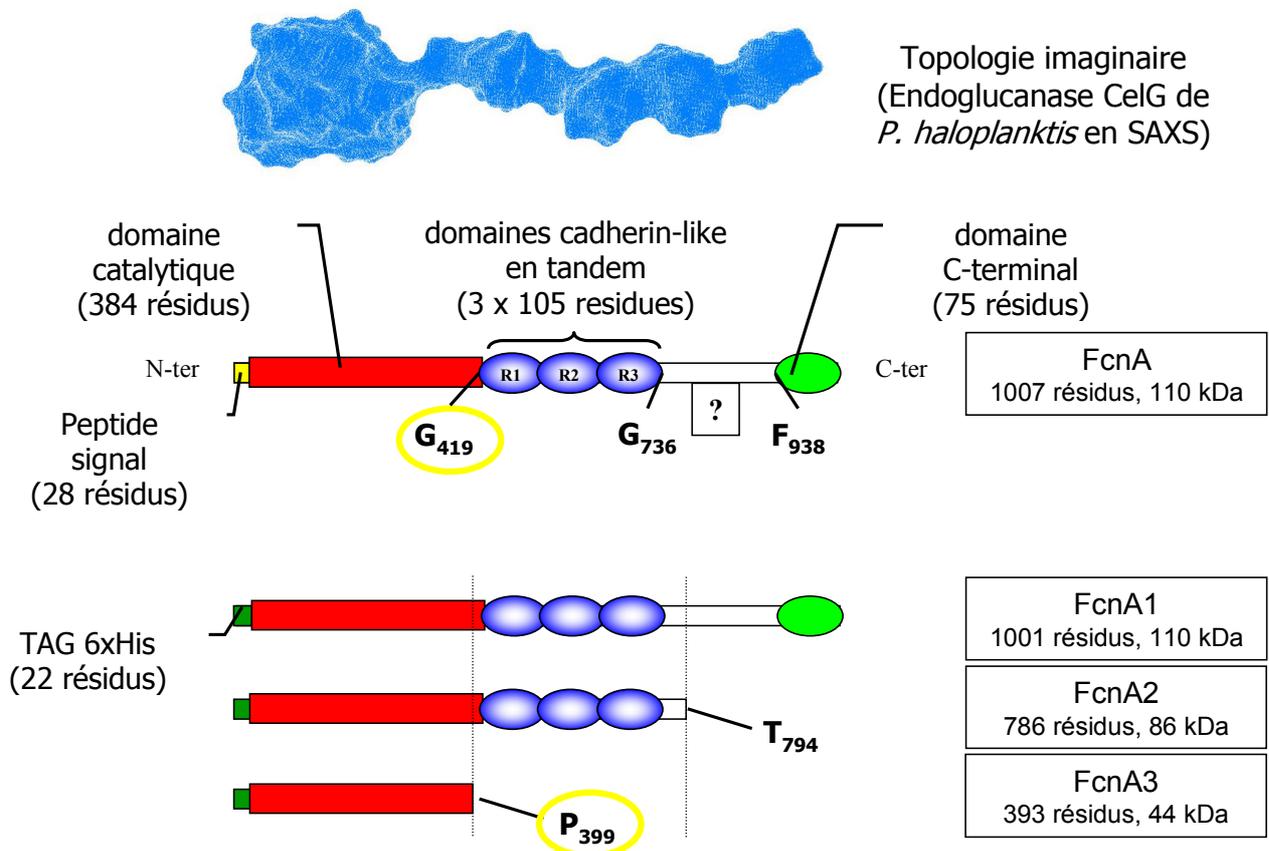


Figure 49 : La modularité de FcnA et son influence sur les constructions réalisées pour sa surexpression FcnA1, FcnA2 et FcnA3. FcnA1, FcnA2 et FcnA3 ont été clonés dans le vecteur de surexpression pDEST17 (technologie Gateway™, Invitrogen). On remarquera que dans FcnA3 le tronquage ne respecte pas la fin du domaine catalytique. La topologie de l'endoglucanase CelG (*P. haloplanktis*) résolue en SAXS (Czjzek, communication personnelle) donne une idée approximative de la forme en marteau que FcnA pourrait adopter.

appuient néanmoins l'analyse de séquence proposée ci-dessus et surtout valident le travail de clonage ainsi que la séquence de FcnA.

A l'inverse de FcnA, la surexpression de Sula en entier à travers la construction pDEST17SulA1 (tag N-terminal de 22 résidus contenant un segment de six histidines) est plus performante dans la mesure où SulA1 est clairement visible dans un extrait total de protéines solubles séparées en SDS-PAGE (Fig. 50). J'ai donc réalisé des tests préliminaires de l'activité de SulA1 sur le fucoïdane FS28, soit seule soit en présence de FcnA1 ou 2. Pour ces tests, j'ai utilisé des extraits total de protéines solubles et j'ai suivi l'activité des enzymes par C-PAGE, autrement dit en souhaitant observer leur effet sur l'apparition/disparition d'oligosaccharides chargés (Fig. 50). La présence de SulA1 seule ne génère pas d'oligosaccharides comme attendu. Mais les mêmes oligosaccharides sont observés sous l'action FcnA1 ou 2 avec ou sans SulA1. Dès lors deux hypothèses sont envisageables. Soit cette dernière n'agit pas sur les produits de FcnA, ce qui n'exclue pas qu'elle agit en amont de la dégradation, soit la protéine n'est pas active puisqu'on sait activité dépendante d'une modification post-traductionnelle. Néanmoins d'autres sulfatases ont déjà été surexprimées actives chez *E. coli*. De plus si l'on compare attentivement les dégradations menées par l'une ou l'autre des hydrolases avec ou sans SulA1, on devine en présence de la sulfatase une intensification de la bande correspondant au produit limite DP2. Le C-PAGE n'est certes pas une méthode quantitative, néanmoins l'effet est notable. Bien que restant à confirmer, une telle action de Sula ne pourrait s'expliquer que par un démasquage des sites de coupures reconnus par FcnA.

Je ne suis pas allé plus loin dans mes investigations sur Sula, préférant accorder un maximum d'effort à l'optimisation de la production de FcnA sous forme recombinante ainsi qu'à sa purification.

3.2. Optimisation de la production de fucoïdanase recombinante

J'ai souhaité optimiser la production de FcnA recombinant afin de rapidement disposer d'une voie de production permettant l'exploitation de l'enzyme. Disposant d'un milieu auto-inducteur (Studier, communication personnelle) dont le principe

repose sur la présence de glucose et de lactose (milieu pour souche DE3 où la T7 DNA polymérase est sous le control du promoteur de l'opéron lactose avec donc une répression catabolique du glucose), j'ai pensé qu'il serait intéressant de l'exploiter dans le cadre de production de routine. Il permet en effet de s'affranchir du control de la croissance avant induction et de favoriser un bon rendement d'expression en induisant à une forte concentration cellulaire. C'est donc ce dernier que j'ai employé pour optimiser la production, bien que la surexpression fonctionne également en M9. Toutefois un tel objectif est ici compliqué par l'absence d'une méthode de dosage de l'activité fucoïdanase. Le suivi de son expression passe donc nécessairement par des méthodes non quantitatives. Plusieurs étapes préliminaires réalisées en batch/erlenmeyer ne seront pas développées ici au profit d'une optimisation de production en fermenteur. Les approches que j'ai adoptées ici consistaient à suivre en parallèle de la culture, la production des protéines solubles totales et des corps d'inclusion, l'activité par C-PAGE, et la production spécifique de FcnA soluble par Western blot (anticorps anti His-tag). La culture de BL21pLysS (DE3) / pDEST17FcnA1 a été menée à 20°C et les résultats sont présentés figure 51. Un soin tout particulier a été porté à la constance des volumes durant le traitement de tous les échantillons de manière à réduire au mieux les biais de variation d'intensité des bandes dans les séparations électrophorétiques.

La culture bactérienne est en phase exponentielle jusqu'à environ 1,7 uDO₆₀₀ (16h) où apparaît un léger infléchissement de la courbe. La croissance se poursuit jusqu'à 13,4 uDO₆₀₀ (28h) avec une diminution du taux de croissance (Fig. 51A). Cet infléchissement coïncide très probablement avec une levée progressive de la répression catabolique du glucose en faveur de la métabolisation de lactose et/ou le début de l'induction de l'expression de FcnA. Cette observation est corrélée avec le début de la production des corps d'inclusion à 18h (Fig. 51A et C) dont la quantité augmente jusqu'à 28h. La figure 51C montre que les corps d'inclusion purifiés ne contiennent pas uniquement la forme entière de FcnA1. Le dosage de la forme insoluble de cette dernière est donc surestimé dans la figure 51A. La sensibilité du test d'activité a été volontairement diminuée (temps de réaction court : 30 min) afin d'éviter des problèmes de saturation du test, autrement dit, de façon à éviter qu'aucun échantillon protéique n'occasionne une digestion à complétion. De cette manière la variation de la

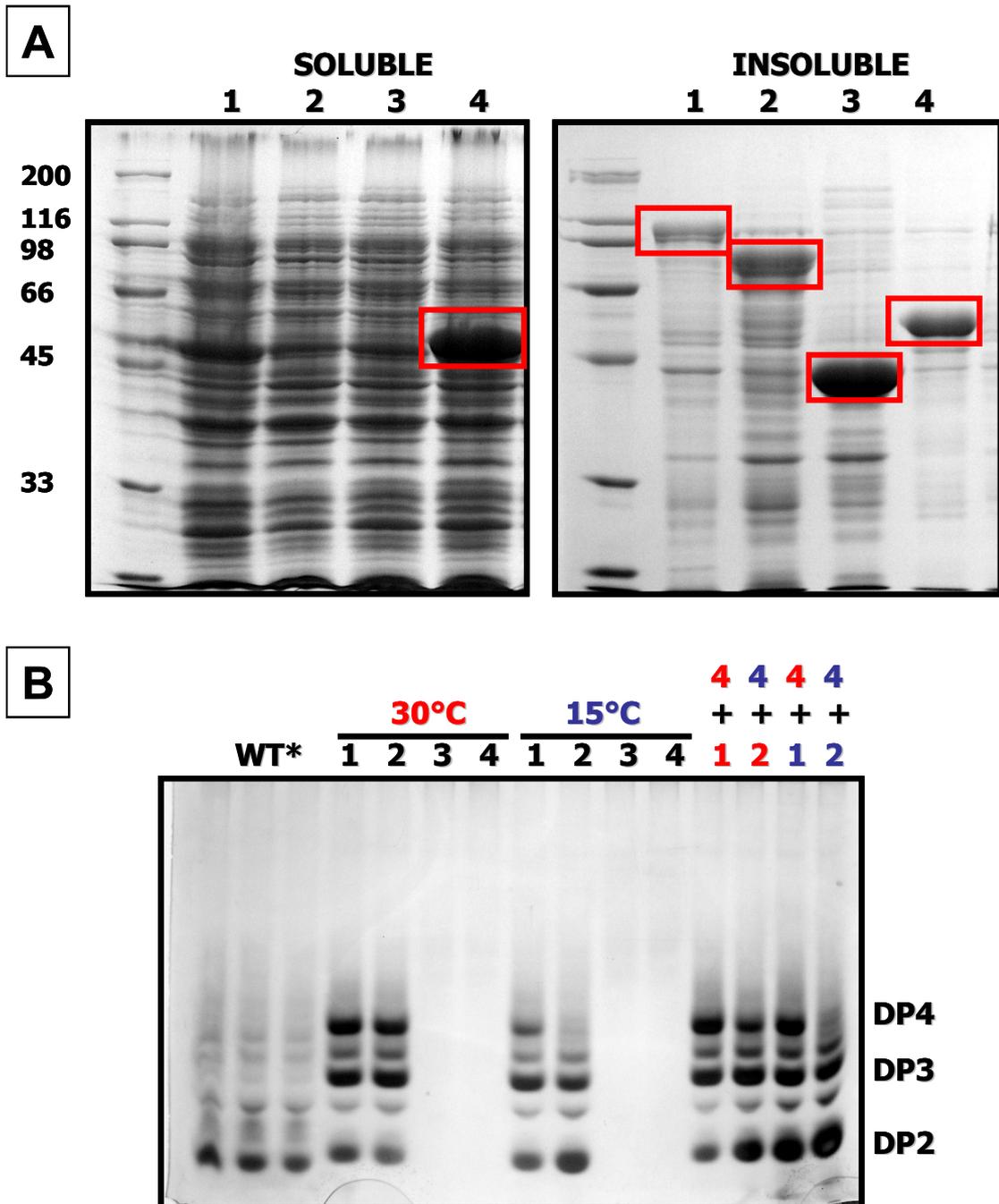


Figure 50 : Surexpression de FcnA1 (1), FcnA2 (2), FcnA3 (3) et Sula1 (4). A : analyses par SDS-PAGE des extractions de protéines solubles et insolubles. Le gel est coloré en bleu de Coomassie. Les différentes constructions sont clonées dans pDEST17 et surexprimées dans BL21pLysS en milieu autoinducteur sur la nuit à 30°C; **B : Analyse par C-PAGE de l'activité recombinante de dégradation du fucoïdane.** Les extraits protéiques employés proviennent de la phase soluble de lysat bactérien. Les surexpressions sont menées à 30 et 15°C dans des conditions identiques à celles décrites précédemment. Des combinaisons d'extraits contenant FcnA1 ou 2 avec Sula1 sont testées (5% d'extrait v/v, 0,2% p/v FS28, 24h de réaction à température ambiante). *Les produits d'hydrolyse sont comparés avec ceux obtenus à complétion avec un extrait de FcnA sauvage partiellement purifié (fractionnement au sulfate d'ammonium).

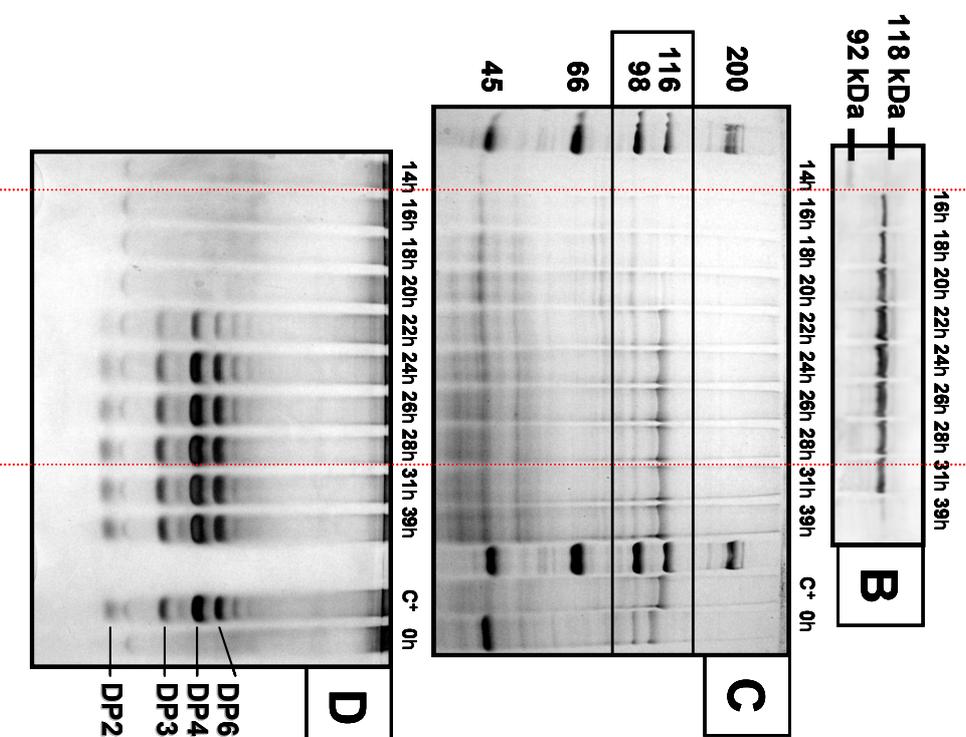
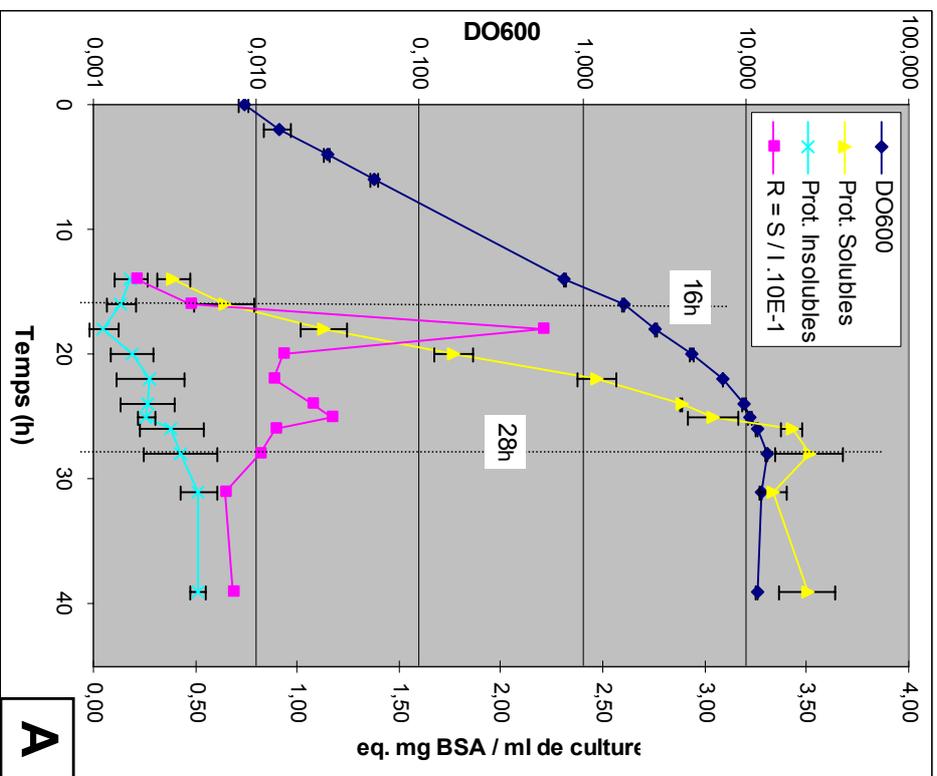


Figure 51 : Optimisation de la production par fermenteur de FcnA1 en milieu autoinducteur. La culture (5L) est menée en fermenteur, avec un milieu auto-inducteur (ZYP5052 ; Studier, communication personnelle) à 20°C, à pH régulé (7,2) et pO₂ contrôlée. Chaque prélèvement est réalisé en triplicat. La fraction soluble des protéines est obtenue par Lyse cellulaire au Bugbuster (Novagen) et élimination de la fraction insoluble par centrifugation. Les corps d'inclusion sont purifiés de cette dernière et solubilisés dans l'urée 8M. **A : suivi quantitatif de la biomasse, des protéines solubles (S) et des corps d'inclusion (I).** La méthode de Bradford est employée pour la mesure des concentrations protéiques. **B : Suivi de FcnA1 dans la fraction soluble par Western blot.** La détection de FcnA1 utilise un anticorps anti His-tag couplé au kit ECF (Amersham). **C : Suivi des corps d'inclusions par SDS-PAGE.** Coloration au bleu de Coomassie. **D : suivi de l'activité par C-PAGE.** Le milieu réactionnel contient 5% v/v d'extrait protéique et 0,2% p/v de fucoïdane FS28 ; la réaction se fait sur 30 min, à température ambiante, en simultané pour tous les échantillons.

teneur en enzyme active reste visible à travers la progression des digestions du substrat (plus les produits sont petits et les produits limites –DP2 et DP3– s’accumulent, plus la réaction est forte). On observe donc une apparition de l’activité à 22h (7,2 uDO₆₀₀). Cette dernière augmente jusqu’à 28h et semble ensuite se stabiliser (Fig. 51D). Avec des réactions de digestion plus longues (2 h) l’activité apparaît dès la 20^{ème} heure de culture bien que très faiblement (résultats non montrés). La détection plus précoce de FcnA1 soluble en Western blot (16h), quoique surprenante comparée à l’analyse de l’activité en C-PAGE, semble révéler une meilleure sensibilité de cette méthode (Fig. 51B). Malgré le dépôt d’un volume constant d’échantillon sur le gel, cette dernière n’est cependant pas quantitative, de par le système de révélation employé et l’absence de standard interne. Cette expérience nous renseigne finalement peu sur l’évolution de l’équilibre des formes soluble / insoluble de FcnA1 au cours de l’induction dans ce milieu. Il semble que l’activité et la quantité de corps d’inclusion progressent conjointement alors que le Western blot semble indiquer un pic de forme soluble précédent la montée de la forme insoluble traduisant un hypothétique basculement de la première vers la seconde.

Il ressort de l’ensemble de ces résultats que la protéine active apparaît dès 2 uDO₆₀₀ et que la quantité d’activité ramenée au volume de culture augmente jusqu’à 13 uDO₆₀₀. Il n’est cependant pas possible de rapporter la quantité d’activité ou de FcnA1 soluble à la biomasse. Or il serait pourtant souhaitable pour faciliter la purification de l’enzyme d’intérêt de récolter les cellules au moment où l’activité spécifique de la phase soluble des protéines intracellulaire est maximale. Néanmoins, si dans le cadre d’application de l’enzyme, seule la quantité d’activité par unité de volume de culture importe alors je préconise, pour ces conditions de surexpression, une récolte à 28h.

3.3. Purification de la fucoïdanase recombinante

FcnA1 et 2 ont été conçues avec un tag N-terminal 6xhis dans le but de pouvoir faciliter leur purification à partir d’un lysat bactérien clarifié, par chromatographie d’affinité sur des matrices portant des ions de Ni²⁺. Le protocole employé est décrit en

détail dans l'article 3. Malgré une élution de l'activité dans un pic franc et unique à environ 150 mM d'imidazole, l'analyse des fractions de ce pic par SDS-PAGE montre, pour FcnA1 et 2, la présence de plusieurs espèces protéiques (Fig. 52). Une d'entre elle, estimée à 75 kDa (P75), est clairement majoritaire. Elle est accompagnée d'une bande de plus faible intensité estimée à 70 kDa (P70). Elles font partie d'un ensemble de bandes contaminantes communes à la purification de FcnA1 et 2, suggérant que ces espèces sont probablement issues d'une protéolyse de la protéine surexprimée. La purification ayant été menée à froid (4°C) dès la récolte des cellules, et en présence d'un cocktail d'antiprotéases, il est probable que, si protéolyse il y a, elle ait lieu pendant la culture. Toutefois le tampon de purification dépourvu d'EDTA ne permettait pas d'inhiber les métalloprotéases.

Diverses tentatives ont été réalisées pour éliminer ces protéines contaminantes de la purification de FcnA1. Dans un premier temps, le pool issu de la chromatographie d'affinité a été utilisé pour une chromatographie sur résine échangeuse d'anions (résine type MonoQ). Là encore les bandes majoritaires de P75 et P70 ont co-éluées avec l'activité. Dans un deuxième temps, j'ai essayé de séparer par leur taille les différentes espèces protéiques présentes dans ce même pool par chromatographie de perméation de gel sur une résine type Superdex 200. De façon surprenante la protéine P75 est sortie avant FcnA1 (11 kDa) qui a elle même co-élué avec P70. Cette dernière chromatographie a réduit la contamination à une seule espèce, mais FcnA1 est clairement minoritaires dans les fractions actives. De surcroît, FcnA1 bien que sortant à la taille attendue d'après la calibration de la colonne est précédé par P75 et accompagnée par P70. Se pose donc le problème de la nature de ces deux protéines qui ne pourra être levé que par un microséquençage de ces dernières. Je n'ai pas poursuivi plus avant la mise en place d'un protocole permettant de purifier à l'homogénéité la fucoïdanase recombinante. Et j'ai donc exploité cette activité à partir du pool d'enzyme issu de la chromatographie d'affinité.

3.4. Analyse de l'activité de FcnA1 et 2

Dans le but de comparer les comportement de FcnA1 et FcnA2 ainsi que d'évaluer le potentiel d'activité contenu dans les pool d'enzymes purifiés par chromatographie d'affinité,

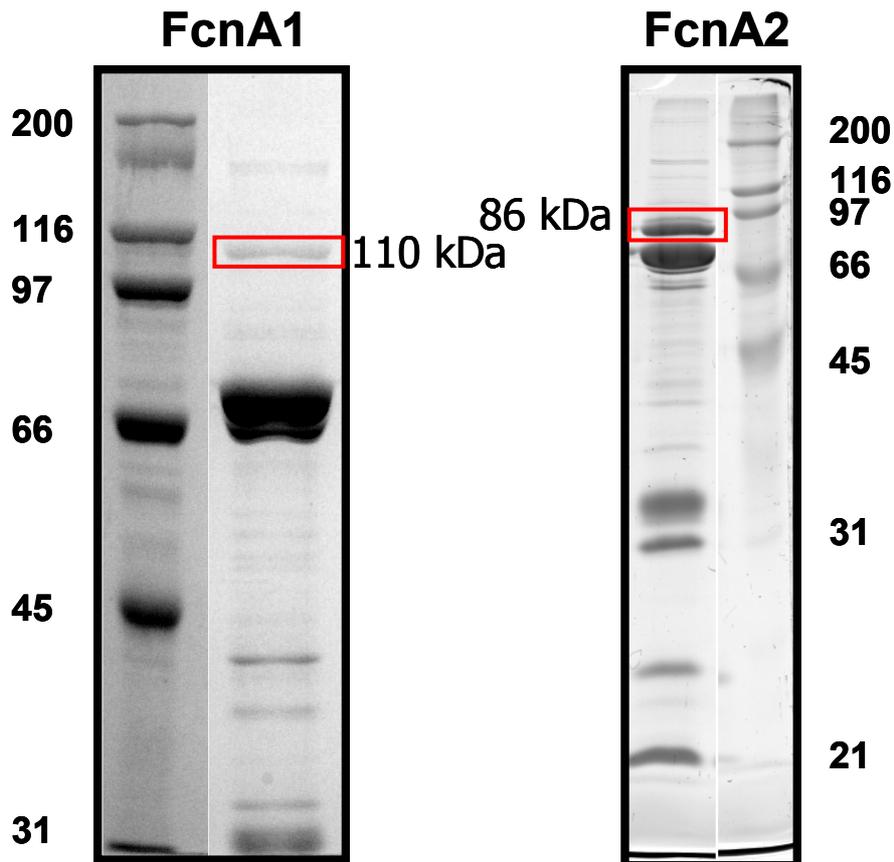


Figure 52 : Purification de FcnA1 et 2 par chromatographie d'affinité. Les protéines recombinante sont produites en milieu autoinducteur à 20°C dans BL21pLysS. Les protéines solubles sont obtenues par lyse du culot cellulaire et clarification du lysat par centrifugation. Elles sont ensuite séparées par chromatographie selon leur affinité pour Ni²⁺. L'activité fucoïdanase élue à environ 150 mM imidazol. Les fractions concernées sont rassemblées et analysées par SDS-PAGE.

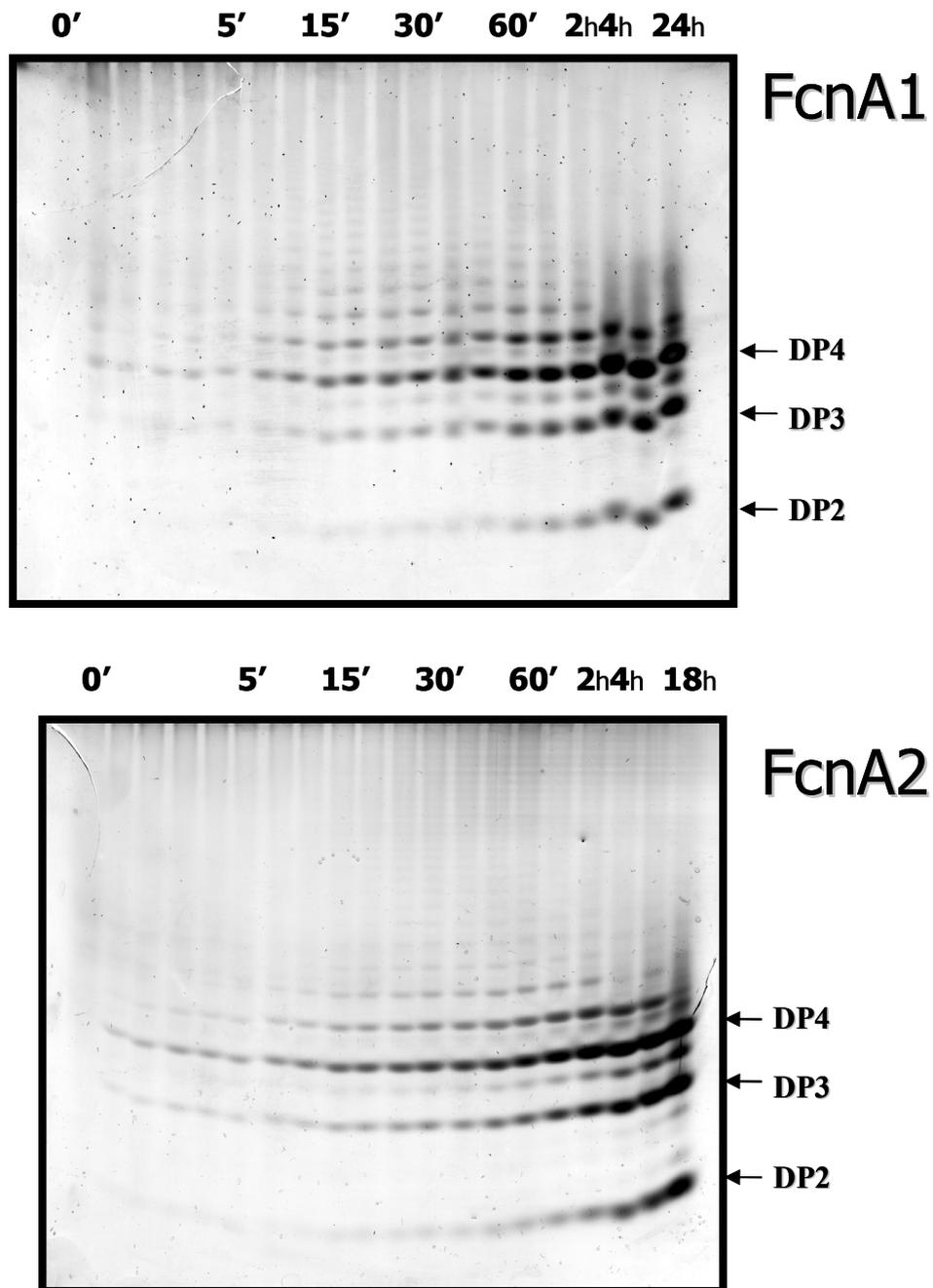


Figure 53 : Cinétique d'hydrolyse du fucoidane FS28 par les deux protéines recombinantes FcnA1 et 2 purifiées par chromatographie d'affinité. Les réactions ont été menées avec 1 μL d'extrait enzymatique provenant d'un pool de purification par chromatographie sur Ni-sepharose, dans 250 μL de substrat à 2 $\text{mg}\cdot\text{ml}^{-1}$ de fucoidane FS28 (*P. canaliculata*). Chaque temps correspond à un prélèvement provenant d'une réaction unique.

j'ai réalisé deux cinétiques d'hydrolyse du fucoïdane FS28 avec chacun des deux enzymes. Ces deux cinétiques ont été suivies par C-PAGE (Fig. 53). Les deux extraits enzymatiques permettent d'arriver à complétion en approximativement 20 h. Compte tenu des conditions de réactions utilisées (400 ml de la culture en fermenteur décrite ci-dessus sont récoltés à 25h ; le culot cellulaire correspondant permet l'obtention d'un pool d'enzyme de 20 mL à l'issue de la chromatographie d'affinité ; 1 μ L d'enzyme provenant de ce pool dans 250 μ L de substrat à 2 mg.mL⁻¹ de FS28), une culture de 1 L permet potentiellement de digérer à complétion 25 g de fucoïdane FS28 en moins de 24h à température ambiante et sans agitation.

Les deux enzymes se comportent de façon similaire avec un mode d'action endo marqué. Toutefois les produits limites apparaissent très tôt de façon identique à ce qui a été observé pour la λ -carraghénase (cf. Chapitre 2) sans que ça ne traduise nécessairement un comportement processifs. Descamps-Lepagnol (1998) au cours de sa thèse a purifié et identifié par RMN deux produits limites de digestions obtenu avec un pool d'enzyme sauvage non pure (simple fractionnement aux sulfate d'ammonium des protéines du surnageant de culture de la souche SW5). Ne disposant pas de ces deux oligosaccharides caractérisés, je n'ai pas pu les faire migrer en parallèle des digestions réalisées avec l'enzyme recombinant. L'attribution des identités DP2, DP3 et DP4 s'est donc faite par comparaison de profils de séparation C-PAGE distincts (ceux réalisés par mes soins, et ceux réalisés par Descamps-Lepagnol). Si ces comparaisons mettent en évidence des correspondances relativement fiables, il ne s'agit évidemment pas d'une preuve catégorique permettant d'identifier les produits de FcnA1 ou 2.

4. Conclusions générales et perspectives

Les objectifs de ce travail sont dans une large partie atteints puisque le gène de FcnA est caractérisé et que sa surexpression hétérologue fonctionne. Néanmoins il faut d'une part reconnaître que l'absence de méthode de dosage de l'activité de FcnA pénalise lourdement la caractérisation de ce catalyseur et l'optimisation de sa production. D'autre part, les problèmes rencontrés lors de la purification de

l'enzymerecombinant entame pour l'instant la perspective d'un programme d'études structurales à court terme. Cependant, d'un point de vue industriel, l'exploitation de l'enzyme au niveau pilote est envisageable.

Un manque de réactivité des extrémités réductrices générées explique probablement l'inefficacité des méthodes de dosage habituelles puisque même lorsque l'hydrolyse est observable en RMN elle demeure non mesurable par les sucres réducteurs. On sait en effet la réactivité de l'extrémité réductrice des méthyl-pentoses plus faible. Similairement, les produits de la λ -carraghénase, eux aussi sulfatés en 2 de l'extrémité réductrice libérée, ne sont eux non plus pas très réactifs (contrairement aux produits de la κ - et ι -carraghénase hydrolysant la même liaison). Sans que je puisse l'expliquer (effet inducteur, perturbation électrostatique ?) ce substituant sulfate de C2 pourrait peut être perturber la réactivité de l'extrémité réductrice voisine ou ralentir la cinétique de conversion entre la forme cyclique du sucre et sa forme aldéhydique réactive.

A l'issu de ce travail, les suites immédiates que je propose, reposent tout d'abord sur une reconstruction du vecteur d'expression visant à n'exprimer que le domaine catalytique. Du succès de cette entreprise dépend les espoirs d'amélioration de l'expression et surtout de la purification de l'enzyme recombinant. Cette étape est indispensable à une étude biochimique fine de l'enzyme, affinage qui ne pourrait qu'être bénéfique à son exploitation industrielle même si il n'est pas absolument indispensable pour celle-ci. A court terme également je propose de tester l'hypothèse des résidus catalytiques putatifs désignés par l'analyse de séquence afin de confirmer cette dernière. Pour se faire la réalisation des mutants D→A et D→N me semble adéquate. Bien qu'une telle étude se doit d'être renforcée par le calcul des constantes cinétiques, ici, l'abolition ou le maintien de l'activité apporterait déjà des éléments de réflexion intéressants qui pourraient jalonner partiellement une approche structurale. La famille de glycoside hydrolases révélée au cours de ce travail justifie de par sa nouveauté et son originalité cet effort d'investigation.

La complexité du fucoïdane fait que pour l'instant la connaissance de ce dernier et celle du comportement de FcnA progressent de façons concomitantes. Cette double progression implique la caractérisation de ses produits. Descamps-Lepagnol au cours

de son travail de thèse (1998) avait caractérisé deux produits d'hydrolyse (cf Article 2). Malencontreusement, alors qu'elle disposait de l'enzyme pure, ces derniers avaient été produits par un pool d'enzyme relativement brut. Si des correspondances sur C-PAGE indiquaient qu'il s'agissait effectivement des produits de FcnA, le doute demeurait. La disponibilité de l'enzyme recombinant même non pure facilite désormais cette approche qui est principalement limitée par la purification en quantité suffisante des produits de l'enzyme. Il s'avère cependant que ce dernier a une action relativement superficielle sur le substrat brut. Les travaux de Deniaud au laboratoire (cf. Article 3) ont ainsi montré qu'une pré-hydrolyse acide du fucoïdane, le sensibilise à l'action de FcnA. Les produits limites de l'enzyme recombinant n'ont pas encore été purifiés. Toutefois, cette amélioration lui a permis d'analyser par RMN les variations de signal du polysaccharide lors de l'hydrolyse occasionnée par FcnA. Ainsi sont apparus différents pics déjà assignés dans la littérature, de sorte que l'hydrolyse de la liaison $\alpha(1\rightarrow4)$ a pu être démontrée au sein de motifs correspondant à ceux décrit par Descamps-Lepagnol dans les tétra- et hexasaccharide qu'elle avait purifié. Cette liaison $\alpha(1\rightarrow4)$ semble distinguer le fucoïdane extrait des Fuciales de celui provenant des Laminariales. En outre l'endo-fucoïdanase extracellulaire caractérisée par la société Takara provenant de la souche bactérienne SN-1009 correspond selon toute vraisemblance (taille de 100 kDa et gène cloné) à l'une des deux protéines Fda1 ou Fda2 brevetées par la même société et dont les gènes ont été clonés à partir de la même souche. Or il s'avère que l'enzyme caractérisé est actif sur le fucoïdane de Laminariales mais pas sur celui de Fuciales. Bien que ces produits ne soient pas encore divulgués, je pense donc que cet enzyme agit sur la liaison $\alpha(1\rightarrow3)$ du fucoïdane. Cela pourrait en outre contribuer à expliquer la forte divergence de séquence existant entre FcnA et Fda1 ou 2, malgré leur appartenance à une même famille de glycoside hydrolase.

La résistance du fucoïdane natif à FcnA souligne la nécessité d'enzymes supplémentaires pour satisfaire le catabolisme de ce polymère. De ce point de vue, l'analyse du contig auquel appartient FcnA est prometteuse. La surexpression réussie de Sula ouvre des possibilités d'investigation directement liées à la connaissance du fucoïdane, de son catabolisme et indirectement de FcnA. Les premiers tests que j'ai pu

réaliser, pour préliminaires qu'ils soient, ne sont pas négatifs. La démarche que je suggère à l'avenir, est de tout d'abord confirmer la modification post-traductionnelle indispensable à l'activité de SulA1. Pour ce faire, la purification de l'enzyme recombinant à homogénéité, suivi d'une analyse en spectrométrie de masse me semble toutes indiquées. Cependant la modification d'une cystéine en formyl-glycine n'entraîne vraisemblablement pas une variation de masse suffisante pour distinguer les deux espèces par cette technique. C'est pourquoi l'oxydation de la cystéine par un composé de masse suffisante pourrait s'avérer nécessaire (SulA1 contient uniquement deux cystéines). Dès lors la détection de son action éventuelle sur le fucoïdane passe probablement par l'analyse HPLC de la libération de sulfate dans le milieu réactionnel avec ou sans FcnA. Par ailleurs, tout comme SulA, la surexpression de EstA me semble intéressante à initier. En effet, que l'on montre ou pas l'implication de ces deux protéines dans la dégradation du fucoïdane, l'originalité de ces deux protéines au sein de leur familles respectives ajoutée à la forte probabilité qu'elles interagissent avec un polysaccharide marin, justifient à mon sens l'effort d'une approche structurale de ces deux protéines. Enfin, concernant les deux gènes partiels du contig, il existe désormais des solutions biotechnologiques permettant l'obtention du complément de séquence sans repasser par une étape de criblage de la banque génomique. Une solution aussi économique en temps est très attractive à défaut de disposer du génome de SW5.

Les perspectives à ce travail sont riches et de nombreuses interrogations restent en suspend. Si elles se proposent d'approfondir la connaissance du fucoïdane d'algue et de son catabolisme, elles projettent aussi l'espoir de disposer de nouveaux biocatalyseurs permettant pourquoi pas de moduler l'activité d'oligofucanes ou de faciliter leur obtention à des fins industrielles. Enfin, l'idée d'une propagation des fonctions du catabolisme des polysaccharides d'algue, par transfert horizontal, au sein de la communauté bactérienne saprophyte marine me semble séduisante. Elle mérite donc d'être approfondie. La méta-génomique couplée à des approches biochimiques offrent de ce point de vue des perspectives intéressantes (Voget et al., 2003 ; Azam et Worden, 2004). Si cette hypothèse se vérifiait, elle impliquerait très probablement une diversité réduite de ces biocatalyseurs.

ARTICLE 2

Isolation and Culture of a Marine Bacterium Degrading the Sulfated Fucans from Marine Brown Algae

Valérie Descamps¹, Sébastien Colin¹, Marc Lahaye², Murielle Jam¹, Christophe Richard¹, Philippe Potin¹, Tristan Barbeyron¹, Jean-Claude Yvin¹ and Bernard Kloareg¹

¹UMR 7139 (CNRS, Laboratoires Goëmar and Université Pierre et Marie Curie), Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, Brittany, France. ²INRA, Unité de Recherche sur les Polysaccharides, leurs Organisations et Interactions, BP 71627, 44316 Nantes Brittany, cedex 03, France

Key words: *fucoidanase*, *fucoidan*, *fucan oligosaccharides*, *fucallean brown algae*, *marine bacteria*, *Flavobacteriaceae*.

Running title: Fucoidan hydrolase activity from a marine *Flavobacteriaceae*

Correspondence to T. Barbeyron, UMR 7139, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, Brittany, FRANCE. Fax: +33 (0)2 98 29 23 24, Phone: +33 (0)2 98 29 23 32, E-mail: barboun@sb-roscoff.fr

Abstract: Fucoidans are matrix polysaccharides from marine brown algae, consisting of an α -L-fucose backbone substituted by sulfate-ester groups and masked with ramifications containing other monosaccharide residues. In spite of their interest as biologically-active compounds in a number of homologous and heterologous systems, no convenient sources with fucanase activity are available yet for the degradation of the fucallean algae. We here report on the isolation, the characterization and culture conditions of a bacterial strain capable of degrading various brown algal fucoidans. This bacterium, a member of *Flavobacteriaceae*, is shown to secrete fucoidan endo-hydrolase activity. An extracellular enzyme preparation was used to degrade the fucoidan from the brown alga *Pelvetia canaliculata*. End products included a tetrasaccharide and a hexasaccharide made of the repetition of disaccharididic units consisting of α -1 \rightarrow 3-L-fucopyranose-2-sulfate- α -1 \rightarrow 4-L-fucopyranose-2,3-disulfate, with the 3-linked residues at the non-reducing end.

INTRODUCTION

Besides alginate, a linear polymer of β -D-mannuronic acid and its C₅ epimer, α -L-guluronic acid, brown algae (Phaeophyta) contain matrix polysaccharides, known as fucoidans, made of sulfated-L-fucose residues (Mabeau and Kloareg, 1987), which, in the fucoid alga *Pelvetia canaliculata*, represent as much as ca. 40% of the cell wall dry weight (Kloareg, 1984). Fucoidans consist of continuous spectrum of highly ramified polysaccharides with a complex and still somewhat elusive structure, ranging from high uronic acid, low-sulfate-containing polymers with significant proportions of D-xylose, D-galactose, and D-mannose to highly sulfated homofucan molecules (Kloareg and Quatrano, 1988; Mabeau et al., 1990).

In brown algal zygotes, sulfated fucans are thought to be involved with adhesion to the substratum as well as with cell polarization, through a trans-membranar complex reminiscent of the focal adhesions of mammalian fibroblast and epithelial cells (Kropf et al., 1988; Goodner and Quatrano, 1993). Sulfated fucans have also been extensively investigated for their anticoagulant properties, which involve the inhibition of thrombin and/or the activation of antithrombin and heparin factor II (eg. Millet et al., 1999; Pereira et al., 1999; Thorlacius et al., 2000). In addition, by interfering with cell-cell recognition in mammalian systems, they display a variety of biological activities. They inhibit the acrosomal reaction in human spermatozoa (Mahony et al., 1991, 1993) and block infections by such viruses as the human immunodeficiency virus (Béress et al., 1993; Hoshino et al., 1998), the vesicular stomatitis virus (Mayer et al., 1987) or the herpes simplex virus (Hoshino et al., 1998), as well as by other microbes, such as *Plasmodium knowlesi* (Dalton et al., 1991). Sulfated fucans are potent inhibitors of smooth muscle cell growth, both in vitro and in vivo (McCaffrey et al., 1992; Logeart et al., 1997), and they display antiproliferative activity in various cancer types, including a non small-cell bronchopulmonary carcinoma line (Riou et al., 1996) and Erhlich carcinoma (Zhuang et al., 1995). Also, sulfated fucans dose-dependently but non specifically inhibit selectin-mediated leukocyte rolling and emigration, thus decreasing organ injuries in non infectious models of inflammation (Ley et al., 1993; Wikström et al., 1995; Ostergaard et al., 2000; Linneman et al., 2000). Finally, sulfated fucans may be used in crop disease control since oligomers from sulfated fucans induce systemic resistance in tobacco against tobacco mosaic virus (Klarzynski et al., 2003).

In this context, a reliable fucoidan-degrading enzyme preparation would be highly desirable to obtain oligofucoidans, both to further elucidate the fine chemical structure of

fucoindans and to investigate their structure-activity relationships in homologous or heterologous systems. Yet, only a few attempts were made so far to isolate fucoindan-degrading enzymes. In the case of herbivorous molluscs, Thanassi and Nakada (1967) partially purified a fucoindanase with a molecular mass comprised between 100 kDa and 200 kDa from *Haliotis rufescens* and *H. corrugata*. These activities hydrolysed the fucoindan from *Fucus gardneri* to oligosaccharides ranging from the decasaccharide to L-fucose. Fucoindanases were also partially purified from *Pecten maximus* (Daniel et al., 1999) but the activity contained in the digestive glands of this mollusc was essentially a highly active and unusually thermal stable exo-fucosidase of 200 kDa (Berteau et al., 2002), mixed with a putative fucoindanase activity (Daniel et al., 1999) as also reported for the bivalve *Patinopecten yessoensis* (Kitamura et al., 1992). As far as bacteria are concerned, a study of Laminaria-decomposing epiflora indicated that fucoindanase activities are inducible, weak, but not rare in marine bacteria (Uchida, 1995). A number of *Vibrio*-type, fucoindan-degrading bacteria were isolated from marine sediments (Morigana et al., 1981) and an exo-acting fucoindanase was purified from *Vibrio* sp. N-5 (Furukawa et al., 1992). Recently, Sakai et al. (2002) isolated from the Japan Sea a marine bacterium, *Fucobacter marina* (Flavobacteriaceae), which cleaved various fucoindans from the Laminariales *Kjellmaniella crassifolia*, *Undaria pinnatifida* and *Lessonia nigrescens* but not from the Fucales *Fucus vesiculosus* and *Ascophyllum nodosum*. The enzyme was further characterized as a sulfated fucoglucuronomannan lyase (Sakai et al. 2003a). Finally, a bacterial strain belonging to the family of Verrucomicrobiaceae was isolated from the gut of the sea cucumber *Stichopus japonicus* (Sakai et al. 2003b). This bacterium, *Fucophilus fucoindanolyticus*, was able to degrade a variety of fucoindans, suggesting it produces a number of fucoindan-digesting enzymes. Interestingly, *F. fucoindanolyticus* was not able to cleave the fucan of the host sea cucumber (Sakai et al. 2003b).

Altogether, no fucoindan-endohydrolase activity is conveniently available yet to degrade the fucoindans from fucallean brown algae. As Fucales are more likely than Laminariales to provide large amounts of sulfated fucoindans (Mabeau et al., 1990), we thus set out to search for convenient sources of such enzymes. Alginate-extraction plants are known to release liquid effluents that are enriched in brown algal sulfated fucoindans (Fleury and Lahaye, 1993). Muds in the water-treatment facilities used for the cleansing of these effluents could therefore be expected to be good candidates for the occurrence of fucoindan-degrading micro-organisms. We here report on the isolation and the characterization, from this latter habitat, of a fucoindan-

degrading bacterium. The fucoidanalytic activity from culture supernatant was capable of degrading sulfated fucoidans from various fucoid brown algae.

MATERIALS AND METHODS

Preparation of Fucoidans

Thalli of *Pelvetia canaliculata* (Dcne et Thur.) (1 kg fresh weight) were collected during low tide at Roscoff (Brittany, France), freed from epiphytes and washed with distilled water. They were crushed in a mortar in the presence of liquid nitrogen and macerated overnight in 1-L ethanol/formaldehyde/H₂O (80:5:15 vol:vol). Algal fragments were then extracted by 2 L of ethanol/formaldehyde/H₂O followed by 2 L of acetone. The resulting pellet was dried at 60 °C and extracted twice for 3 h at 70 °C by a 0.01 N HCl solution supplemented with 4% (wt:vol) CaCl₂. The extract was filtered, concentrated with a rotary evaporator, neutralized with ammonium carbonate and fucoidans were precipitated with 2.5 volumes of ethanol. The precipitate was redissolved in water and freeze-dried (Lyolab ALSL, Secfroid). This crude fucoidan fraction (FS28) typically exhibited a total carbohydrate content of ca. 55%, as assayed colorimetrically according to Tillmans and Phillipi (1929), and the proportion of fucosyl residues was ca. 30% in dry weight, as measured with the cystein method (Disches and Schettles, 1948). Sulfate dry weight was determined as from the sulfur % content by elementary analysis (CNRS, Service Central d'Analyse, Vernaison, France), whereas the uronic acid content was 2%, using the modified m-hydroxy-diphenyl-sulfuric acid method with mannuronic acid lactone as standard (Blumenkrantz and Asboe-Hansen, 1973). Alternatively, fucans were purified from *Pelvetia canaliculata*, *Ascophyllum nodosum* and *Fucus spiralis* as described previously (Mabeau et al., 1990). They contained 32%, 28% and 37% of fucose and 33%, 27% and 35% of sulfate, respectively.

Isolation of Fucoidan-Degrading Bacteria

Muds were collected from the water treatment facility of an alginate extraction plant (Danisco-Ingredients, Landerneau, Brittany, France) and precultivation of fucan-degrading bacteria was carried out as follows. Erlenmeyer flasks (100 ml) containing 10 ml of SeaWaterFucan (SWF) medium, consisting of sterile sea water supplemented with 0.2% (wt:vol) fucoidan from *P. canaliculata*, were inoculated with 400 µl of the muds and incubated aerobically at 22 °C and 200 rpm on a rotary shaker (New Brunswick Scientific). Aliquots were taken at daily intervals for 7 days and tested for the presence of fucoidanalytic activity as follows. The culture supernatant (200 µl) was added to 2 ml of acid albumin solution (3.26 g of sodium acetate, 4.56 ml of glacial acetic acid and 1.0 g of bovine serum albumin (Sigma) dissolved in 1 L of water, pH adjusted 3.72 to 3.78). Failure to develop a white turbidity due to the polysaccharide-albumin interaction indicated the presence of fucoidan-degrading activity (Kitamikado et al., 1990).

Active cultures were then plated onto Petri dishes with ZoBell medium (ZoBell, 1941) containing 0.3% (wt:vol) fucoidan and solidified by 0.7% (wt:vol) agar, and incubated at 22 °C for one week. The various colonies were then transferred aseptically into 10 ml of ZF medium, consisting of one volume of filtered sea water supplemented with 0.4% (wt:vol) fucoidan and one volume of ZoBell medium, incubated at 22 °C and assayed daily for fucoidanase activity (see below). The fucoidan-degrading isolates were submitted to repeated cultivation on ZF medium, eventually leading to the isolation of two different fucoidan-degrading strains, referred to as SW1 and SW5. They were stored at - 80 °C in ZF medium.

Biochemical and Molecular Characterization of SW5 Bacterial Strain

The morphological features of strain SW5 were investigated by microscopy (A100PL, Olympus BH-2) with cells in the exponential phase in ZoBell medium. To determine the respiratory type, bacteria were inoculated in Veillon tubes containing ZoBell medium solidified with 0.6% (wt:vol) agar. Oxygen was removed from the medium by boiling. To determine the oxidative or fermentative behaviour, bacteria were inoculated into a modified Hugh and Leifson O-F medium (Hugh and Leifson, 1953; Smibert and Krieg, 1981) containing 0.5% glucose. Oxidase activity was assayed with disks impregnated with dimethylparaphenylene diamine oxalate (Diagnostic Pasteur). Catalase activity was assayed by mixing one colony from a ZoBell agar plate with a drop of 10% (v/v) hydrogen peroxide. Production of flexirubin was assessed by flooding a 4-day plate culture with 5 N potassium hydroxide followed by the observation of changes in colony color from yellow to red or brown (Reichenbach et al., 1974). Other phenotypical tests were performed using API 20 NE strips (API System, Bio-Mérieux) and Biolog GN microplates (Micomer, France).

The 16S ribosomal DNA sequence of the strain SW5 was determined as previously described for another polysaccharide-degrading marine bacterium (Barbeyron et al., 2001). The 16S rDNA gene was amplified by polymerase chain reaction from the bacterial genomic DNA with the 8F primer (5'-AGAGTTTGATCCTGGCTCAG-3') (Hicks et al., 1992) and the 1492R primer (5'-GGTTACCTTGTTACGACTT-3') (Kane et al., 1993). PCR products were cloned in vector pCRII2.1 and both strands of 16S rDNA were sequenced using Texas-red labelled primers with a Vistra 725 sequencer. The sequence, 1241 bp in length, was analysed for similarities to rDNA sequences in the DNA database. The phylogenetic tree was constructed as previously described (Barbeyron et al., 2001).

Culture of SW5 and Detection of Fucoidanalytic Activity

SW5 was inoculated from a -80 °C stock into 10 ml of ZF medium in 250-ml Erlenmeyer flask. After incubation for 48 hours at 20 °C with vigorous rotary shaking, the culture was transferred to 250 ml of ZF medium in a 1-L Erlenmeyer flask and incubated under the same conditions. The culture was then centrifuged at 2,000 g for 15 min and the pellet was suspended in 10 ml of ZF medium and inoculated into 5 L of ZF medium, in a 6.6-L fermenter jar. The culture was maintained for 6 days at 20 °C (Bioflo 3000, New Brunswick Scientific). The pH was regulated at 7.85 with HCl and NaOH solutions (1 N) and aeration was 1 vvm (5 L/min), regulated at 70% by rotary shaking ranging from 200 to 800 rpm.

Attempts to quantitatively assay for fucoidanalytic activity by monitoring the release of fucoidan oligosaccharides by a conventional reducing sugar assay (Kidby and Davidson, 1973) were unsuccessful. Fucoidan-degrading activity was thus monitored by a carbohydrate-polyacrylamide gel electrophoresis (C-PAGE) assay of the release of anionic oligosaccharides, according to the procedure of Zablackis and Perez (1990). Briefly, 0.2% (wt:vol) fucoidan in 20 mM Tris/HCl buffer (pH 7.5) was incubated at room temperature with the enzyme and the products of hydrolysis (20 µl) were frozen at -20°C to stop the reaction. Samples were then mixed with the loading buffer (10% sucrose and 0.01% phenol red) and electrophoresed through a 6% (wt:vol) stacking and a 27% running, 1 mm-thick, polyacrylamide gel in 50 mM Tris/HCl, 2 mM EDTA buffer (pH 8.7) and stained with alcian blue followed by silver nitrate (Min and Cowman, 1986). Fucoidanase activity was detected by the occurrence of anionic oligosaccharide bands in the running gel. At the end of culture, the medium was centrifuged at 5,000 g for 30 min and the supernatant (5 L) was concentrated by ultrafiltration with an Aminco cassette (10-kDa cutoff). The cassette was washed with 1 L of fucoidanase buffer pH 7.5 containing 20 mM Tris-HCl, 50 mM NaCl, 5 mM MgCl₂ and 5 mM CaCl₂. The retentate (500 ml) was brought to 40% (wt:vol) saturation with (NH₄)₂SO₄ and precipitated overnight at 4 °C. The suspension was centrifuged at 12,000 g for 30 min and the

resulting supernatant was brought to 60% (wt:vol) saturation with $(\text{NH}_4)_2\text{SO}_4$. After precipitation overnight at 4°C, the precipitate was collected by centrifugation at 20,000 g for 1 hour, dissolved in ca 20 ml of fucoidanase buffer, dialyzed (10-kDA cutoff) against the fucoidanase buffer for 2 or 3 days at 4°C and stored at -20°C. The enzyme optimal pH and temperature were estimated using the C-PAGE assay and the ammonium sulfate fraction over the range 5.0-8.0 and 17-40°C, respectively.

Preparation of Fucoidan Oligosaccharides

An aliquot (8 ml) of the partially purified enzyme fraction (40-60% $(\text{NH}_4)_2\text{SO}_4$ fraction) was added to 1 L of 20 mM Tris-HCl buffer pH 7.5, 5 mM MgCl_2 , 5 mM CaCl_2 and 50 mM NaCl containing 5 g of fucoidan from *P. canaliculata* (FS28 fraction) and the mixture was incubated at 25 °C for 24 h. The hydrolysate was diluted in 20 L of distilled water and then ultrafiltered on a 10 kDa membrane (Millipore). The filtrate (17 L) was refiltered on a 500 Da membrane (Pall Filtron) with a 2 bars pressure. The filtrate (ca.15 L) was concentrated on a rotary evaporator and freeze-dried.

Aliquots (250 mg) of the resulting powder, referred to as OF fraction, were resuspended in distilled water (5 ml) and applied onto a DEAE Sepharose CL6B (Pharmacia, 1.6 x 10 cm, 1.1 ml/min) equilibrated with distilled water. Elution was first performed with water (330 ml), then with a linear gradient of 0-2 M NaCl (660 ml). Fractions (11 ml each) were assayed colorimetrically for total sugars (Thibault, 1979) and the carbohydrate-containing fractions were pooled, concentrated by evaporation, then applied on a BioGel P6 column (Biorad, 4.4 x 100 cm, 1 ml/min) equilibrated with 50 mM sodium nitrate, NaN_3 0.01 % (1500 ml), and eluted with the same buffer. Detection of carbohydrates was performed by refractometry (ERC 7510). Carbohydrate-containing peaks, referred to as peaks 1- 4, were collected, concentrated, desalted on a Sephadex G10 column (Pharmacia, 3.2 x 100 cm, 1.9 ml/min) and freeze-dried.

The neutral monosaccharide composition of the purified oligosaccharide fractions was analysed after hydrolysis and reduction by gas chromatography (Autosystem XL, Perkin Elmer) on an OV-225 column (30 m x 0.32 mm, 0.25 μm , DB 225J&W, USA) according to Blakeney et al. (1983). These fractions (250 $\mu\text{g}/\text{ml}$) were further analysed by high performance anion exchange chromatography (HPAEC) and conductivity detection, using a Dionex chromatograph in the configuration DX500 (Dionex, Sunnyvale, CA) and equipped with an anion micromembrane suppressor AMMS, a 50 μl -injection loop and an AS11A column (4 x 250 mm). Elution was performed at 0.5 ml/min by a linear gradient from 125 mM to 350 mM NaOH for 25 min and 350 mM NaOH for 5 minutes. Desalted fucoidan oligosaccharides were freeze-dried twice with deuterium oxide (99.9% D, Aldrich) and dissolved in 1 ml of D_2O (100% D, Aldrich) to replace exchangeable protons by deuterium. ^{13}C and ^1H NMR spectra were recorded at 25 °C with a Bruker ARX 400 spectrometer operating at 100.62 MHz and 400.13 MHz. Chemical shifts were assigned relative to internal acetone at 2.225 ppm and 31.45 ppm for ^1H and ^{13}C , respectively. Pulse sequences for 1D and 2D experiments were used as specified by the manufacturer.

RESULTS

Isolation and Characterization of a Fucoidan-degrading Bacterium

Muds collected from a water-treatment facility for the recycling of the effluents of an alginate extraction plant were searched for the occurrence of fucoidan-degrading bacteria. Of the 51 bacterial strains that were isolated for their ability to grow on fucoidan-enriched media, only two, referred to as SW1 and SW5, were active in the crude fucoidan-BSA enzymatic test (Kitamikado et al., 1990). These two strains were maintained on fucoidan-enriched ZoBell medium solidified with agar and then tested in liquid medium for their ability to degrade fucoidans. The SW5 strain, which displayed the higher extracellular fucoidanolytic activity, was retained for further studies.

SW5 appeared as an orange-pigmented, Gram-negative, rod-shaped, non-flagellated bacterium, unable to grow on the fresh-water Luria Bertani medium (Maniatis et al., 1982). Its biochemical characteristics are reported in Table 1. Strain SW5 was shown to be a strictly aerobic, seawater-requiring, chemo-organotrophic and heterotrophic organism, with an oxidative metabolism which used oxygen, but not nitrate, as the electron acceptor. The bacterium synthesized β -glucosidase and β -galactosidase but did not synthesize flexirubin, cytochrome oxydase or gelatinase and it was not able to degrade agarose and kappa-carrageenan. In contrast, iota-carrageenan was hydrolysed.

The 16S rDNA sequence of strain SW5 was found to be related, with sequence identities of 96% and 93%, respectively, to those of the strains BSD RB 42 (GenBank AY259505) and BSA CS 02 (GenBank AY259501), which belong to the Flavobacteriaceae family. In the phylogenetic analysis of its 16S rDNA gene, SW5 16S rDNA clusters with heterologous genes from unidentified strains. This group forms a clade with the psychrophilic bacteria *Psychroserpens burtonensis*, *Gelidibacter algens* and *G. mesophilus* (Figure 1).

Production of Fucoidanolytic Activity from SW5

To produce fucoidanolytic activity, the SW5 strain was grown in a fermenter in a fucoidan-enriched ZoBell medium. The generation time was of 12.5 hours at 20 °C and the culture medium reached an OD_{600nm} of ca 1.0 after 3 days of cultivation (Figure 2A). Fucoidanase activity was detected by C-PAGE in the culture supernatant as early as at the end of the exponential phase and was maintained during the stationary phase (Figure 2B). In contrast, intracellular fucoidanolytic activity was very low throughout the culture (data not shown).

The protein ammonium sulfate precipitate from the culture supernatant of SW5 extensively degraded the sulfated fucoidan from the fucoid alga *Pelvetia canaliculata* as well as those from two other fucoid algae, *Fucus spiralis* and *Ascophyllum nodosum* (Figure 3A). Based on the C-PAGE assay of the release of fucoidan oligomers, the optimal pH and temperature were estimated at 7.5 and 20-25°C, respectively. Although the digestion profiles depended on the fucoidan under investigation, they were all found to share two, low-molecular weight components with similar electrophoretic motilities (referred to as bands 3 and 4 on Figure 3A). The kinetics of the degradation of *P. canaliculata* fucoidan by the enzyme ammonium sulfate fraction is shown in Figure 3B. Oligofucoidans were detected as early as after 2.5 min of enzymatic digestion, in the form of several discrete bands, referred to as 3, 2, 8, 9 and 10. With the exception of band 2, the amount of the high molecular-weight fucoidan oligosaccharides rapidly decreased, while new oligofucoidan products appeared (bands 1, 5, 6 and 7). However, the terminal product (band 4) was always present in the reaction medium. No significant increase was observed in the OD_{235 nm} of the reaction mixture (data not shown).

Fractionation of Fucoidan Oligosaccharides

The crude fucoidan from the fucoid alga *Pelvetia canaliculata* (FS28) was hydrolysed to completion by the ammonium sulfate enzyme fraction and the products were ultrafiltered through a 500 Da membrane (OF fraction in Figure 3C), with a recovery of 66% in mass relative to the mass of the initial substrate. The OF fraction was fractionated by anion-exchange chromatography on a DEAE Sepharose CL6B column, resolving three major peaks (Figure 4A), which, on the basis of their levels in total sugars, represented ca. 87% of the initial oligosaccharide mixture. The main fucoidan oligosaccharide fraction, which was eluted at 0.6-1.6 M NaCl (fraction numbers 85-100) and represented 55% of OF in total carbohydrates, was further fractionated by gel filtration onto a Biogel P6 column (Figure 4B). It consisted of a mixture of unresolved fucan oligosaccharides eluting close to the column void volume as well as of four distinct low-molecular-mass fractions. Based on their sugar content, these latter oligosaccharides represented ca. 10% of OF fraction in total carbohydrates, i. e., about 7% of the mass of initial polysaccharide. Whereas the OF fraction contained, besides fucose (83.8 mol%), some xylose (5.9 mol%) and galactose (10.4 mol%) as well as traces of mannose and glucose, fucose was the only monosaccharide detected by gas chromatography in the purified oligosaccharide fractions 2, 3 and 4.

After desalting on a Sephadex G10 column, fractions 3 and 4 appeared homogenous upon C-PAGE analysis (Figure 3C). This was confirmed by high performance anion exchange chromatography (HPAEC) analysis. They were eluted essentially as single peaks, with retention times of 18.5 and 13.2 min, respectively (Figure 4C). In contrast, although fractions 1 and 2 appeared as homogenous by C-PAGE analysis (Figure 3C), they featured several peaks upon HPAEC chromatography, with their main components eluting at 19.9 and 20.9 min, respectively (Figure 4C). Discrete bands on C-PAGE gels may thus contain more than one fucoidan oligosaccharide, with various structures, but having the same electrophoretic motility.

NMR Analysis of Fucoidan Oligosaccharides

From its COSY and HMQC NMR spectra, oligosaccharide 4 featured chemical shifts assignable to five different α -L-fucopyranosyl residues (Bock et al., 1984), referred from **a** to **e** (Table 2). The ROESY spectrum (Figure 5) showed cross-peaks from H1 of residue **b** with H3 of residue **d** (**b1/d3**), from H1 of residue **d** to H4 of residue **c** (**d1/c4**), and from H1 of residue **c** to H3 of residues **a** and **e**. As previously reported for 3-linked sugars in the *galacto*-configuration (Lipkind et al., 1988; Bock and Thøgerson, 1982), cross-peaks were also observed between H1 of residue **b** and H4 of residue **d** and from H1 of **c** and H4 of residue **a**. The HMBC spectrum (data not shown) corroborated the presence of α -1 \rightarrow 3- and α -1 \rightarrow 4-linked fucose. The anomeric protons of **b**, **c**, **d**, at 5.40, 5.39 and 5.30 ppm, correlated with C3 of **d** and **a**, at 73.6 and 74.8 ppm, and with C4 of **c**, at 80.4 ppm (Table 2). Conversely, the anomeric carbons **b1**, **c1** and **d1** correlated with protons **d3**, **a3** and **c4**, respectively. From these results, oligosaccharide 4 consists of the following carbon backbone: α -L-Fucp-1 \rightarrow 3- α -L-Fucp-1 \rightarrow 4- α -L-Fucp-1 \rightarrow 3- α/β -L-Fucp.

Protons of α -L-fucopyranosyl units were identified from the COSY spectrum of oligosaccharide 3 and ^{13}C chemical shifts were deduced from the HMQC spectrum (Table 3). Chemical shifts were similar to those of oligosaccharide 4, with the exception of two groups of signals (**c'** and **d'**), indicating the presence of an additional disaccharide unit in this oligosaccharide. The actual sequence of oligosaccharide 3 was deduced from its HMBC spectrum (Figure 6) and by comparison with ^{13}C and ^1H of oligosaccharide 4. HMBC showed cross-peaks from H3 of residues **d**, **d'**, and **a** to C1 of residues **b**, **c** and **c'**, and from H4 of residues **c** and **c'** to C1 of residues **d** and **d'**. Cross-peaks between **c**, **c'** and **d**, **d'** corroborated the presence of α -1 \rightarrow 4-linkages. However, it was not possible to specify whether the **c'**-**d'** units were close to the reducing end or the non-reducing end in the oligosaccharide. The

proposed sequence for oligosaccharide 3 is α -L-Fucp-1 \rightarrow 3- α -L-Fucp-1 \rightarrow 4- α -L-Fucp-1 \rightarrow 3- α -L-Fucp-1 \rightarrow 4- α -L-Fucp-1 \rightarrow 3- α/β -L-Fucp.

The position of sulfate ester substituents in oligosaccharides 4 and 3 was deduced from the downfield shifts of H2 (0.7-0.9 ppm), which were consistent with those reported for the 2-O-sulfated fucans from the egg jelly layer echinoderms (Vilela-Silva et al., 2002), and of C2 (4-5 ppm), relative to unsubstituted α -L-Fucp residues (Table 4; Bock and Thøgerson, 1982). Sulfation on H3 of the 4-linked α -L-Fucp 2-sulfate was deduced from the downfield shift observed for H3 (~0.5 ppm) and C3 (~9 ppm) of this residue compared with the same signals in 4-linked α -L-Fucp-2-sulfate in *Strongylocentrotus droebachiensis* sea urchin fucan (Vilela-Silva et al., 2002). Furthermore, the proton and carbon chemical shifts for oligosaccharides 3 and 4 were in good agreement with those reported for alternating 3-linked α -L-Fucp-2-sulfate and 4-linked α -L-Fucp-2,3-disulfate in fucoidan oligomers from *Ascophyllum nodosum* (Chevolot et al., 2001).

Based on the attribution of the signals for proton of oligosaccharides 4 and 3, the mean degree of polymerization (DP) was estimated from the sum of the integral of H1 of the reducing-end sugar **a** (α -configuration) and H5 of the reducing-end sugar **e** (β -configuration) or that of H1 of **a** and (H5 + H3)/2 of **e**, in order to establish the integral value of one proton (**H**). The number of internal fucose residues (i.e., excluding non-reducing residue) (**B**) was estimated from the integral of signals between 5.25 and 5.40 ppm divided by the integral value for one proton (**B/H**). The DP value was then calculated as (**B/H**) + 1. Using this calculation, oligosaccharides 4, 3 and 2, have a mean DP of 3.9, 5.9 and 6.8, respectively. The proton NMR spectrum of oligosaccharide 2 indicated that the mixture did not contain oligosaccharides in the series homologous to oligosaccharides 4 and 3.

DISCUSSION

The SW5 Strain Is a Fucoidanolytic Flavobacteriaceae

We here report the isolation from the effluent treatment facility of an alginate plant of a novel bacterium referred to as SW5, which exhibits fucoidanase activity when grown in the presence of sulfated fucans from the fucoid *P. canaliculata* (Figure 2). The occurrence of such a fucoidanolytic bacterium in this habitat is likely to be accounted for by the presence of significant amounts of high molecular-weight sulfated fucoidans in these effluents, as by-

products of the alginate-manufacturing process from fucoidan-containing brown algae such as *Laminaria digitata* and *Ascophyllum nodosum* (Fleury and Lahaye, 1993).

Based on its physiological and biochemical characteristics (Table 1) as well as on the phylogenetic analysis of its 16S rDNA gene (Figure 1), SW5 belongs in the order of *Flavobacteriales* and in the family of *Flavobacteriaceae*. However, it is only remotely related to the only other fucoidan-degrading *Flavobacteriaceae* characterized so far, *Fucobacter marina* (Sakai et al., 2002). It is closely related to, but distinct from, the genera *Gelidibacter* and *Psychroserpens*, which were isolated from an Antarctic sea-ice habitat (Bowman et al., 1997). All of these bacterial strains require the presence of salt for optimal growth, suggesting that they are marine in origin. Therefore, SW5 may have been a member of the natural epiflora of brown algae carried throughout the alginate-extraction process and seeded into the water-treatment muds, where it contributes to the depolymerisation of the sulfated fucoidans in the effluents. Alternatively, since the water treatment facility is located close to an estuary, SW5 may have been naturally introduced from the nearby estuarine bacterial communities.

SW5 Secretes Fucoidan-hydrolase Activity

The bacterium extracellular enzymes were subjected to fractionation by ammonium sulfate precipitation. The sulfated fucoidan from *Pelvetia canaliculata* was extensively degraded by the fraction of SW5 extracellular enzymes which precipitates between 40-60% ammonium sulfate saturation (Figure 3A). Fucoidan was first cleaved into a number of high molecular-weight fucoidan oligosaccharides (Figure 3B), indicating an endo-acting mechanism.

As shown by DEAE chromatography and gel filtration followed by C-PAGE, HPAEC, and NMR analysis, the hydrolysis products contained a tetrasaccharide, α -L-Fucp-2,3-(diOSO₃)-1→3- α -L-Fucp-2-(OSO₃)-1→4- α -L-Fucp-2,3-(diOSO₃)-1→3- α -L-Fucp-2-(OSO₃), and a hexasaccharide, α -L-Fucp-2,3-(diOSO₃)-1→3- α -L-Fucp-2-(OSO₃)-1→4- α -L-Fucp-2,3-(diOSO₃)-1→3- α -L-Fucp-2-(OSO₃), as well as higher molecular-weight fucoidan oligosaccharides. These findings are similar to those reported from a low molecular weight fucan fraction of *A. nodosum* (Chevolot et al., 2001), which harbors repetitions of the disaccharide 1→3- α -L-Fucp-2(OSO₃)-1→4- α -L-Fucp-2,3(diOSO₃), referred to as B-A and where the B fucosyl residue can more or less regularly bear a substituent at C-4 (sulfate, xylose, fucose or sulfated fucose). Similar fucan oligosaccharides with repeating units were also recognized in the fucoidans of *F. vesiculosus* (Chevolot et al., 2001), *F. evanescens* (Bilan et al., 2002) and *F. distichus* (Bilan et al., 2004).

Since SW5 extracellular enzymes can degrade a variety of sulfated fucoidans from fucoid algae (Figure 3A), it is thus likely that the bacterium attacks fucoidan by cleaving linkages within blocks of alternating α -1 \rightarrow 3 and α -1 \rightarrow 4 L-sulfated fucose residues. This conclusion is only preliminary, however, since SW5 is likely to secrete more than one fucoidan-degrading activity, as suggested by the presence of different low molecular-weight fucoidans at completion of the hydrolysis of *P. canaliculata* fucoidan (Figure 3B). Altogether it appears that SW5 features a fucoidan-degrading enzyme machinery distinct from that of *Fucobacter marina*, which exhibits a fucoglucuronomannan lyase (Sakai et al., 2002; Sakai et al. 2003a). Other fucoidan-degrading activities were reported from intracellular extracts of *Fucophilus fucoidanolyticus*, namely a fucoidan hydrolase, an α -D-glucuronidase and a fucose-deacetylase (Sakai et al., 2003b, 2003c). The isolation of SW5 provides an alternative source for the production of extracellular enzymes to degrade fucoidan.

In conclusion, on the basis of its capacity to degrade brown algal sulfated fucoidans, we have isolated a flavobacteriacean strain, referred to as SW5, which secretes fucoidan-hydrolase activity in its culture medium. The end products of the degradation of *Pelvetia canaliculata* fucoidan comprised a tetrasaccharide, made of the repetition of two units of α -1 \rightarrow 3-L-fucopyranose-2-sulfate, and α -1 \rightarrow 4-L-fucopyranose-2,3-disulfate, with the 3-linked residues at the non-reducing end, and a hexasaccharide in the same homologous series. SW5 has now been successfully maintained for more than three years in our laboratory, where it is routinely used for the preparation of fucoidan-degrading enzymes. The strain was patented (Descamps et al., 1998) and deposited in the DSMZ bacterial collection, under the accession number 12171. SW5 fucoidan endo-hydrolase(s) should prove a useful tool for the structural analyses of fucoidan as well as for the preparation of biologically active low molecular weight fucans.

ACKNOWLEDGMENTS

VD and SC were the recipients of a PhD fellowship co-funded by the ANRT, whose help is gratefully acknowledged. The authors also are grateful to the late J.-C. Pasquier, for access to the water treatment facilities of Danisco Cultor Landerneau SA, and to Erwan Corre, for his expertise in phylogenetic analyses.

REFERENCES

- Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., and Potin, P. (2001). *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of Flavobacteriaceae isolated from a red alga, and classification of [Cytophaga] *uliginosa* (ZoBell and Upham, 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov.. *Int J Syst Evol Microb* 51:985-997.

- Béress, A., Wassermann, O., Bruhn, T., Béress, L. (1993). A new procedure for the isolation of anti-HIV compounds (polysaccharides and polyphenols) from the marine alga *Fucus vesiculosus*. *J. Nat. Prod* 56: 478-488.
- Berteau, O., McCort, I., Goasdoue, N., Tissot, B., and Daniel, R. (2002). Characterization of a new alpha-L-fucosidase isolated from the marine mollusk *Pecten maximus* that catalyzes the hydrolysis of alpha-L-fucose from algal fucoidan (*Ascophyllum nodosum*). *Glycobiology* 12:273-282.
- Bilan, M.I., Grachev, A.A., Ustuzhanina, N.E., Shashkov, A.S., Nifantiev, N.E. and Usov, A.I. (2002). Structure of a fucoidan from the brown seaweed *Fucus evanescens* C. Ag. *Carbohydr Res* 337:719-730.
- Bilan, M.I., Grachev, A.A., Ustuzhanina, N.E., Shashkov, A.S., Nifantiev, N.E. and Usov, A.I. (2004). A highly regular fraction of a fucoidan from the brown seaweed *Fucus distichus* L. *Carbohydr Res* 339:511-517.
- Blakeney, A.B., Harris, P.J., Henry, R.J. and Stone, B.A. (1983). A simple and rapid preparation of alditol acetates for monosaccharide analysis. *Carbohydr Res* 113:291-299.
- Blumenkrantz, N., and Asboe-Hansen, G. (1973). New method for quantitative determination of uronic acids. *Anal Biochem* 54:484-486.
- Bock, K. and Thøgersen, H. (1982). Nuclear magnetic resonance spectroscopy in the study of mono- and oligosaccharides. *Ann Rep NMR Spect* 13:2-57.
- Bock, K., Pedersen, C. and Pedersen, H. (1984). Carbon-13 nuclear magnetic resonance data for oligosaccharides. *Adv Carbohydr Chem Biochem* 42:193-225.
- Bowman, J.P., Mc Cammon, S.A., Brown, J.L., Nichols, P.D., and Mc Meekin, T.A. (1997). *Psychroserpens burtonensis* gen.nov., sp. nov., and *Gelidibacter algens* gen. nov., sp. nov., psychrophilic bacteria isolated from antarctic lacustrine and sea ice habitats. *Int J Syst Bact* 54:670-677.
- Chevotot, L., Mulloy, B., Ratiskol, J., Foucault, A. and Collic-Jouault, S. (2001). A disaccharide repeat unit is the major structure in fucoidans from two species of brown algae. *Carbohydr Res* 330:529-535.
- Dalton, J.P., Hudson, D., Adams, J.H. and Miller, L.H. (1991). Blocking of the receptor-mediated invasion of erythrocytes by *Plasmodium knowlesi* malaria with sulfated polysaccharides and glycosaminoglycans. *Eur. J. Biochem.* 195: 789-794.
- Daniel, R., Berteau, O., Jozefonvicz, J., and Goasdoue, N. (1999). Degradation of algal (*Ascophyllum nodosum*) fucoidan by an enzymatic activity contained in digestive glands of the marine mollusc *Pecten maximus*. *Carbohydr Res* 322:291-297.
- Descamps, V., Klarzynski, O., Barbeyron, T., Kloareg, B., Fritig, B., Joubert, J.-M., Plesse, B., and Yvin, J.-C. (1998). "Fuco-oligosaccharides, enzyme pour leur préparation à partir des fucanes, bactérie productrice de l'enzyme et applications des fuco-oligosaccharides à la protection des plantes". Patent n° FR9811756, n° WOFR9902243.
- Disches, Z., and Schettles, L.B. (1948). A specific color reaction of methyl-pentoses and a spectrophotometric micromethod for their determination. *J Biol Chem* 175:595-603.
- Fleury, N., and Lahaye, M. (1993). Studies on by-products from the industrial extraction of alginate. *J Appl Phycol* 5:605-614.
- Furukawa, S., Fujikawa, T., Koga, D., and Ide, A. (1992). Production of fucoidan-degrading enzymes, fucoidanase and fucoidan sulfatase by *Vibrio* sp. N-5. *Nippon Suisan Gakkaishi* 58:1499-1503.
- Goodner, B.W. and Quatrano, R.S. (1993). *Fucus* embryogenesis: a model to study the establishment of polarity. *Plant Cell* 5: 1471-1481.
- Hicks, R.E., Amann, R.I., and Stahl, D.A. (1992). Dual staining of natural bacterioplankton with 4',6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. *Appl Environ Microbiol* 58:2158-2163.
- Hoshino, T., Hayashi, T., Hayashi, J., Lee, J.B. and Sankawa, U. (1998). An antivirally active sulfated polysaccharide from *Sargassum horneri* (Turner) C. Agardh. *Biol. Pharm. Bull.* 21: 730-734.
- Hugh, R., and Leifson, E. (1953). The taxonomic significance of fermentative versus oxydative metabolism of carbohydrates by various Gram-negative bacteria. *J Bacteriol* 66:24-26.
- Kane, M.D., Poulsen, L.K., and Stahl, D.A. (1993). Monitoring the enrichment and isolation of sulfate-reducing bacteria by using oligonucleotide hybridization probes designed from environmentally derived 16S rRNA sequences. *Appl Environ Microbiol* 59:682-686.
- Kidby, D.K., and Davidson, D.J. (1973). A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal Biochem* 55:321-325.
- Kitamikado, M., Yamaguchi, K., Tseng, C.H., and Okabe, B. (1990). Method designed to detect alginate-degrading bacteria. *Appl Environ Microbiol* 9:2939-2940.
- Kitamura, K., Matsuo, M., and Yasui, T. (1992). Enzymic degradation of fucoidan by fucoidanase from the hepatopancreas of *Patinopecten yessoensis*. *Biosc Biotechnol Biochem* 56:490-494.

- Klarzynski, O., Descamps, V., Plesse, B., Yvin, J.-C., Kloareg, B., and Fritig, B. (2003). Sulfated fucan oligosaccharides elicit defense responses in tobacco and local and systemic resistance against tobacco mosaic virus. *Mol Plant-Microbe Interact* 16:115-122.
- Kloareg, B. (1984). Isolation and analysis of cell walls of brown algae *Pelvetia canaliculata* and *Ascophyllum nodosum*. *Physiol Veg* 22:47-56.
- Kloareg, B., and Quatrano, R.S. (1988). Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanography Marine Biology Annual review* 26:259-315.
- Kropf, D.L., Kloareg, B. and Quatrano, R.S. (1988). Cell wall is required for fixation of the embryonic axis in *Fucus zygotes*. *Science* 239:187-189.
- Ley, K., Linneman, G., Meinen, M., Stoolman, L.M. and Gaehtgens, P. (1993). Fucoidin, but not yeast polyphosphomannan PPME, inhibits leukocyte rolling in venules of the rat mesentery. *Blood* 81: 177-185.
- Linneman, G., Reinhart, K., Parade, U., Philipp, A., Pfister, W., Straube, E. and Karzai, W. (2000). The effects of inhibiting leukocyte migration with fucoidin in a rat peritonitis model. *Intensive Care Med.* 26: 1540-1546.
- Lipkind, G.M., Shashkov, A.S., Mamyan, S.S. and Kochetkov, N.K. (1988). The nuclear overhauser effect and structural factors determining the conformations of disaccharide glycosides. *Carbohydr Res* 181:1-12.
- Logeart, D., Prigent-Richard, S., Jozefonvicz, J. and Letourneur, D. (1997). Fucans, sulfated polysaccharides extracted from brown seaweeds, inhibit vascular smooth muscle cell proliferation. I. Comparison with heparin for antiproliferative activity, binding and internalization. *Eur. J. Cell Biol.* 74: 376-384.
- Mabeau, S., and Kloareg, B. (1987). Isolation and analysis of the cell walls of brown algae: *Fucus spiralis*, *F. ceranoides*, *F. vesiculosus*, *F. serratus*, *Bifurcaria bifurcata* and *Laminaria digitata*. *J Exp Bot* 38:573-1580.
- Mabeau, S., Kloareg, B., and Joseleau, J.-P. (1990). Fractionation and analysis of fucans from brown algae. *Phytochemistry* 29:2441-2445.
- Mahony, M.C., Oehninger, S., Clark, G.F., Acosta, A.A. and Hodgen, G.D. (1991). Fucoidin inhibits the zona pellucida-induced acrosome reaction in human spermatozoa. *Contraception* 44:657-665.
- Mahony, M.C., Clark, G.F., Oehninger, S., Acosta, A.A. and Hodgen, G.D. (1993). Fucoidin binding activity and its localization on human spermatozoa. *Contraception* 48:277-288.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Mayer, A.M.S., Diaz, A., Pesce, A., Criscuolo, M., Groisman, J.F. and de Lederkremer, R.M. (1987). Biological activity in *Macrocystis pirifera* from Argentina : sodium alginate, fucoidan and laminaran. III. Antiviral activity. *Hydrobiologia* 151/152: 497-500.
- McCaffrey, T.A., Falcone, D.J., Borth, W., Brayton, C.F. and Weksler, B.B. (1992). Fucoidan is a non-coagulant inhibitor of intimal hyperplasia. *Biochem. Biophys. Res. Commun.* 184: 773-781.
- Millet, J., Collic-Jouault, S., Mauray, S., Theveniaux, J., Sternberg, C., Boisson-Vidal, C. and Fischer, A.M. (1999). Antithrombotic and anticoagulant activities of a low molecular weight fucoidan by the subcutaneous route. *Thromb Haemostasis* 81:391-395.
- Min, H., and Cowman, M.K. (1986). Combined alcian blue and silver staining of glycosaminoglycans in polyacrylamide gels : application to electrophoretic analysis of molecular weight distribution. *Anal Biochem* 155:275-285.
- Morigana, T., Araki, T., Ito, M., and Kitamikado, M. (1981). A search for fucoidan-degrading bacteria in coastal sea environments of Japan. *Bull Japanese Soc of Scientific Fisheries* 47:621-625.
- Ostergaard, C., Yieng-Kow ,R.V., Benfield, T., Fridodt-Moller, N., Espersen, F. and Lundgren, J.D. (2000). Inhibition of leukocyte entry into the brain by the selectin blocker fucoidin decreases interleukin-1 (IL-1) levels but increases IL-8 levels in cerebrospinal fluid during experimental pneumococcal meningitis in rabbits. *Infect. Immun.* 68: 3153-3157.
- Pereira, M.S., Mulloy, B. and Mourão, P.A.S. (1999). Structure and anticoagulant activity of sulfated fucans. *J Biol Chem* 274:7656-7667.
- Reichenbach, H., Kleinig, H. and Achenbach, H. (1974). The pigment of *Flexibacter elegans*: novel and chemosystematically useful compounds. *Archiv Microbiol* 101:131-144.
- Riou, D., Collic-Jouault, S., Pinczon du sel, D., Bôsch, S., Siavoshian, S., Le Bert, V., Tomasoni, C., Sinquin, C., Durand, P. and Roussakis, C. (1996). Antitumor and antiproliferative effects of a fucan extracted from *Ascophyllum nodosum* against a non-small-cell bronchopulmonary carcinoma line. *Anticancer Res.* 16: 1213-1218.
- Sakai, T., Kimura, H., and Kato, I. (2002). A marine strain of Flavobacteriaceae utilizes brown seaweed fucoidan. *Mar Biotechnol* 4:399-405.
- Sakai, T., Kimura, H., Kojima, K., Shimanaka, K., Ikai, K., and Kato, I. (2003a). Marine bacterial sulfated fucoglucuronomannan (SFGM) lyase digests brown algal SFGM into trisaccharides. *Mar Biotechnol* 5:70-78.

- Sakai, T., Ishizuka, K., and Kato, I. (2003b). Isolation and characterization of a fucoidan-degrading marine bacterium. *Mar Biotechnol* 5:409-416.
- Sakai, T., Ishizuka, K., Shimanaka, K., Ikai, K. and Kato, I. (2003c). Structures of oligosaccharides derived from *Cladosiphon okamuranus* fucoidan by digestion with marine bacterial enzymes. *Mar Biotechnol* 5:536-544.
- Smibert, R.M., and Krieg, N.R. (1981). General characterization. In *Manual of methods for general bacteriology*, pp. 409-443. Edited by P. Gerhardt, R.G.E. Murray, R.N. Costilow, E.W. Nester, W.A. Wood, N.R. Krieg, and G.B. Phillips. Washington, DC 20006, USA: American Society for Microbiology.
- Thanassi, N.M., and Nakada, H.I. (1967). Enzymic degradation of fucoidan by enzymes from the hepatopancreas of abalone, *Haliotis* species. *Arch Biochem Biophys* 118:172-177.
- Thibault, J.F. (1979). Automatisation du dosage des substances pectiques par la méthode du méta-hydroxyphényl. *Lebens Wiss Technol* 12:247-251.
- Tillmans, J., and Phillipi, K. (1929). The carbohydrate content of the important protein of foodstuff on colorimetric procedure for the determination of nitrogen free sugars in protein. *Biochem Z* 215:36-60.
- Thorlacius, H., Vollmar, B., Seyfert, U.T., Vestweber, D. and Menger, M.D. (2000). The polysaccharide fucoidan inhibits microvascular thrombus formation independently from P- and L-selectin function in vivo. *Eur J Clin Invest* 30:804-810.
- Uchida, M. (1995). Enzyme activities of marine bacteria involved in *Laminaria*-thallus decomposition and the resulting sugar release. *Marine Biology* 123:639-644.
- Vilela-Silva, A.C.E.S., Castro, M.O., Valente, A.P., Biermann, C.H. and Mourao, P.A.S. (2002). Sulfated fucans form the egg jellies of the closely related sea urchins *Strongylocentrotus droebachiensis* and *Strongylocentrotus pallidus* ensure species-specific fertilization. *J Biol Chem* 277:379-387.
- Wikström, T., Braide, M., Bagge, U. and Risberg, B. (1995). Leukocyte margination during hemorrhagic shock correlates to preshock margination and is reduced by fucoidin. *Shock* 3: 40-45.
- Zablackis, E., and Perez, J. (1990). A partially pyruvated carrageenan from Hawaiian *Grateloupia filicina* (Cryptonemiales, Rhodophyta). *Bot Mar* 33:273-276.
- Zhuang, C., Itoh, H., Mizuno, T. and Ito, H. (1995). Antitumor active fucoidan from the brown seaweed, *Umitoranoo* (*Sargassum thunbergii*). *Biosci. Biotech. Biochem.* 59: 563-567.
- ZoBell, C.E. (1941). Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *J Mar Res* 4:41-75.

LEGENDS OF FIGURES

Figure 1. Phylogenetic relationships of strain SW5 to some marine representatives of the family *Flavobacteriaceae*. Square brackets indicate a generically misnamed taxon and quotation marks indicate a name not yet validated. Accession numbers of 16S rDNA sequences are given in brackets. The topology shown is the tree obtained using the neighbour-joining method (Jukes and Cantor distance correction). Numbers at the nodes refer to the bootstrap values (100 replicates) as obtained in distance, maximum-likelihood, and maximum parsimony analyses, respectively, while dashes instead of numbers indicate that the node was not observed in the corresponding analysis. The scale bar represents the expected number of changes per sequence position.

Figure 2. Growth of the SW5 strain (A), as seen from the culture optical density measured at 600 nm. SW5 was grown in a 5-L fermenter in the presence of ZoBell medium supplemented with fucoidan from *Pelvetia canaliculata*. (B) Monitoring of the fucoidanolytic activity in the culture supernatant, using C-PAGE analysis. Culture aliquots (2 ml) were centrifuged, supernatant aliquots (20 µl) were incubated for 3 h with 100 µl of 0.2% (wt:vol) fucoidan from *Pelvetia canaliculata*, and the hydrolysates (5 µl aliquots) were analyzed by C-PAGE. Numbers above the lanes refer to the culture duration (in days), with T corresponding to unhydrolysed fucoidan.

Figure 3. (A) C-PAGE electrophoresis of the hydrolysates by SW5 ammonium sulfate fucoidanase fraction of various brown algal sulfated fucans (0.5%). An analytical-grade fraction of *Pelvetia canaliculata* fucoidan (Pc1), a purified fraction from the same species (Pc2); and fucoidans purified from *Fucus spiralis* (Fs) and *Ascophyllum nodosum* (An). The

control (T) consisted of the Pc1 fraction incubated with boiled enzyme. Bands numbers were assigned relative to their elution order upon gel filtration (see Figure 4B) and by comparison with Figure 3C. (B) C-PAGE analysis of the hydrolysis kinetics of *P. canaliculata* fucoidan by SW5 ammonium sulfate fucoidanase fraction. The fucoidan (400 μ l, 0.2%) was incubated with the ammonium sulfate enzyme fraction (40 μ l) and aliquots (20 μ l) of this mix were boiled and electrophoresed as above. Bands numbers were assigned by comparison to figure 3A and 3C. Numbers at the top of the gel refer to the duration of hydrolysis (in min and h). (C) C-PAGE analysis of purified low molecular weight fucoidans. *Pelvetia canaliculata* fucoidan (FS28) was hydrolysed by SW5 fucanase and the products were recovered by ultrafiltration through a 500 Da membrane (OF fraction). This fraction were further purified by chromatography on DEAE Sepharose CL6B then on Biogel P6 and carbohydrate-containing peaks 1, 2, 3 and 4 (see Figure 4B) were submitted to C-PAGE (lanes 1 - 4). Arrows indicate the corresponding oligosaccharide bands in the OF fraction.

Figure 4. Purification of the end-products of the hydrolysis of *Pelvetia canaliculata* fucoidan by SW5 ammonium sulfate fucanase fraction. (A) Fractionation of the hydrolysate (OF fraction) from *P. canaliculata* fucan on DEAE Sepharose. The OF fraction (250 mg) was chromatographed on DEAE Sepharose CL6B with a 0-2 M NaCl gradient, eluted fractions were assayed for total sugars, and carbohydrate-containing fractions (fractions 85 to 100) were collected. (B) The fractions 85 to 100 were pooled and then chromatographed on Biogel P6 using 50 mM NaNO₃ as eluent. The four peaks resolved as fractions 1 – 4 were collected. (C) HPAEC elution profiles of peaks 1 to 4 from fractions 85 to 100 upon Biogel P6 chromatography. Fractions were desalted on Sephadex G10 and injected at 250 μ g.ml⁻¹ on an AS11A anion-exchange column, using a NaOH linear gradient and conductivity detection.

Figure 5. Expansion of the double ¹H-¹H ROESY spectrum of oligosaccharide 4, at 400 MHz, 25 °C, in D₂O. Cross-peaks between the various sugar residues are indicated as, for example: **d1/c4** for H1 of unit **d** with H4 of unit **c**. 1024 experiments of 2K data points and 8 transients each with a recycling time of 1.16 s were recorded and transformed on 2 x 2 K data points with unshifted sine bell multiplication in both dimensions. The ROESY spinlock pulse duration was 700 ms.

Figure 6. Expansion of the double ¹H-¹³C HMBC spectrum of oligosaccharide 3, at 400 MHz, 25 °C, in D₂O. Cross-peaks between the various sugar residues are indicated using the same nomenclature as in Figure 5. 512 experiments of 2K data points and 64 transients, each with a recycling time of 1.2 s, were recorded and transformed on 1 x 2 K data points on the F1 and F2 dimensions, respectively after exponential multiplication in F1 (line broadening of 0.1 Hz) and unshifted sine bell multiplication in F2. The evolution time for long range interactions was set at 75 ms.

Table 1. Characteristics of the Fucoidan-Utilizing Bacterium, Strain SW5

Reduction of nitrate to nitrite	-	D-Alanine	-	Acetic acid	-
Indole production	-	L-Alanine	+	cis-Aconitic acid	-
Flexirubin production	-	L-asparagine	+	Adipic acid	-
Cytochrome oxidase	-	L-Aspartic acid	+	Capric acid	-
Catalase	+	L-Glutamic acid	+	Citric acid	-
Arginine dihydrolase	-	L-Pyroglutamic acid	-	Formic acid	-
Urease	-	L-Histidine	-	Malic acid	-
β -Galactosidase (PNPG*)	+	L-Leucine	-	Malonic acid	-
Aesculin hydrolase	+	L-Phenylalanine	-	D/L-Lactic acid	+
□□□□□□□□□□□□□□	-	L-Proline	+	Itaconic acid	-
Tween 40	-	Hydroxy-L-proline	-	Propionic acid	-
Tween 80	-	D-Serine	-	Quinic acid	-
N-Acetyl-D-galactosamine	-	L-Serine	-	Succinic acid	-
N-Acetyl-D-glucosamine	-	L-Threonine	+	Bromosuccinic acid	-
Adonitol	-	Putrescine	-	D-Galactonic acid	-
D-Glucose	+	L-Ornithine	+	D-Galacturonic acid	+
D-Galactose	+	D/L-Carnitine	-	D-Gluconic acid	-
D-Fructose	+	Glycyl-aspartic acid	+	D-Glucuronic acid	+
D-Fucose	+	Glycyl-glutamic acid	+	D-Glucosaminic acid	-
L-Rhamnose	-	L-Alanyl-glycine	+	γ -Aminobutyric acid	-
D-Melibiose	-	D-Arabitol	-	α -Hydroxybutyric acid	-
Maltose	+	L-Erythritol	-	β -Hydroxybutyric acid	-
D- Psicose	-	m-Inositol	-	γ -Hydroxybutyric acid	-
D-Mannose	+	D-Mannitol	-	α -Ketobutyric acid	-
L-Arabinose	-	D-Sorbitol	-	α -Ketoglutaric acid	-
D-Trehalose	-	Xylitol	-	α -Ketovaleric acid	-
Sucrose	-	D/L-glycerophosphate	-	D-saccharic acid	-
D-Raffinose	-	Glucose-1-phosphate	+	Sebacic acid	-
Gentobiose	+	Glucose-6-phosphate	+	Succinamic acid	-
Cellobiose	+	β -Methyl-D-glucoside	+	Urocanic acid	+
D-Lactose	+	Inosine	-	Methyl pyruvate	-
Lactulose	-	Thymidine	-	Monomethylsuccinate	+
Turanose	+	Uridine	-	Alininamide	-
Agarose	-	2-Aminoethanol	-	Glucuronamide	-
κ -Carrageenan	-	2,3-butanediol	-	p-hydroxyphenylacetic acid-	
ι -Carrageenan	+	Glycogen	+		
Dextrin	+	α -Cyclodextrin	+		

* : Paranitrophenol galactoside

Table 2. Chemical Shifts (ppm) for the Fucan Oligosaccharide 4

residue	H1 ^a	H2	H3	H4	H5	H6
a (α)	5.51	4.55	4.08	4.11	4.25	1.25
e (β)	4.73	4.35	3.91	4.06	3.83	1.30
c	5.39	4.66	4.78	4.30	4.57	1.42
d	5.30	4.60	4.22	4.14	4.46	1.31
b	5.40	4.58	4.77	4.24	4.61	1.26
	C1 ^b	C2	C3	C4	C5	C6
a (α)	91.68	74.53	74.84	70.28	67.08	16.60
e (β)	96.32	78.34	77.55	69.28	71.76	~16.6
c	96.18	73.61	75.05	80.38	69.03	16.78
d	99.86	76.53	73.61	70.56	68.20	16.46
b	95.81	73.52	75.05	71.76	67.693	16.38

^aThe ¹H spectrum was recorded at 400 MHz, 298°K, chemical shifts are relative to internal $\delta^1\text{H}$ acetone 2.225. ^bThe ¹³C spectrum was recorded at 400 MHz, 298°K, chemical shifts are relative to internal $\delta^{13}\text{C}$ acetone 31.45.

Table 3. Chemical Shifts (ppm) for the Fucan Oligosaccharide 3

residue	H1 ^a	H2	H3	H4	H5	H6
a (α)	5.49	4.53	4.07	4.10	4.23	1.24
e (β)	4.72	4.33	3.90	4.05	3.82	1.28
c'	5.39	4.64	4.76	4.28	4.56	1.40
d'	5.30	4.58	4.22	4.13	4.39	1.30
c	5.40	4.65	4.79	4.29	4.62	1.40
d	5.29	4.58	4.22	4.13	4.43	1.30
b	5.39	4.57	4.73	4.22	4.59	1.25
	C1 ^b	C2	C3	C4	C5	C6
a (α)	91.64	74.57	74.57	70.16	67.06	16.56
e (β) [*]	-	-	-	-	-	-
c'	95.73	73.57	75.06	80.87	68.98	16.76
d'	100.04	74.57	74.57	70.53	68.15	16.43
c	95.96	73.57	75.06	80.58	68.98	16.76
d	99.87	74.57	74.57	70.53	68.15	16.43
b	95.96	73.57	76.18	71.73	67.68	16.43

^aThe ¹H spectrum was recorded at 400 MHz, 298°K, chemical shifts are relative to internal $\delta^1\text{H}$ acetone 2.225,

^bThe ¹³C spectrum was recorded at 400 MHz, 298°K, chemical shifts are relative to internal $\delta^{13}\text{C}$ acetone 31.45,

* not determined.

Table 4. Comparison of the Chemical Shifts of Various Fucans.

	H1	H2	H3	H4	H5	H6
<i>Desulfated sea cucumber fucan</i> ^a	5.08	3.92	4.00	nd	4.27	1.21
<i>native sea cucumber fucan, residue A</i> ^a	5.40	4.58	4.39	4.91	4.37	1.25
<i>α-L-Fucose</i>	5.2	3.8	3.8	3.8	4.2	1.2
<i>α-L-Fucose residues of fucan oligosaccharides 3 and 4</i>	4.7-5.5	4.4-4.7	3.9-4.0	4.0-4.1	3.8-4.2	1.2-1.3

^a data from *Ludwigothurea grisea* (59)

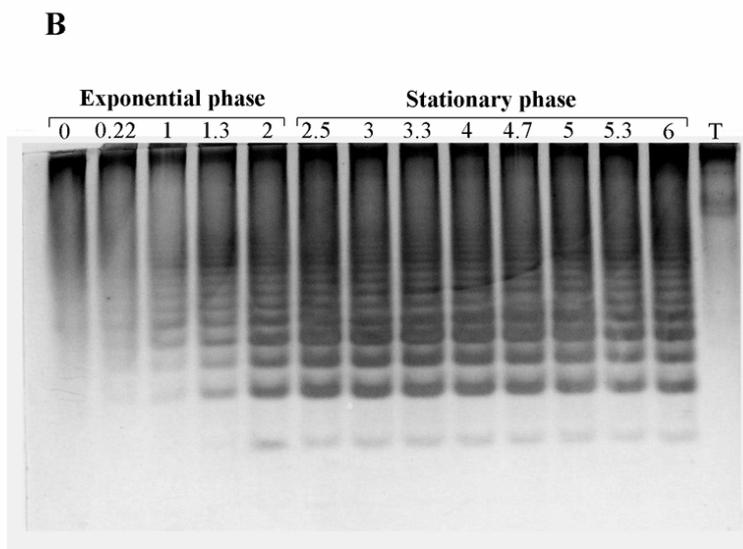
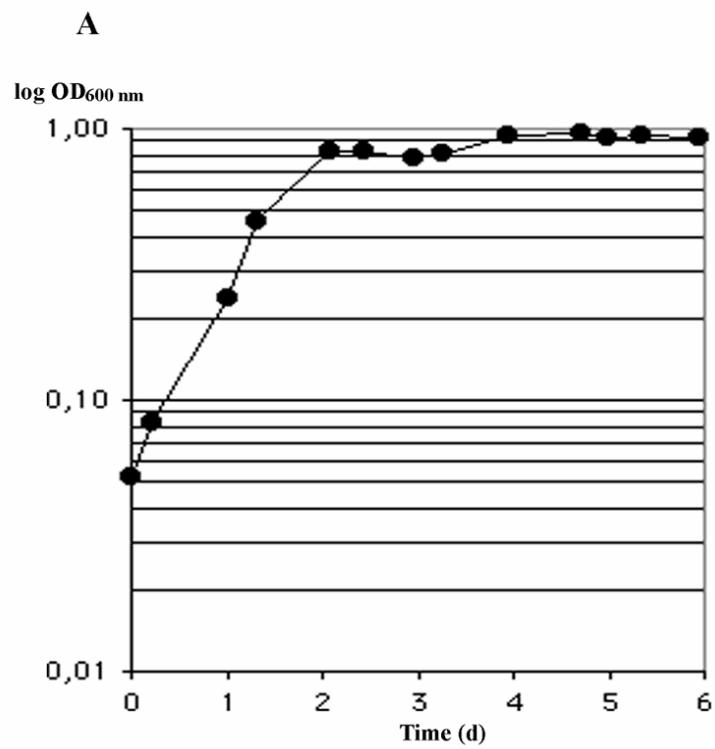


Fig. 2

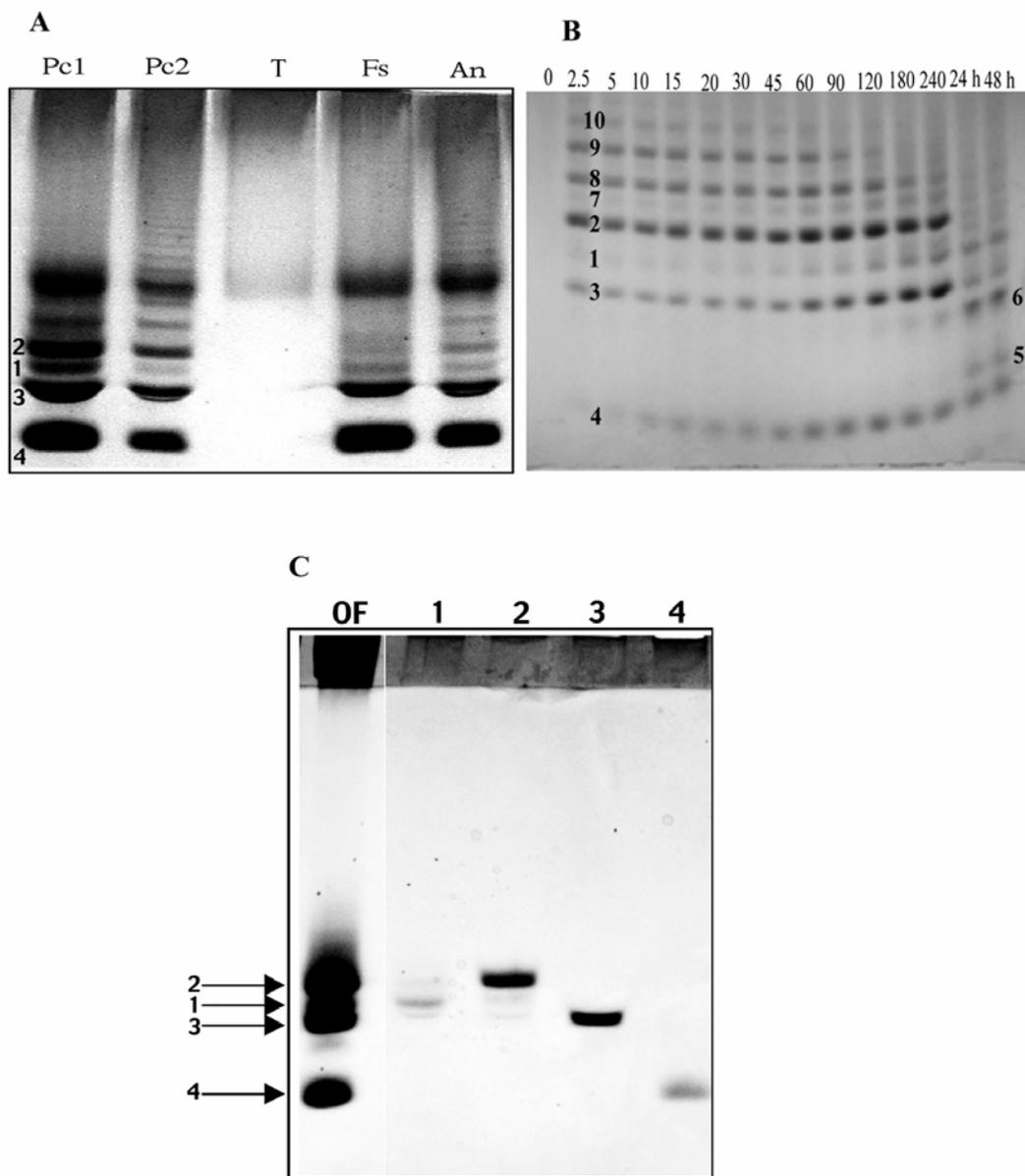


Fig. 3

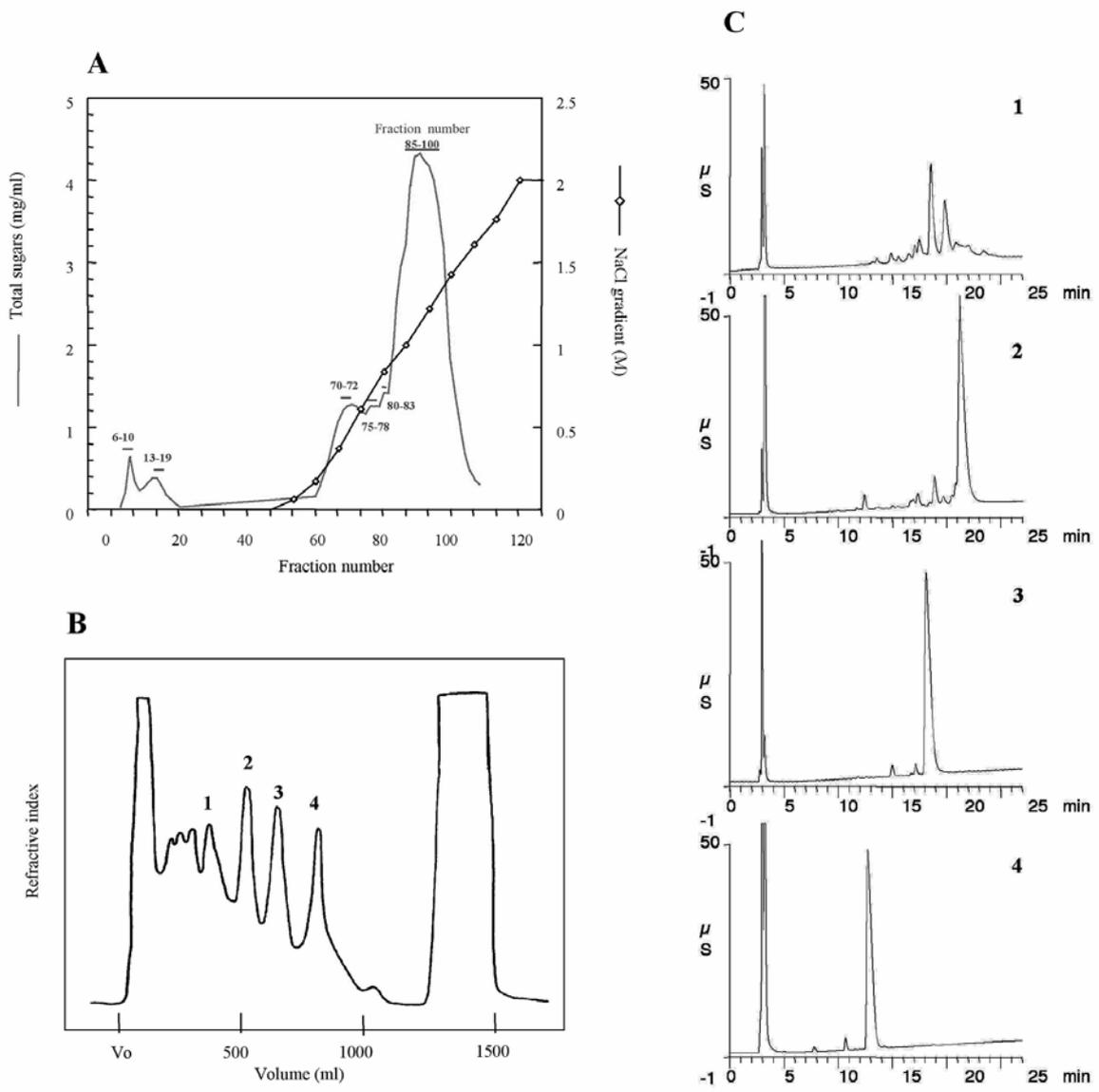


Fig. 4

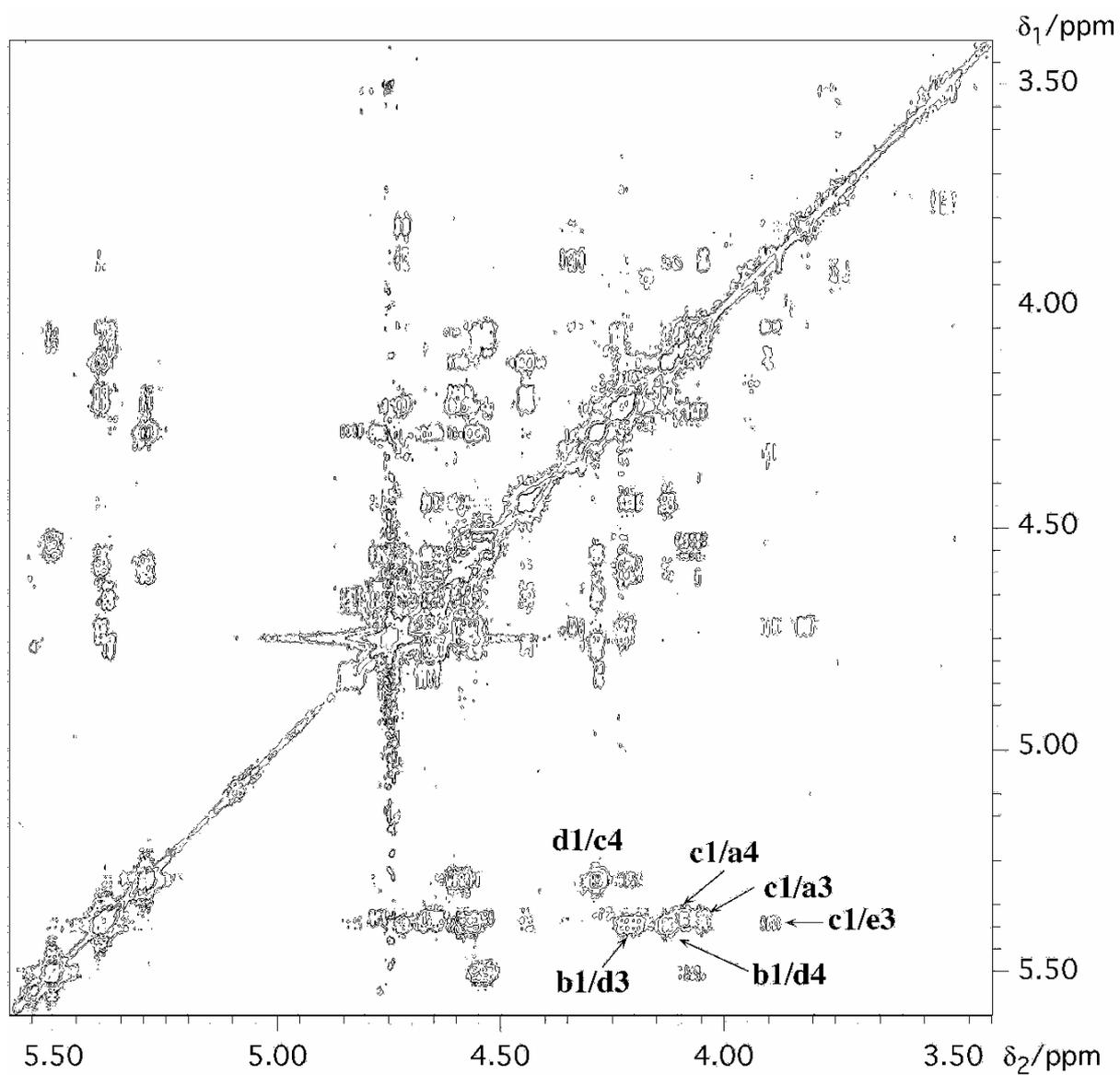


Fig. 5

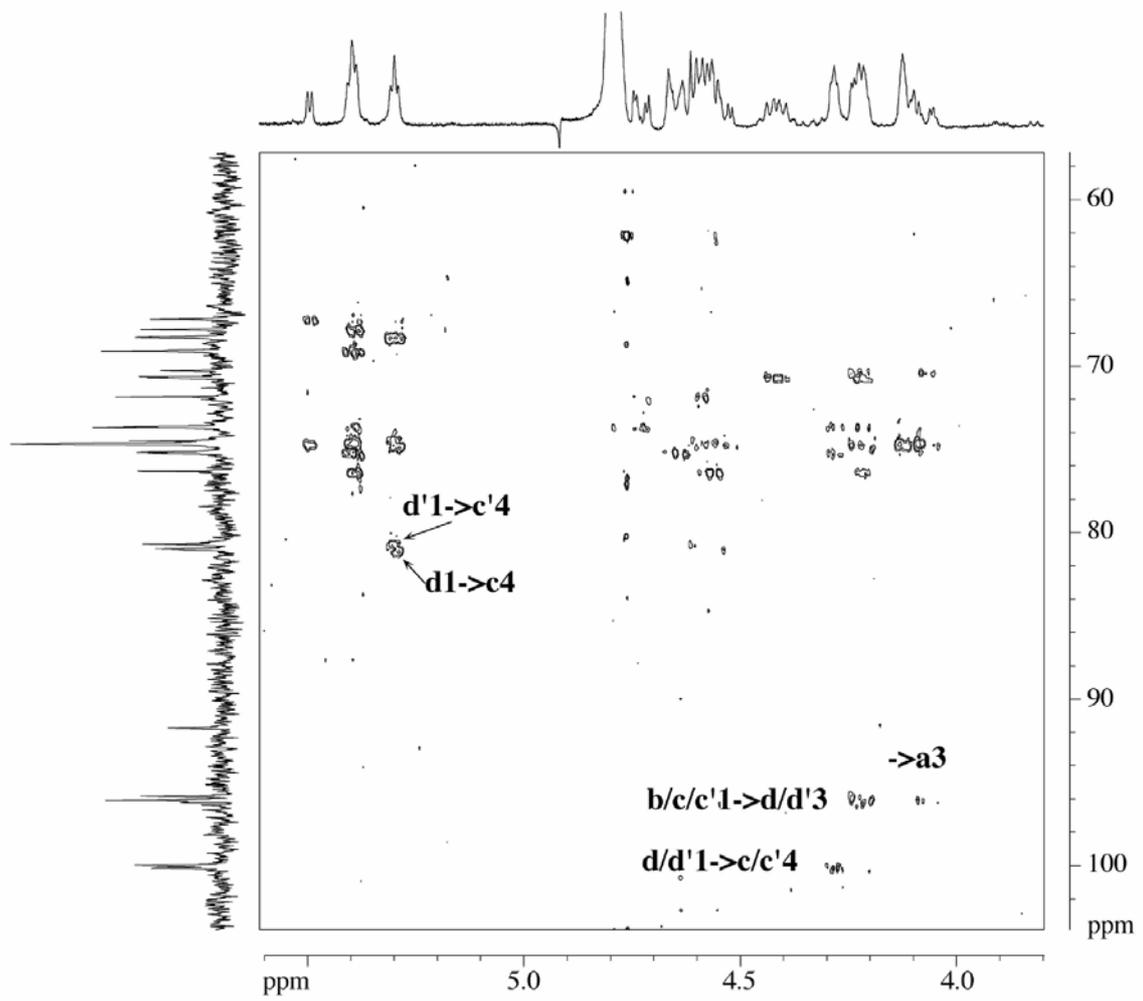


Fig. 6

ARTICLE 3

PURIFICATION AND EXPRESSION OF FUCOIDANASE, A NOVEL GLYCOSIDE HYDROLASE*

Sébastien Colin¹, Estelle Deniaud¹, Valérie Descamps¹, Murielle Jam¹, Nelly Kervarec², Jean-Claude Yvin¹, Tristan Barbeyron¹, Gurvan Michel¹ and Bernard Kloareg¹

From ¹Végétaux Marins et Biolomolécules, UMR 7139 (CNRS / UPMC / Goëmar), Station Biologique, Roscoff, France, ²Service Commun de Résonance Magnétique Nucléaire, Université de Bretagne Occidentale, Brest, France

Running Title: Fucoidanase, a novel family of glycoside hydrolase

Address correspondence to: Gurvan Michel, Végétaux Marins et Biolomolécules, UMR 7139 (CNRS / UPMC / Goëmar), Station Biologique, Place Georges Teissier, 29680 Roscoff, France, Tel. +33 298 292 330; Fax. +33 298 292 324 ; E-mail: gurvan@sb-roscoff.fr

Fucoidans are matrix polysaccharides from marine brown algae, consisting of an α -L-fucose backbone substituted by sulfate-ester groups, masked with ramifications, and containing other monosaccharide residues. We here report on the characterization of a novel glycoside hydrolase (FcnA) specific for the degradation of fucoidans. This glycoside hydrolase was purified to electrophoretic homogeneity from a Flavobacteriaceae referred to as SW5. The gene *fcnA* was cloned and sequenced (1007 amino acids) and the protein was produced in *E. coli*. FcnA exhibited a modular architecture, consisting of a 400 residues - long N-terminal domain followed by three contiguous cadherin-like domains and a 80 amino acids - long C-terminal domain. A truncated protein encompassing the N-terminal domain and the cadherin-like repeats was shown to retain the enzyme activity. The N-terminal catalytic domain shared about 20% of sequence identity with two other, potential fucoidanase genes and these three fucoidanases delineate a new family of glycoside hydrolases. As shown by size exclusion chromatography and ¹H-NMR analyses, the fucoidanase FcnA proceeds according an endo mode of action and cleaves the α -(1→4) glycosidic linkages within blocks of repeating motifs [→4)- α -L-fucopyranosyl-2,3-disulfate-(1→3)- α -L-fucopyranosyl-2-sulfate-(1→)]_n.

Protein-polysaccharide interactions have been extensively characterized at the structural level in the case of neutral glycans such as amylose and cellulose as well as of carboxylated polysaccharides such as pectins. With the exception of the proteoglycans of Metazoans, less is known on the structural bases of the

recognition and modification of sulfated polysaccharides. One can expect that the presence of bulky, acidic substituents on the polysaccharide backbone has affected the evolution of the proteins which modify these carbohydrates, at least in the fine topology of their catalytic sites.

Marine macroalgae feature a variety of cell wall sulfated polysaccharides most often recycled by marine bacteria, providing additional polysaccharide-polysaccharidase systems to gain insight into the structural biology of the hydrolases which degrade sulfated glycans. We have so far unraveled the crystal structure and mode of action of some bacterial glycoside hydrolases which depolymerize the neutral and sulfated galactans of red algae, agars and carrageenans. The β -agarases and κ -carrageenases of the family GH-16, which degrade neutral or moderately sulfated galactans (1, 2), adopt a β jelly roll fold (3, 4), only differing in the fine topology of their substrate binding site (4, 5). In contrast, ι -carrageenases, which are specific of the degradation of carrageenans with two sulfate substituents per repeating unit, constitute a distinct family of glycoside hydrolases, the family GH-82 (6), and fold into a β -helix (7). This fold is adopted by other enzymes which degrade highly acidic polysaccharides, such as PL-1 pectate lyases (8), PL-6 chondroitin lyases (9, 10) and GH-28 polygalacturonases (11). The GH-82 ι -carrageenases also differ from GH-16 β -agarases and κ -carrageenases by the recognition mode of their respective substrates (12).

The cell wall of brown algae also contains sulfated polysaccharides, the fucoidans. These anionic polymers encompass a continuous spectrum of highly ramified polysaccharides, ranging from high uronic acid, low-sulfate-

containing polymers with significant proportions of D-xylose, D-galactose, D-mannose to highly sulfated homofucan molecules with ca. three substituents per disaccharide (13, 14). In the cell wall matrix they associate with alginate, a linear polymer of β -D-mannuronic acid and α -L-guluronic acid (13). Fucoidans exhibit a variety of biological activities (15), both as cell wall ligands (16, 17) and as anticoagulant (18), antiviral, and antiproliferative agents. They were also shown to be recognized as defense elicitors in tobacco, where they induce systemic resistance against tobacco mosaic virus (19). These properties have attracted a renewed interest to their fine chemical structure in the last decade. Sulfated fucan oligosaccharides of about 8-14 fucose units, with a regular repeating disaccharide structure built on two residue types, 2-sulfated fucose and 2,3-disulfated fucose and with α -(1 \rightarrow 3) and α -(1 \rightarrow 4) glycosidic linkages were isolated by chemical methods (20), indicating a significant degree of repetitiveness in the fucoidan backbone. Yet, the general complex structure of fucoidans has remained somewhat elusive.

In this context the isolation and expression of hydrolases degrading fucoidans would be of special interest, to further elucidate the fine chemical structure these polysaccharides as well as their activity in homologous or heterologous systems, and to shine light on the structure of hydrolases which degrade polysaccharides with high linear densities of sulfate ester substituents. To date, only two, homologous fucoidan hydrolases, referred to as Fda1 and Fda2, were cloned and expressed, from the marine bacterium *Alteromonas sp.* SN-1009 (21), but they were not subjected to a detailed biochemical analysis. On the basis of its capacity to degrade sulfated fucans from fucoid algae, we recently isolated a flavobacteriacean strain, SW5, which secretes fucoidan-hydrolase activity in its culture medium. A crude enzyme preparation from SW5 culture supernatant was used to degrade the fucoidan from the brown alga *Pelvetia canaliculata*. End products included a tetrasaccharide and a hexasaccharide made of the repetition of disaccharidic units consisting of [\rightarrow 4)- α -L-fucopyranosyl-2,3-disulfate-(1 \rightarrow 3)- α -L-fucopyranosyl-2-sulfate-(1 \rightarrow)]_n, with the 3-linked residues at the non-reducing end (22). We here report the purification, the cloning and the overexpression of SW5 fucoidan hydrolase, referred to as FcnA.

The protein consists of an N-terminal, catalytic domain followed by three cadherin-like repeats and it harbors a C-terminal domain conspicuous of the hydrolases of Flavobacteriaceae. The enzyme depolymerizes fucoidans according a endolytic mode of action and cleaves the α -(1 \rightarrow 4) glycosidic linkage between 2-sulfated fucose and 2,3-disulfated fucose residues. We also show that, together with Fda1 and Fda2, FcnA belongs to a new family of glycoside hydrolases (23, 24).

Materials and Methods

Preparation and Analysis of Fucoidan Fractions

– Preparation of the crude fucoidan fraction from the fucoid brown alga *Pelvetia canaliculata*, referred to as FS28, was described previously (22). Briefly, thalli were macerated overnight in 1 L ethanol/formaldehyde/water (80:5:15; v/v) and exhausted by 2 L of ethanol/formaldehyde/water followed by 2 L of acetone. The resulting pellet was dried at 60°C and extracted twice for 3 h at 70°C by a 0.01 N HCl solution supplemented with 4% (w/v) calcium chloride. The extract was filtered, neutralized with ammonium carbonate and fucans were precipitated with 2.5 volumes of ethanol, re-suspended in water and freeze-dried. The resulting FS28 fraction contained 36 % fucose (22), was hygroscopic and displayed a brownish color, indicating the presence of polyphenols. In order to eliminate polyphenols, FS28 was further fractionated as follows. Aliquots (ca 500 mg) were re-suspended in water (1% w/v) and applied onto a Sephadex LH 20 column (Amersham Biosciences; total volume, 100 mL) equilibrated with water. The water-eluted fractions (total volume, 250 mL) were pooled, extensively dialysed (cut-off 3500 Da) against distilled water and freeze-dried. This white-colored fucoidan fraction contained 41 % fucose and is referred to as FC. In order to eliminate possible branching osidic residues (25), FC was dissolved in 0.5 M sulphuric acid (0.5%; w/v) and incubated at 40°C for 30 min. After neutralisation with 4 M sodium hydroxide, the sample was dialysed (cut-off 500 and 3500 Da) against water and freeze-dried. This fucoidan fraction is referred as FC_{H+}.

Assays for Fucoidanase Activity – Attempts to quantitatively assay for fucoidanolytic activity using the release of fucan oligosaccharides by a conventional reducing sugar assay (26, 27) were unsuccessful. Fucoidan degrading activity was

thus monitored by a carbohydrate-polyacrylamide gel electrophoresis (C-PAGE) assay of the release of anionic oligosaccharides (28). Briefly, 100-250 μL of 0.1% (w/v) FS28 fucoidan in 20 mM Tris-HCl buffer (pH 7.5) were incubated with 1-5 μL of enzyme fraction for 1 h at room temperature. The products (5 μL) were electrophoresed through a 6% (m/v) stacking and a 27% running, 1 mm-thick, polyacrylamide gel in 50 mM Tris-HCl, 2 mM EDTA buffer (pH 8.7) and stained with Alcian blue followed by silver nitrate (29). Fucoidanolytic activity was detected by the occurrence of anionic oligosaccharide bands in the bottom part of the running gel.

Production and Purification of Fucoidanase FcnA – A two day-old culture in ZF medium (22) of SW5 strain was inoculated (10 mL) in a 1-liter Erlenmeyer flask containing 250 mL of ZF and cultured at 22°C and 200 rpm for one day. Aliquots (40 mL) were then grown for five days in 5-liters Erlenmeyer flasks in 6 x 1 liters of the same medium under the same conditions and bacterial growth was monitored by the absorbance at 600 nm. The culture supernatant was collected by centrifugation at 12000 g for 20 min, filtered using a Pellicon system with 0.45 μm membrane (Millipore) and concentrated by ultrafiltration on a 30 kDa cut-off membrane (Millipore). The retentate (ca. 620 mL) was brought to 40% saturation with $(\text{NH}_4)_2\text{SO}_4$, the suspension was centrifuged at 12000 g for 15 min and the resulting supernatant was brought to 60% (w/v) saturation with $(\text{NH}_4)_2\text{SO}_4$. The precipitate was then collected by centrifugation at 12000 g for 15 min and dissolved in 60 mL of 20 mM Tris-HCl buffer pH 7.5, 5 mM MgCl_2 , 5 mM CaCl_2 and 50 mM NaCl. This fraction was diluted 20-fold in a 20 mM Tris-HCl buffer pH 7.5, resaturated at 40% with $(\text{NH}_4)_2\text{SO}_4$ and deposited on a Phenyl Sepharose CL4B column (Amersham) equilibrated with 100 mL of 20 mM Tris-HCl buffer pH 7.5, 40% $(\text{NH}_4)_2\text{SO}_4$. The enzyme was eluted with the same buffer followed by a linear gradient of $(\text{NH}_4)_2\text{SO}_4$ from 1.8 M to 0 M (180 mL) in the same buffer. The active fractions (determined by their ability to release oligosaccharides from the FS28 fucoidan fraction as assayed by C-PAGE) were pooled and diluted 5-fold in a 20 mM Tris-HCl buffer (pH 7.5) containing 5 mM NaCl. The solution was applied on a DEAE Sepharose CL6B column (Amersham) equilibrated with 100 mL of 20 mM Tris-HCl buffer (pH 7.5, 5 mM NaCl). The column was washed with the same

buffer (140 mL) and eluted with a linear gradient of 5 mM-1 M NaCl (210 mL). The active fractions were pooled (ca. 27 mL) and concentrated down to ca. 1 mL by ultrafiltration on a 10 kDa cut-off membrane (Centricon, Millipore). The concentrate (1 mL) was applied on a Superdex 200 (Amersham) equilibrated with 50 mM sodium phosphate buffer pH 7.0, 150 mM NaCl (100 mL) and the enzyme was eluted with the same buffer (30 mL). All purification steps were performed at 4°C. The apparent molecular weight of the purified enzyme was estimated by SDS-PAGE analysis, using a 12.5% polyacrylamide gel and a Phast System[®] (Amersham), with standard proteins ranging from 14.5 kDa to 200 kDa (broad range, BioRad). Protein contents were measured according to the Bradford method (30), using a BioRad protein assay reagent with bovine serum albumin (Sigma) as standard. Internal peptide microsequences were determined by Edman degradation at the Pasteur Institute (Paris, France).

Isolation and Analysis of Fucoidanase Clones – Total DNA from SW5 strain was prepared as previously described (31), cut by the restriction endonuclease *Sau3AI*, and fractionated on a sucrose gradient. DNA fragments of approximately 4-10 kb were cloned at the *BamHI* site of plasmid vector pAT153 which was used to transform *E. coli* strain DH5 α . The genomic library contained approximately 6000 clones. Using the microsequences obtained from internal peptides of the purified fucoidanase (in bold and underlined Fig.2), degenerated DNA primers were designed and an internal DNA probe of 203 nt (underlined, Fig. 2) was obtained by polymerase chain reaction. The probe was labelled (kit Megaprime DNA labelling systems, Amersham Lifescience), and used to screen the library according to (32) then the clones of interest were mapped by restriction endonucleases. The probe was used for a Southern Blot experiment (32) carried out on genomic DNA, using various endonuclease combinations (*HinDIII/BamHI*, *HinDIII*, *HinDIII/XcmI*, *EcoRI/NdeI*, *PstI/NdeI*, *EcoRI/PstI/NdeI*, *ClaI/NdeI*, *XcmI/NdeI*, *XcmI/PvuII*, *EcoRI/SalI*). Sequencing was carried out by gene walking on both strands, using a 3100 Genetic Analyser (Applied Biosystem Instruments) with BigDye Terminator V3.0 chemistry (DNA sequencing kit, Applied Biosystem) and synthetic oligonucleotides as

primers. Sequences were verified at least 5 times and most of them more than 10 times.

Sequence Analysis – Nucleotide sequences were searched for open reading frames, as well for putative ribosomal binding sequences, putative promoters, and putative transcription terminators hairpins via the *mfold* v.3.1 prediction server (33). The signal peptide was predicted with PSORT (34) and SIGNALP v.1.1 and v.2.0 (35). The searches of homology were run with GENBANKnr and PAT via the NCBI web server using BLASTp and PSI-BLAST (36) and with UNIPROT from the EBI web server using WU-BLAST and MPsrch (37). Evidence for specific motifs and domains was queried using INTERPRO via InterProScan (EBI; 38, 39) and via CDD domain databases (NCBI). The fucanase sequence was checked for repeats using the program RADAR (EBI; 40).

Production of Recombinant Fucoidanase – The overexpression vector was constructed using GATEWAY™ cloning technology (Invitrogen), according to the supplier's instructions. The coding region of SW5 fucoidanase, either as a whole or truncated after the cadherin repeats (see Fig. 4), was amplified by PCR from genomic DNA with high fidelity PLATINUM Pfx DNA polymerase (Invitrogen). PCR products were inserted into pDONR201 vector, the sequence was checked, and a valid clone was inserted into the pDEST17 vector with a N-terminal 6xHistidine tag. Plasmid vectors were used to transform the DE3 *E. coli* BL21 pLysS strain (Novagen) and overexpression was carried out in the M9 medium supplemented with 2% casaminoacids, 100 µg.ml⁻¹ ampicillin and 35 µg.ml⁻¹ chloramphenicol. Recombinant bacteria were grown in a Bioflow 3000 fermentor at 37°C, pH 7.2 and under regulated oxygen concentration (70% of air). When cultures reached 0.8 OD₆₀₀, expression was induced at 20°C with 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). Partially purified recombinant fucanases were produced as follows. After an overnight induction, bacterial cultures (500 mL) were centrifuged for 15 min at 10000 g and pellets were either frozen (-20°C) or re-suspended in 50 mL of 20 mM Tris buffer, pH 7.5 containing 250 mM NaCl, 80 mM imidazol and one tablet of Complete™ protease inhibitors (Roche Diagnostics), then disrupted by a French Press (4°C). The lysate was ultracentrifuged at 40000 g for 1 hour and the supernatant was loaded on a Fast Flow Chelating Sepharose (Amersham) column saturated with

100 mM NiSO₄. The affinity chromatography column (V = 10 mL) was washed with lysis buffer (5 V) and eluted with an 80-300 mM imidazol gradient (12 V), then by 1 M imidazol (2 V). Protein contents were estimated by SDS-PAGE and aliquots of the recombinant enzyme were assayed for fucoidan-degrading activity by C-PAGE, as above.

Size exclusion chromatography and ¹H-NMR monitoring of the enzymatic degradation of fucoidan fractions – The enzymatic degradation of fucoidan was monitored by size exclusion chromatography as follows. Fucoidan aliquots (20-50 mg, 0.5-1% in water) were extensively digested with native or recombinant fucoidanase for 42 h at room temperature, then boiled for 10 min and freeze-dried. Aliquots (10 mg) of the digestion products were re-suspended (0.5-1% w/v) in water and chromatographed with 50 mM ammonium carbonate using a Superdex 200 column (Amersham), equipped with refractometric detection (*TSP*). Liquid ¹H-spectra were recorded at 60°C on a Bruker DRX 500 Avance (Bruker, Wissembourg, France) at 500 MHz. Samples (ca. 5-10 mg) were previously exchanged twice in D₂O (100%) before solubilization in 100% D₂O.

RESULTS

Purification of an Extracellular Fucoidan-Degrading Enzyme from SW5 – Purification of the fucan-degrading activity of the SW5 strain was based on a carbohydrate-polyacrylamide gel electrophoresis (C-PAGE) assay of the release of anionic oligosaccharides from the fucoidan fraction FS28. From the concentrated bacterial culture supernatant, purification to electrophoretic homogeneity was achieved in four steps, namely protein precipitation with ammonium sulfate at 40-60% saturation followed by hydrophobic interaction chromatography on Phenyl Sepharose, ion-exchange chromatography on DEAE-Sepharose and gel filtration on Superdex 200. The active protein fractions corresponded to 50, 35, 0.50 and 0.008% of the initial protein content, respectively. The latter protein fraction migrated as a single band in SDS-PAGE analysis, corresponding to an apparent molecular weight of 105 kDa (Fig. 1). This fucoidan-degrading enzyme is referred to as FcnA. The SDS-PAGE 105 kDa protein band was excised and three internal peptide sequences were determined by Edman degradation (see Fig. 2).

The Fucoidanase Gene fcnA – An internal gene probe of 203 nt (underlined in Fig. 2) was generated by PCR from the peptidic microsequences A and B and used to screen a genomic library prepared with SW5 total DNA. Of the 6000 clones in the library, three clones, referred to as pAT153fu11, pAT153fu14 and pAT153fu33 and with respective insert sizes of 4.0 kb, 13.5 kb and 4.0 kb, hybridized with the probe. Southern Blot experiments showed that SW5 strain contains only one copy of the *fcnA* gene (data not shown). Restriction analyses indicated that the inserts from pAT153fu11 and pAT153fu33 were the same genomic DNA fragment, cloned in opposite directions, and that the pAT153fu14 insert overlapped these fragments by ca. 1.5 kb. The pAT153fu11 and pAT153fu14 inserts were sequenced on both strands. The two sequences perfectly overlapped over a length of 1610 nt, the fucanase open reading frame being interrupted at the 3' end of pAT153fu11 insert and finishing on the pAT153fu14 insert. The physical validity of this ORF was proven by PCR amplification from the genomic DNA, leading to the cloning of the entire, 3021 nt-long fucanase gene on a contig of 7485 nt. The *fcnA* gene was surrounded, at 157 nt upstream and 44 nt downstream on the same strand (Fig 2), by two complete ORFs as well as, at the contig ends, by two truncated ORFs. Sequence similarities suggest that these ORFs consist of two sulfatases, one esterase and one fucosidase (data not shown). No obvious Shine-Dalgarno ribosome binding site consensus sequence (41) was found upstream of the *fcnA* gene nor upstream of the other genes in this contig. In contrast, 15 nt upstream of the ATG codons, a conserved motif, AAxTAAAT, was present in all of the four available 5' ends. No obvious hairpin was identified as transcription terminator in any of the four gene 3' ends in this contig.

The Fucoidanase FcnA Displays a Modular Architecture – The predicted product of the fucoidanase gene from SW5 strain, FcnA, is a preprotein of 1007 amino acids, with a theoretical molecular mass of 110.3 kDa (Fig 2), consistent with the apparent molecular mass of the secreted protein upon SDS-PAGE analysis, 105 kDa (Fig. 1). It includes the three internal peptides determined by microsequencing. Based on the prediction program SignalP V2.0, a signal peptide is likely to be cleaved between A28 and Q29. Using BLASTp, PSI-BLAST, WU-BLAST, and MPsrch including with high

sensitivity settings, 60% of the sequence did not share any significant similarity with the protein sequences available in GENBANKnr and UNIPROT. Only the core and the C-terminal end of FcnA brought significant matches (see below). However, a BLASTp search of the PAT database (NCBI) provided two significant hits with two yet unpublished fucoidanase sequences from *Alteromonas sp.* SN-1009 (21), referred to as Fda1 (AAO00508) and Fda2 (AAO00509). The three proteins shared a large domain in their N-terminal ends (approximately 380 residues long). Fda1 and Fda2 are highly similar, displaying 70% identity over their entire length, 814 and 881 amino acids, respectively, whereas they displayed 20% and 18% sequence pairwise identities, respectively, with FcnA over their common N-terminal domain (Fig 3).

In the protein core, 400 residues from the N-terminal end, RADAR searches also detected three tandemly repeated domains of 105 residues each (repeats R1, 419 to 526; R2, 526 to 632; R3, 632 to 736; Fig. 4), which share 31% (R1 / R2), 27% (R1 / R3) and 48% (R2 / R3) pairwise sequence identities. In this region, domain recognition methods (Interproscan) identified via SMART a domain family (SM00736) matching repeat 2, referred to as CDAG (Dystroglycan-type cadherin-like domains). This domain has known homologues in Metazoan, Yeasts and Bacteria (42). PSI-BLAST analyses confirmed the close relationship between these repeats and cadherin-like domains. In GENBANKnr, the best hits, representing 9 to 28% identity, were with bacterial proteins which also display a tandem organization of cadherin-like domains, namely from Proteobacteria (*Shewanella oneidensis*, NP_715831; *Microbulbifer degradans*, ZP_00067001; *Geobacter sulfurreducens*, NP_953122; *Magnetococcus sp.*, ZP_00045566; *Vibrio parahaemolyticus*, NP_798012), Cyanobacteria (*Nostoc sp.*, NP_490234), and Planctomyces (*Pirellula sp.*, NP_864105). Interestingly only one hit was found from the Bacteroidetes phylum (*Cytophaga hutchinsonii*, ZP_00116997) and it was much more distant (see supplemental data for alignment).

In the C-terminal region, a stretch of 75 residues displayed significant matches, ranging from 14% - 31% identity, exclusively with the C-terminal domains of several polysaccharidases and proteases from bacteria of the *Bacteroidetes* phylum, including the genera *Porphyromonas*, *Cytophaga*, *Zobellia*, *Flavobacterium*,

Microscilla and *Rhodothermus*. In particular the YPNP motif at the beginning of these domain sequences is highly conserved.

Altogether, FcnA consists of a 1007 amino acid-long protein (110.3 kDa) with a modular architecture. This protein encompasses a 28-aa long signal peptide, a 390-aa long N-terminal domain, three successive cadherine-like domains with ca. 105 residues each, followed by a region of 196 amino acids without any significant sequence similarity with known proteins and, finally by a 75-aa long C-terminal domain (Fig. 4).

Overexpression of the Fucoidanase FcnA Gene from SW5 Strain – The fucanase FcnA was expressed under two different forms (Fig. 4). The first construct (FcnA1) encoded the whole FcnA sequence without its predicted signal peptide in frame with the N-terminal His tag of the pDEST17 vector. The second construct (FcnA2) encompassed the N-terminal domain and the three cadherin-like repeats of the protein. The expected products were proteins with 1001 residues (109.7 kDa), and 788 residues (86.5 kDa), respectively. As indicated by SDS-PAGE analysis (data not shown), products with the expected protein sizes were obtained from the two constructs, although FcnA2 was expressed in a greater amount than FcnA1. Contaminant proteins co-eluted with FcnA1 and FcnA2 in the elution peak at ca. 150 mM imidazole. Further purification steps combining anion exchange and size exclusion chromatography did not significantly improve the purity of FcnA1 and FcnA2 (data not shown).

Action of Fucoidanase FcnA on Fucoidan fractions – The crude fucoidan fraction, FS28, was incubated in the presence of either one or the other partially purified recombinant enzymes, FcnA1 and FcnA2, and the enzymatic degradation kinetics was monitored by carbohydrate-PAGE analysis. As illustrated for FcnA2 (Fig. 5), both recombinant enzymes yielded the same profiles in C-PAGE analysis of the degradation products. However, as monitored by size exclusion chromatography (SEC), FcnA2 did not extensively degrade the FS28 fucoidan fraction (not shown). Similar results were obtained when using the FC fucoidan fraction. The ¹H-NMR spectrum (data not shown) and the size profile of FC did not significantly change upon incubation in the presence of the recombinant fucoidanase (Fig. 6A). In contrast, the fucoidan fraction FC_{H+}, resulting from the treatment of mild acid

hydrolysis, was amenable to significant degradation by the recombinant fucoidanase. In SEC analyses, the size profile of FC_{H+} showed the presence of high molecular weight polymers, which eluted close to the exclusion volume, with only a small proportion of oligosaccharides. After incubation in the presence of FcnA2, the size distribution was markedly shifted towards the lower molecular weights (Fig. 6B).

¹H-NMR analyses of the enzymatic degradation of FC_{H+} – As shown in Fig. 7A, the ¹H-NMR spectrum of fucoidan which had been subjected to mild acid hydrolysis, FC_{H+}, was comparable to that of polymeric fucoidan. In the region of the anomeric protons, some signals are found with chemical shifts similar to those described by (20) as alternating α -(1→4) linked 2,3-disulfated-fucose and α -(1→3) linked 2-sulfated-fucose residues (F2,3S-F2S). Therefore, this pattern, which was proposed to be the main structure in the fucoidans from *Ascophyllum nodosum* and *Fucus vesiculosus* fucoid algae (20), is indeed present in the fucoidan of the fucalcaean alga *Pelvetia canaliculata*. After incubating FC_{H+} with FcnA2, the ¹H-NMR spectrum was markedly modified (Fig. 7B), particularly in the regions of the α -anomeric protons (5.6-5.2 ppm) and of the methyl groups (1.4-1.0 ppm). The α -anomeric signals corresponding to the repeating motif F2,3S-F2S were affected by the enzyme action. The signal assignable to the H1 of the internal α -(1→3) linked 2-sulfated-fucose residues (5.30-5.34 ppm) was refined, but its intensity did not decrease upon hydrolysis. In contrast, the signal of the H1 of the internal α -(1→4) linked 2, 3-disulfated-fucose residues (5.45 ppm) significantly diminished. A doublet also appeared, with a chemical shift close to that of the α -(1→3) linked 2-sulfated-fucose reducing-end residues (5.51 ppm).

To further investigate the identity of the linkages cleaved by FcnA2, DQF-COSY spectra were recorded on FC_{H+} and FC_{H+} digested by FcnA2 (Fig. 7 C, D). On the spectrum of FC_{H+} (Fig. 7C), a H1-H2 cross-peak was identified for the internal α -(1→4) linked 2,3-disulfated-fucose residues (5.45 ppm; 4.69 ppm). This H1-H2 cross-peak was no longer present after the fucoidanase treatment, indicating the disappearance of internal α -(1→4) linked 2,3-disulfated-fucose residues (Fig. 7D). In contrast, the H1-H2 cross-peak of the internal α -(1→3) linked 2-sulfated-fucose residues (5.30 - 5.35

ppm; 4.62 - 4.65 ppm) was modified but remained apparent. Although the signals for reducing and non-reducing ends fucose residues were weak in the polymeric FC_{H+}, a H1-H2 cross-peak was nevertheless identified for the reducing-end α -(1→3) linked 2-sulfated-fucose residues (5.54 ppm; 4.55 ppm). After the action of FcnA2, the intensity of this cross-peak clearly increased. Altogether, the above results indicate that FcnA cleaves the α -(1→4) linkages within the repeating motifs [\rightarrow 4)- α -L-fucopyranose-2,3-disulfate-(1→3)- α -L-fucopyranose-2-sulfate-(1→)]_n of the fucoidan polysaccharidic backbone.

DISCUSSION

FcnA is a Sulfated Fucan α -1,4 Endohydrolase – We here report the purification of the fucoidanolytic activity secreted by a marine Cytophagale referred to as SW5 (22). Using a four-step procedure, the protein content was decreased 12500-fold from the culture medium down to the enzyme electrophoretic homogeneity. As the bioassay used to monitor protein purification, observation by carbohydrate-PAGE of the release of oligosaccharides from a complex fucan substrate, was not amenable to quantification, the ratio of enzyme recovery was not assessed. Based on internal peptides sequenced from the purified protein the enzyme was cloned and successfully produced in *E.coli*, both as full-length and truncated proteins and tagged with an N-terminal histidine tail.

Several lines of evidence indicate that the full-length recombinant protein is indeed a fucoidanase: i) its sequence featured the three internal peptide sequences found in the wild-type protein (Fig. 2); ii) based on C-PAGE analysis it released oligosaccharides from the FS28 fucoidan fraction (Fig 5); iii) as seen by SEC (Fig. 6) and NMR (Fig. 7) analyses, the recombinant protein significantly depolymerized a fucan fraction obtained by mild acid hydrolysis of fucoidan, FC_{H+}; and iv) the N-terminal domain of FcnA exhibited low yet significant sequence similarities with two patented fucoidanases (21). Based on the marked size decrease of the initial polysaccharide substrate (Fig. 6) leading to the release of intermediary sized fucan oligosaccharides (Fig 5), the fucoidanase proceeds according to an endo mode of action. ¹H-NMR analyses of the hydrolysis

products (Fig. 7) demonstrate that FcnA cleaves α -(1→4) glycosidic linkages within blocks of the repeating disaccharidic motif [\rightarrow 4)- α -L-fucopyranosyl-2,3-disulfate-(1→3)- α -L-fucopyranosyl-2-sulfate-(1→)]_n. This result is consistent with our former identification of the tetra- and the hexasaccharides of this homologous series in the degradation products of fucoidans (22).

However, as shown by SEC analyses (Fig. 6), the recombinant fucoidanase did not extensively hydrolyze the purified fucoidan fraction FC, whereas it significantly degraded FC_{H+}, the fucan fraction obtained by mild acid hydrolysis of FC. The fucoidans of fucallean algae consist of a polysaccharidic backbone of alternating α -(1→4) linked fucose and α -(1→3) linked fucose residues, masked by various substitutions, mainly ester-sulfate substituents and lateral branches (18, 20, 43, 44, 45). Mild acid hydrolysis is likely to have unmasked the fucoidan core structure and increased the accessibility of fucoidanase to its substrate. Altogether, SW5 FcnA fucoidanase appears to specifically cleave the α -(1→4) glycosidic linkages of the polysaccharidic backbone of fucoidans. The extensive degradation of complex fucoidans seems to require that this glycoside hydrolase acts in synergy with other enzymes. This hypothesis is consistent with our finding that the *fcnA* gene of SW5 is surrounded by genes encoding a variety of proteins potentially acting on sulfated fucans, namely two sulfatases, one putative carboxyl esterase, and one fucosidase, suggesting that the *fcnA* locus is a gene cluster involved with fucoidan degradation.

Sulfated Fucan Endohydrolases Constitute a New Family of Glycoside Hydrolases – FcnA features five remarkable domains, the N-terminal region, three cadherin-like repeats of the CDAG family and a C-terminal stretch. The C-terminal region of FcnA is conserved at the C-terminus of several proteins expressed by bacteria from the Bacteroidetes phylum, mainly including degrading enzymes such as polysaccharidases or proteases. Nothing is known yet on the function of this specific peptidic sequence in these bacteria. Given the retention of fucan-degrading activity by the truncated form of the enzyme, FcnA2, this C-terminal domain does not participate in the catalytic machinery of FcnA. As hinted for the β -agarases and κ -carrageenase from *Zobellia*

galactanivorans (2), we tentatively propose that this conspicuous C-terminal extension is involved in the enzyme localization.

In Metazoans, cadherin repeats are known to be involved with calcium-dependant homophilic protein binding. When aligned with cadherin sequences from the PDB, the cadherin-like domain of FcnA featured, at conserved positions within the three repeats, most of the Ca^{2+} binding sites of cadherins (see supplemental data). Therefore the cadherin-like domains of FcnA likely are calcium-binding domains. The tandem repetition of these domains suggests that they may adopt a rigid, rod-like shape upon calcium binding (46). In bacteria the domains which match best the repeats of FcnA are found in such large proteins as putative RTX toxins (COG2931), important virulence factors produced by a wide range of Gram-negative bacteria.

The N-terminal region of FcnA, which was alignable over a length of 384 residues with the N-terminal domains of the ca. 800 aa-long fucoidan hydrolases of *Alteromonas* sp SN-1009, Fda1 and Fda2 (Fig 3), likely is the catalytic domain of the fucoidanase. This observation is further substantiated by the fact that the truncated protein FcnA2, which comprises the N-terminal domain and the three cadherin-like repeats only, retained the enzyme activity. It also follows that the N-terminal portions of Fda1 and Fda2, which share 20 and 18% sequence identity with that of FcnA, respectively (Fig. 3), also are the catalytic

domains of these fucoidanases. These three sequences do not match significantly with any proteins in the public databases and no other fucoidanase sequences are known so far. We thus propose that they constitute a novel family of glycoside hydrolases (23, 24). Eleven acidic residues, which could serve as catalysts in the hydrolysis mechanism (47), are conserved in all of the three fucoidanases (Fig. 3). Based on the conservation of their immediate surroundings, we propose that D221, D226 or D289 (numbering of FcnA, Fig. 3) are the most likely candidates to act as catalytic residues in this family.

In conclusion, the fucoidanase from SW5, referred to as FcnA, should help cracking the structure of the fucoidans of brown algae, a task which so far has proven difficult because of the lack of appropriate, specific fucan hydrolases. This fucan α -(1→4) endohydrolase depolymerizes the fucoidan backbone, made of stretches of alternating α -(1→4) linked 2,3-disulfated-fucose and α -(1→3) linked 2-sulfated-fucose residues. FcnA thus provides another case study of the degradation of polysaccharides with high densities of sulfate-ester substituents. We show here that, together with two other bacterial fucanases, FcnA delineates a novel structural family, thus opening a new opportunity to further decipher the structural bases of the degradation of sulfated polysaccharides.

REFERENCES

1. Barbeyron, T., Gerard, A., Potin, P., Henrissat, B., and Kloareg, B. (1998) *Mol. Biol. Evol.* **15**, 528-537
2. Jam, M., Flament, D., Allouch, J., Potin, P., Thion, L., Kloareg, B., Czjzek, M., Helbert, W., Michel, G., Barbeyron, T. (2005) *Biochem. J.*, **385**, 703-713
3. Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) *Structure* **9**, 513-525
4. Allouch, J., Jam, M., Helbert, W., Barbeyron, T., Kloareg, B., Henrissat, B., and Czjzek, M. (2003) *J. Biol. Chem.* **278**, 47171-47180
5. Allouch, J., Helbert, W., Henrissat, B., and Czjzek, M. (2004) *Structure* **12**, 623-632
6. Barbeyron, T., Michel, G., Potin, P., Henrissat, B., and Kloareg, B. (2000) *J. Biol. Chem.* **275**, 35499-35505
7. Michel, G., Chantalat, L., Fanchon, E., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) *J. Biol. Chem.* **276**, 40202-40209
8. Yoder, M. D., Keen, N. T., and Jurnak, F. (1993) *Science* **260**, 1503-1507
9. Huang, W., Matte, A., Li, Y., Kim, Y. S., Linhardt, R. J., Su, H., and Cygler, M. (1999) *J. Mol. Biol.* **294**, 1257-1269
10. Michel, G., Pojasek, K., Li, Y., Sulea, T., Linhardt, R. J., Raman, R., Prabhakar, V., Sasisekharan, R., and Cygler, M. (2004) *J. Biol. Chem.* **279**, 32882-32896

11. Pickersgill, R., Smith, D., Worboys, K., and Jenkins, J. (1998) *J. Biol. Chem.* **273**, 24660-24664
12. Michel, G., Helbert, W., Kahn, R., Dideberg, O., and Kloareg, B. (2003) *J Mol Biol.* **334**, 421-433
13. Kloareg, B., and Quatrano, R. S. (1988) *Oceanogr. Mar. Biol. Annu. Rev.* **26**, 259-315
14. Mabeau, S., Kloareg, B., and Joseleau, J. P. (1990) *Phytochemistry* **29**, 2441-2445
15. Berteau, O., and Mulloy, B. (2003) *Glycobiology* **13**, 29-40
16. Kropf, D. L., Kloareg, B., and Quatrano, R. S. (1988) *Science* **239**, 187-189
17. Goodner, B. W., and Quatrano, R. S. (1993) *Plant Cell* **5**, 1471-1481
18. Pereira, M. S., Mulloy, B. and Mourão, P. A. S. (1999) *J. Biol. Chem.* **274**, 7656-7667
19. Klarzynski, O., Descamps, V., Plesse, B., Yvin, J. C., Kloareg, B., and Fritig, B. (2003) *Mol. Plant Microb. Interact.* **16**, 115-122
20. Chevolut, L., Mulloy, B., Ratiskol, J., Foucault, A., and Collic-Jouault, S. (2001) *Carbohydr. Res.* **330**, 529-535
21. Takayama, M., Koyama, N., Sakai, T., and Kato, I. (2002) US Patent n°US6489155 B1
22. Descamps, V., Colin, S., Jam, M., Richard, C., Potin, P., Barbeyron, T., Yvin, J. C., and Kloareg, B. (2005) *Mar. Biotechnol.* **submitted for publication**
23. Henrissat, B. (1991) *Biochem. J.* **280**, 309-316
24. Coutinho, P. M., and Henrissat, B. (1999) *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G., Henrissat, B., and Svensson, B., eds.), pp. 3-12, The Royal Society of Chemistry, Cambridge, United Kingdom.
25. Chevolut, L., Foucault, A., Chaubet, F., Kervarec, N., Sinquin, C., Fisher, A. M., and Boisson-Vidal, C. (1999) *Carbohydr. Res.* **319**, 154-165
26. Kidby, D. K., and Davidson, D. J. (1973) *Anal. Biochem.* **55**, 321-325
27. Nelson, N. (1944) *J. Biol. Chem.* **153**, 375-380
28. Zablackis, E., and Perez, J. (1990) *Bot. Mar.* **33**, 273-276
29. Min, H., and Cowman, M. K. (1986) *Anal. Biochem.* **155**, 275-285
30. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254
31. Barbeyron, T., Kean, K., and Forterre, P. (1984) *J. Bacteriol.* **160**, 586-590
32. Sambrook, J., and Russel, D. W. (2001) *Molecular cloning: a laboratory manual*, pp. 1.126-1.128, 1.135-1.142, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
33. Zuker, M., Mathews, D. H., and Turner, D. H. (1999) *RNA Biochemistry and Biotechnology* (Barciszewski, J., and Clark, B. F. C., eds.), pp. 11-43, Kluwer Academic Publishers, Dordrecht, Netherlands
34. Nakai, K., and Kanehisa, M. (1991) *Proteins* **11**, 95-110
35. Nielsen, H., Brunak, S., and von Heijne, G., (1999) *Protein Eng.* **12**, 3-9
36. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389-3402
37. Smith, T. F., and Waterman, M. S. (1981) *J. Mol. Biol.* **147**, 195-197
38. The InterPro Consortium* (*Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M. D. R., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N. J., Oinn, T. M., Pagni, M., Servant, F., Sigrist, C. J. A., and Zdobnov E. M.) (2001) *Nuc. Ac. Res.* **29**, 37-40
39. Zdobnov, E. M., and Apweiler, R. (2001) *Bioinformatics* **17**, 847-848
40. Heger, A., and Holm, L., (2000) *Proteins* **41**, 224-237
41. Kozak, M. (1999) *Gene* **234**, 187-208
42. Dickens, N. J., Beatson, S., and Ponting, C. P. (2002) *Curr. Biol.* **12**, 197-199
43. Daniel, R., Berteau, O., Jozefonvicz, J., and Goasdoue, N. (1999) *Carbohydr. Res.* **322**, 291-297
44. Bilan, M. I., Grachev, A. A., Ustuzhanina, N. E., Shashkov, A. S., Nifantiev, N. E., and Usov, A. I. (2002) *Carbohydr. Res.* **337**, 719-730
45. Bilan, M. I., Grachev, A. A., Ustuzhanina, N. E., Shashkov, A. S., Nifantiev, N. E., and Usov, A. I. (2004) *Carbohydr. Res.* **339**, 511-517
46. Koch, A. W., Manzur, K. L., Shan, W. (2004) *Cell. Mol. Life Sci.* **61**, 1884-1895
47. Henrissat, B., and Davies, G. (1995) *Structure* **3**, 853-859
48. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15**, 305-308

FOOTNOTES

* S.C. and V.D. were the recipients of PhD fellowships co-funded by the ANRT, whose help is gratefully acknowledged.

The nucleotide sequence(s) reported in this paper has been submitted to the Genbank™/ EBI Data Bank with accession number(s).

FIGURE LEGENDS

Fig. 1. Purification to homogeneity of SW5 fucanase, FcnA. The enzyme fraction after Superdex 200 HR was concentrated (Centricon 10 kDa), submitted to electrophoresis on a 12.5% SDS polyacrylamide gel (Phast-system®, Pharmacia) and stained with Coomassie blue (lane 1). The molecular weight markers (lane 2) consisted of lysozyme, soybean trypsin, carbonic anhydrase, ovalbumin, serumalbumin, phosphorylase B, E. coli β -galactosidase and myosine (BioRad).

Fig. 2. Nucleotide and deduced amino acid sequences of SW5 fucoidanase gene, *fcnA*. The putative ribosomal binding site is shown in bold. The open reading frame starts at nucleotide 226, encoding a pre-protein of 110.3 kDa. A 28-aa peptide, shown in bold and italics is proposed as the signal peptide. The amino acid sequences shown underlined and in bold (A-C) were confirmed by peptide sequencing, and the peptides A and B were used to design PCR primers to amplify an internal gene DNA probe. The C-terminal amino-acid sequence which was truncated in the FcnA2 construct is indicated in italics.

Fig. 3. Alignment of FcnA with the fucoidanases from *Alteromonas* SN-1009. The amino acid sequences were aligned to optimize the alignment of predicted secondary structure elements. The figure was drawn using the program ESPript 2.0 (48), with the residues colored according to the conservation of their physico-chemical properties. Grey triangles underline basic residues homologies and black triangles indicate basic residues identities. The grey circles indicate conserved acidic residues while the black ones refer to potential catalytic amino acids (D221, D226, and D289, referring to the numbering of FcnA).

Fig. 4. The modular architecture of FcnA. FcnA consists of a 400 residues - long N-terminal catalytic domain (black) followed by three contiguous cadherin-like domains (R1-R3, hatched), a ca. 200 residues long region of unknown function (white) and by a 75 amino acids - long C-terminal domain (grey). The fucoidanase gene, either as a whole (FcnA1) or truncated after the cadherin-like repeats (FcnA2), was cloned in frame with the N-terminal His tag of the pDEST17 vector (GATEWAY™ Cloning Technology) for expression in *E. coli*.

Fig. 5. Carbohydrate- PAGE analysis of the degradation products of fucoidan with recombinant fucoidanase. Hydrolysis kinetics of crude, native fucoidan (FS28) by the recombinant fucoidanase (FcnA2), showing the endolytic mechanism of the hydrolase and the end products of the enzyme.

Fig. 6. Size exclusion chromatography analysis of the degradation of fucoidans with recombinant fucoidanase. A. SEC profile (Superdex 200, detection by refractometry) of purified, native fucoidan (FC, bold line) and after degradation by the recombinant enzyme (FC-FcnA2). B. SEC profile of FC after mild-acid hydrolysis (FC_{H+}, bold line) and followed by degradation with the recombinant enzyme (FC_{H+}-FcnA2). Note that FC resisted enzyme hydrolysis much more than FC_{H+}.

Fig. 7. ¹H-NMR analysis of the degradation of fucoidans with recombinant fucoidanase. A. 1-D spectrum of FCH⁺. B. 1-D spectrum of FCH⁺ after degradation by the recombinant enzyme (FcnA2). C. COSY DQF spectrum of FCH⁺. D. COSY DQF spectrum of FCH⁺ after degradation by FcnA2. Spectra were recorded at 500 MHz and at 60°C in D₂O.

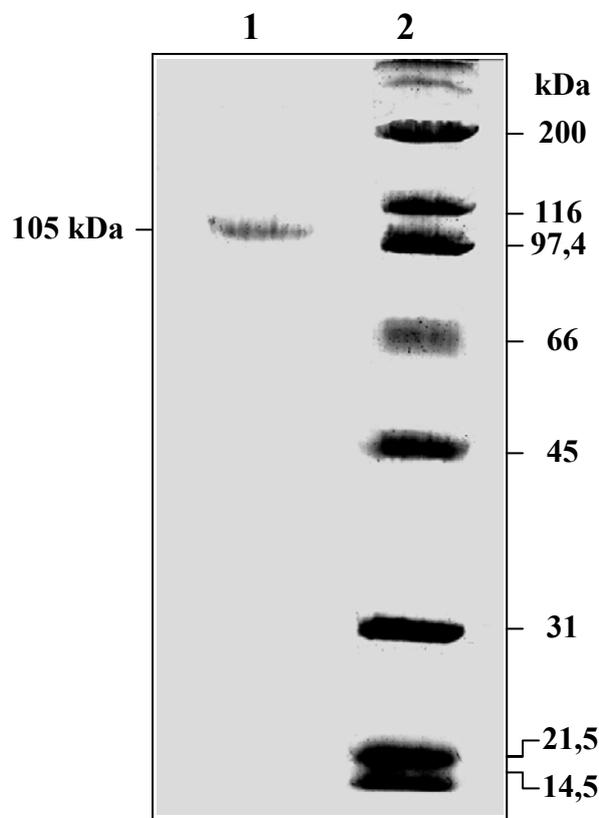


Fig. 1.

1 CAACCCAACCTCTTGAAGAAATCACAAAAAACAGTAGATTTTTCTTGAGGTACAAATAACTATTTTTTATTA
 N P T L E E I T K K T V D F F L R Y K
 76 AAGGTAATGAGAGGATATGATTTGCTAGTGAAAATATTTCAAGGAACATATTTATTATTAGAAAAATCAAAGT
 151 GCTTAAACATTTGTTCTTAAATATGAATTGAATATCAAACCTAACTATATACTAACTAAATAAATCTTTTTTATT
 226 ATGAAAAAACACTACCAACTAAAAAGTCTAATTTATGGTTCCTTATGGCTTGTTTAATAATAACCTTCCATAAG
 1 M K K T L P T K K S N L W F L M A C L I I T F H K
 301 GTAGAGGCACAAAGTACCAGATCCAAACCAAGGATTACGTGCCGAATGGATGCGAGGCGCATTGGGCATGCTTTGG
 26 V E A Q V P D P N Q G L R A E W M R G A L G M L W
 376 TTGCCCGAGCGTACTTTCAACGGTAACATTGAAGGCATAAGAATTGATGACTTTTTAACCCAAATTAAGATATT
 51 L P E R T F N G N I E G I R I D D F L T Q I K D I
 451 AGAACGGTTGATTATGTTCAACTTCCATTAACAAGCCAAATTTTTTCTCCAACCTCATGTAGCTCCACATCCT
 76 R T V D Y V Q L P L T S P N I F S P T H V A P H P
 526 ATCATTGAAAGTCTATGGCAAGGTGATACAGATGCTAATGGAGATCCAATTAATTTGGTGGCTCCACGGGAATCA
 101 I I E S L W Q G D T D A N G D P I N L V A P R E S
 XcmI EcoRI
 601 GTTGATGATCCACTTTTAAGTTGGCTCAAGGCCCTTAGGGCTGCGGGTTTAAAGAACAGAGATTTATGTGAATTC
 126 V D D P L L S W L K A L R A A G L R T E I Y V N S
 676 TATAATTTATTAGCAAGGATCCCGAAGATACGCAAGCAGATTACCCAGACGTTTCAGCACGCTGGATGGAGTGG
 151 Y N L L A R I P E D T Q A D Y P D V S A R W M E W
 HindIII
 751 TGTGATACTAACACTGAGGCTCAAGCTTTTATTAATAGCCAACTTATCATGAAGGTAATGGCAGACGTAATAC
 176 C D T N T E A Q A F I N S Q T Y H E G N G R R K Y
 826 ATGTTCTGTTATGCAGAATTTATCCTTAAAGAATATGCCCGCGTTATGGAGATTTAATTGATGCGTGGTGT
 201 M F C Y A E F I L K E Y A Q R Y G D L I D A W C F
 901 GATTCGGCAGATAATGTTATGGAAGATGAGTGTGGAGACGATCCTGCTCCGAAGATGTAACGATCAACGAATT
 226 D S A D N V M E D E C G D D P A S E D V N D Q R I
 976 TATCAGGCCTTTCAGATGCATGTCATGCAGGGAATCCGAACGCGGCCATTGCATTTAATAACAGCGTTGGT
 251 Y Q A F A D A C H A G N P N A A I A F N N S V G D
 1051 AGAGAGGGAAACCCATTCACATCAGCAACTCTTTTTGATGATTATACGTTTGGCCATCCTTTTGGCGGCGCAGGT
 276 R E G N P F T S A T L F D D Y T F G H P F G G A G
 HindIII
 1126 AACATGGTGGTTCCAGAAGCTTTATACACTTACAATCATGATTTAGTTGTGTTTCATGCAAACCAACAATGGTTAT
 301 N M V V P E A L Y T Y N H D L V V F M Q T N N G Y
 1201 GCTTTTAGAGATGATACAAGAACTTGAACGACAATGTAGTTGCGCATTTCTTTCCTAAACAGAGTACCACATCG
 326 A F R D D T R T W N D N V V A H F F P K Q S T T S
 1276 TGGAATGCAGGAAATACACCTTGTCTTACAGATGAGCAATTCGTAGAATGGACTAGTACAGGCATTGTAATGGA
 351 W N A G N T P C L T D E Q F V E W T S T G I V N G
 1351 GGTGGCATCACTTGGGGAACACCGCTTGTAGAACAAATTTGGAAAATGCTCCAGTGTCTACTTTACAACCTTAT
 376 G G I T W G T P L V R T N L E N A P V L T L Q P Y
 1426 GCGCTTAATCAGTTCCGAACCTAACAGATACTTATTTGAAAGAATTTCAATCACCAGGAAAGCCTAATTTGGTCAAGA
 401 A L N Q F E L T D T Y L K E F Q S P G K P N W S R
 1501 CAATATACAATCTTGCCAGCCATTTACCCAGGTGAGCCTTATTTCTCATAATTTGGTTGAAGGCGTAGATTTTTGG
 426 Q Y T I L P A I Y P G Q P Y S H N L V E G V D F W
 BamHI
 1576 GATCCAGAAGGAGTCGGTATAACCGGGCTTACAGCGTCTGGAACCTTACCAGCATGGTTAACAATTAGTCAAACA
 451 D P E G V G I T G L T A S G T L P A W L T I S Q T
 1651 GCTACAGGAACTTGGACGTTAAGTGAACCTCCGCTGTATCTGAGGCTTCTAATTATACCTTTGAACCTTATGGCT
 476 A T G T W T L S G T P P V S E A S N Y T F E L M A
 PstI
 1726 CAGGATAGTGATGGAGTAACAAACAGAGAAGTGAAGTTAGAGGTGATTTACATCCTGCAGGTTTACAAATCCT
 501 Q D S D G V T N R E V K L E V I S H P A G F T N P
 1801 GGAGATGGTACTCCAGTATGGTTTTCGAATCCCATGGTACTTGCAGAAACAGCATTAAAGGATTATGGAAGC
 526 G D G T P V W F S N P M V L A K A T A L K D Y G S
 1876 CTTTTGAAACTTGGTGTGATTTTTATGATTTTGAAGGAGATGTTCTTACTATAACTAAAACATCAGGACCTGAT
 551 L L K L G V D F Y D F E G D V L T I T K T S G P D

EcoRI *PvuII*

1951 TGGCTTGGTTTTAACTCAGAATTCAGATGATACTTGGCGTTTTAAGTGGGATGCCAACAGCTGCTGATGCAGGTGAA
576 W L V L T Q N S D D T W R L S G M P T A A D A G E

2026 AATTCATTTTACATTTAATGTTAGCGATGGTATTCTTTCATCAGATACAGAGATTAATAATTACAGTTGATCATGTT
601 N S F T F N V S D G I L S S D T E I K I T V D H V
L A J

2101 GCAGGTTTTACTAATTTGGGTAATGGAGCACCTGTTTGGTCTTCACCTATACTTAATCTTACCGATGGAAAAGGA
626 A G F T N L G N G A P V W S S P I L N L T D G K G

2176 TCATTGCGCTATAATTTATACTTTGCAATTAGGAACCGATTATTATGATTTTTGAAGGTGATGCACCTTACTATTACT
651 S F A Y N Y T L Q L G T D Y Y D F E G D A L T I T

2251 AAAACATCAGGACCTGATTGGCTTACTATTCAACAACCCGATGCAAAATAGTTGGAATTAAGTGGAACTCCAATA
676 K T S G P D W L T I Q Q T D A N S W K L S G T P I
L B J

2326 AATTCAGATGCGGGTGAATAATCTTTTACGTTCAATCTCAGCGATGATACCAATTCACAACCTGCTGAAATTCG
701 N S D A G E N S F T F N L S D D T N S T T A E I L

2401 ATCAACGTAATTCGAACCTATTATTAGTGATGGAATGTTGAAATTAAGCAACTGCCAATACAACCTTATGGTATT
726 I N V I A T I I S D G N V E I K A T A N T T Y G I
L C J

2476 AATACTGTTGCTAGTATGTATTCCGCAGTACAAACCGCTCCAGATGGACTAGCGACTTATAGAATTTCTATCGAC
751 N T V A S M Y S A V Q T A P D G L A T Y R I S I D

2551 GTTACACCTCCAACCTGATAAAGGTATTTATTCTGGTCTTCAGGAGGTATAACGACTACGACGTCATGGGGAATC
776 V T P P T D K G I Y S G S S G G I T T T T S W G I

2626 GGTGATGGTACAGACGCTATTCAAAATACTATATCCGAGGGAGTGATAATGAATGGACAGAAAGTATTAACAAT
801 G D G T D A I Q N T I F R G S D N E W T E S I N N

2701 ATTAATAATGGTTGATTTTAAATGCAAAATGGAGGATCTCTTACGACGGATAATGTTACCATGTTTTTAAATCTATT
826 I K M V D F N A N G G S L T T D N V T M F F K S I

EcoRI

2776 TCTATAGGGAATTCACAATCTGTAAATGATTTTGTATCGCTCAAAGTTGGAGGAGTAAATATCAAATCCAGGGAGA
851 S I G N S Q S V N D F V S L K V G G V I S N P G R

Clai

2851 TCAGCAAACCAATACGAAACCATCGATTTAACCTTAGCTACATCAGTGAGTAATATGCAAATTTTGCTATTGGA
876 S A N Q Y E T I D L T L A T S V S N I A N F A I G

2926 ACAGGTAATGATTCGGATACTAATAAATGGTCTGTTGAAGGTATAACAGTATTCGTAGACTTTCGAGGCACGCTT
901 T G N D S D T N K W S V E G I T V F V D F A G T L

3001 TCCGTTACTAATCCAATACAGGATGATGTTGACTCTTTCAGCTATATCCTAACCCAGCAAAAGATAGAATTTTC
926 S V T N P I Q D D V D S F K L Y P N P A K D R I F

HinDIII

3076 ATAACAACAACCTGTTACTGTTTCAAGTATTTTGTGTAAGTGGTAAAGTTGATTTTAAAGGAAAA
951 I N K Q P V T V Q I F D V T G K L V K V D F K G K

3151 AATGAACCTGATATTTCAACTTTAAAGCAAGGATTATATATTTGAAAATACAAACCGCTGAAGGGAATATGTTG
976 N E L D I S T L K Q G L Y I L K I Q T A E G N M L

3226 TTCGAAAAATTCATTAAGAAATAAATATCTAATAAATAGATAATAACTGTATTTATAAATTAATGATGATTG
1001 F E K F I K K M M I

3301 ATAAAAAATAAATCTAGTTGTATGTTGCCTCATGGCATCCATAGGTATTTATGCGCAAAACCAATCAAAAAGTC
D K K I I L V V C C L M A S I G I Y A Q T Q S K S

Fig. 2.

1 10 20 30 40 50
FcnA QGLRAEWMRGALGMLWLP...ERTFNNGN.TEGIRIDDFLTQIKDIRTVDYVQLPLTSPNIF
Fda1 ...KADWMQGNWGISYRIPGGDINYSGSHVVAEYNVRAAVEQISAIPLGLKVVQINLTN.GAS
Fda2 ..IRASWMQGNWGISFRISGGDISQNESHSVNEYQVAPAVEQIAAIPGLKWLQVNLNSN.GAF

60 70 80 90 100 110
FcnA SPTHVAPHPITESLWQGDTDANGDPINLVAPRESVDPLLSWTKALRAAGLRTEIYVNS..
Fda1 GDRFIVPVTEVEAIN..PLSAPNSINDLYDPTLPGRDLFEQLALAFKAKGIRVVAYIATQG
Fda2 GDRFIVPVPEVEAIN..PNSAPNSSADLFDPALPGDDLFEQIALGLQAKGIKVVAYIATQG

120 130 140 150 160
FcnA YNLLARIPEDTQADYPD.VSARWMEW.CDTNTEAQAFINSQTYHEGN.....
Fda1 PGMILKH.GAENSMDDEDSDITD.....CKSSKPLVTDLDTQVYCSANMNRWRDYVLEQYPS
Fda2 PAMILKH.GAERSMDFDSDIIVDESDGSAKSSRPVVSDDPDQVYCSANMNRWRDYVLLQYPS

170 180 190 200 210 220
FcnA .GRRKYMFCYAEFILLKEYAQRYGDLIDAWCFDSADNVMEDECDDPASEDQVNDQRITYQAFAA
Fda1 TSLYRSFELAMVNIIVETLSLRYGSTIDGWWFDHS.....GFGDSEILLHAAA
Fda2 TSLHHSFQLGLVNIIVETLSLRYGTLIDGWWFDHS.....IYGDYNLLPDA

230 240 250 260 270
FcnA DACHAGNPNAAIAFNNSVGDREGNPFTSATLFD DDTYFGHPFGGAGNMVV....PEALYTY
Fda1 LA...GNNDAAVAFN..EGDK.VPLVNNPETLDDTYFGHPT.PIGSEVSSDDKNLPLMTSI
Fda2 RA...GNSNAAVSLNL.EGD..IFLSNNPEVMEDEFTHGHP.PIARVVSSDDTNLPLMTAI

280 290 300 310 320 330
FcnA NHDLVVFMTQTNNGYAFRDDTRTWNDNVVAHFF.PKOSTTSWNAGNTPCLTDEQFVEWTSSTG
Fda1 EATLDGILT.....GSGDD.....VGSVGHMFPLQ..ESWNGG.TVVFSEAKGSDWLNRA
Fda2 EDAPNGIFT.....GTGDD.....VDALGHMFLPLQ..ETWNGG.TVVFSEAKGTWLNRV

340 350 360 370 380
FcnA IVNGGITWGTPLVVRTN.LENAPVLT LQ.P....YALNQFELTDTYLKEFQSP...
Fda1 LKAGGAF TWALSQDSNDELGGGARLISEPQVKMLERMSFNIGKQLHMNLDGSDGD
Fda2 TRAGGAL TWALS..HEGSVSGGEAMLISAPQAKMLARMQLNIGKQLDMDLDGADG.

Fig. 3.

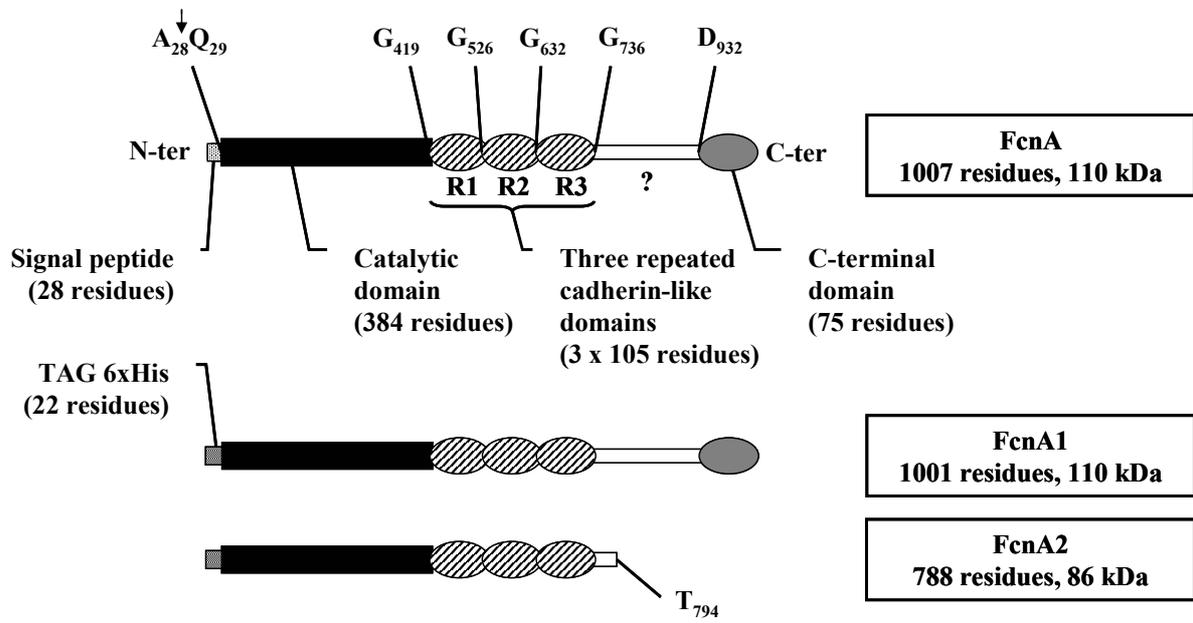


Fig. 4.

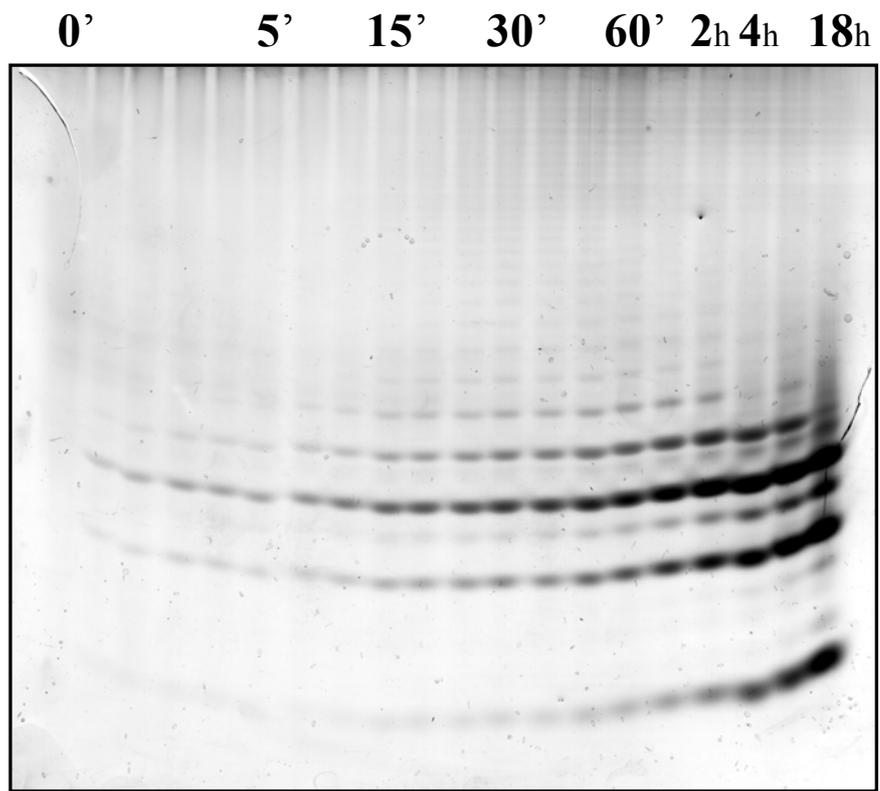


Fig. 5.

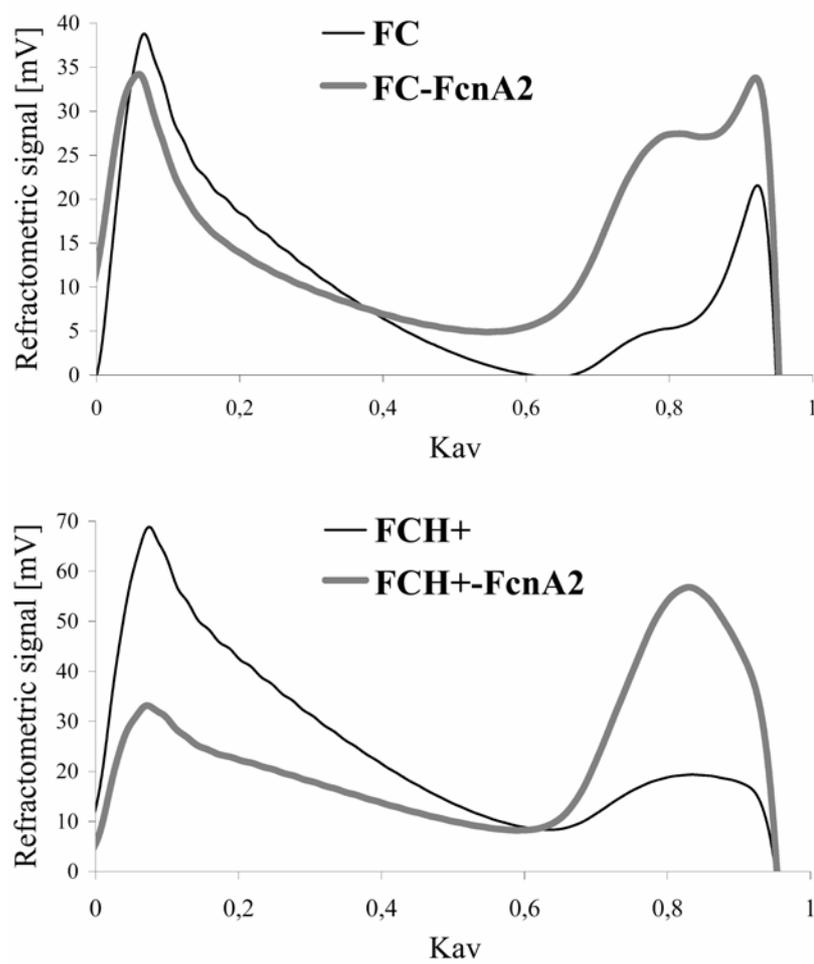


Fig. 6.

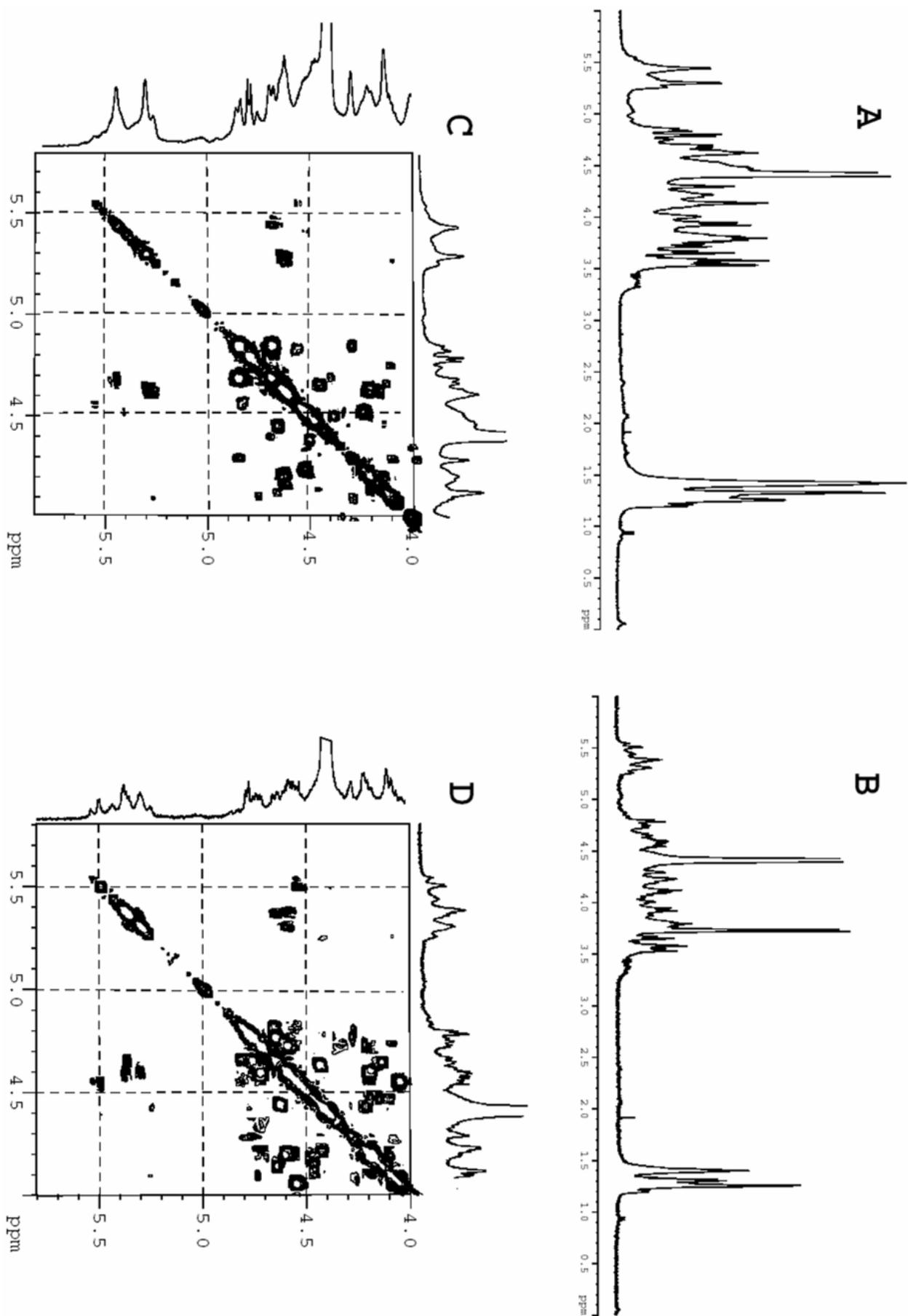


Fig. 7.

BIBLIOGRAPHIE

-A-

- Ajisaka, K, Fujimoto, H, Miyasato, M (1998) An alpha-L-fucosidase from *Penicillium multicolor* as a candidate enzyme for the synthesis of alpha (1→3)-linked fucosyl oligosaccharides by transglycosylation. *Carbohydr Res.*, **309**, 125-129.
- Akatsuka, I, and Iwamoto, K (1979) Histochemical localization of agar and cellulose in the tissue of *Gelidium pacificum* (Gelidiaceae, Rhodophyta). *Botanica Marina*, **22**, 367-370.
- Albersheim, P, and Darvill, AG (1985) Oligosaccharins. *Sci. Am.*, **253**, 44-52.
- Alberto, F, Bignon, C, Sulzenbacher, G, Henrissa, B, and Czjzek, M (2004) The three-dimensional structure of invertase (β -fructosidase) from *Thermotoga maritima* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases. *J. Biol. Chem.*, **279**, 18903-18910.
- Alexeeva, YV, Ivanova, EP, Bakunina, IY, Zvaygintseva, TN, and Mikhailov, VV (2002) Optimization of glycosidases production by *Pseudoalteromonas Issachenkonii* KMM 3549. *Lett. App. Microbiol.*, **38**, 38-42.
- Allouch, J, Helbert, W, Henrissat, B, Czjzek, M (2004) Parallel substrate binding sites in a beta-agarase suggest a novel mode of action on double-helical agarose. *Structure (Camb.)*, **12**, 623-632.
- 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.
- Alves, AP, Mulloy, B, Diniz, JA, and Mourão, PA (1997) Sulfated polysaccharides from the egg jelly layer are species-specific inducers of acrosomal reaction in sperms of sea urchins. *J. Biol. Chem.*, **272**, 6965-6971.
- Alves, AP, Mulloy, B, Moy, GW, Vacquier, VD, and Mourão, PAS (1998) Females of the sea urchin *Strongylocentrotus purpuratus* differ in the structures of their egg jelly sulfated fucans. *Glycobiology*, **8**, 939-946.
- Amimi, A, Mouradi, A, Givernaud, T, Chiadmi, N, and Lahaye, M (2001) Structural analysis of *Gigartina pistillata* carrageenans (Gigartinaceae, Rhodophyta). *Carbohydr. Res.*, **333**, 271-279.
- Anderson, NS, Campbell, JW, Harding, MM, Rees, DA, and Samuel, JW (1969) X-ray diffraction studies of polysaccharide sulphates: double helix models for k- and l-carrageenans. *J. Mol. Biol.*, **45**, 85-99.
- Anderson, NS, Dolan, TCS, and Rees, DA (1965) Evidence for a common structural pattern in the polysaccharides sulfates of the Rhodophyceae. *Nature*, **205**, 1060-1062.
- Anderson, NS, Dolan, TCS, Lawson, CJ, and Rees DA (1968) The carrageenan Part V: the masked repeating structures of λ - and μ -carrageenan. *Carbohydr. Res.*, **7**, 468-473.

- Andrade, LR, Salgado, LT, Farina, M, Pereira, MS, Mourão, PA, and Amado Filho, GM (2004) Ultrastructure of acidic polysaccharides from the cell walls of brown algae. *J. Struct. Biol.*, **145**, 216-225.
- Aquino, RS, Landeira-Fernandez, AM, Valente, AP, Andrade, LR, and Mourão, PA. (2005) Occurrence of sulfated galactans in marine angiosperms: evolutionary implications. *Glycobiology*, **15**, 11-20.
- Azam, F, and Long, RA (2001) Sea snow microcosm. *Nature*, **414**, 495-498.
- Azam, F, and Worden, AZ (2004) Microbes, molecules and marine ecosystems. *Science*, **303**, 1622-1624.

-B-

- Bailey ST (1955) X-ray and infrared studies on carragenin. *Biochim. Biophys. Acta*, **17**, 195-205.
- Baker, B, Zambryski, P, Staskawicz, B, and Dinesh-Kumar, SP (1997) Signaling in plant-microbe interactions. *Science*, **276**, 726-733.
- Bakunina, IY, Nedashkovskaya, OI, Alekxeeva, SA, Ivanova, EP, Romanenko, LA, Gorshkova, NM, Isakov, VV, Zvaygintseva, TN, and Mikhailov, VV (2002) Degradation of Fucoidan by the Marine Proteobacterium *Pseudoalteromonas citrea*. *Microbiology (Moscow)*, **71**, 41-47.
- Bakunina, IY, Shevchenko, LS, Nedashkovskaya, OI, Shevchenko, NM, Alekxeeva, SA, Mikhailov, VV, and Zvaygintseva, TN (2000) Screening of marine bacteria for fucoidanases. *Microbiology (Moscow)*, **69**, 370-376.
- Baldauf SL, Roger AJ, Wenk-Siefert I, and Doolittle WF. (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science*, **290**, 972-977.
- Barbeyron, T, Gerard, A, Potin, P, Henrissat, B, and Kloareg, B (1998) The kappa-carrageenase of the marine bacterium *Cytophaga drobachiensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.*, **15**, 528-537.
- Barbeyron, T, Henrissat, B, and Kloareg, B (1994) The gene encoding the kappa-carrageenase of *Alteromonas carrageenovora* is related to beta-1,3-1,4-glucanases. *Gene*, **139**, 105-109.
- Barbeyron, T, L'Haridon, S, Corre, E, Kloareg, B, and Potin, P (2001) *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of Flavobacteriaceae isolated from a red alga, and classification of [*Cytophaga*] *uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.*, **51**, 985-997.
- Barbeyron, T, Michel, G, Potin, P, Henrissat, B, Kloareg, B (2000) Iota-Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of kappa-carrageenases. *J. Biol. Chem.*, **275**, 35499-35505.
- Bartnicki-Garcia, S (1968) Cell wall chemistry, morphogenesis, and taxonomy of Fungi. *Ann. Rev.*, **22**, 87-108.

- Bayer, EA, Belaich, JP, Shoham, Y, Lamed, R (2004) The Cellulosomes: Multienzyme Machines for Degradation of Plant Cell Wall Polysaccharides. *Annu. Rev. Microbiol.*, **58**, 521-554.
- Bellion, C, Hamer, GK, and Yaphe W (1982) The degradation of *Euchema spinosum* and *Euchema cottonii* carrageenans by ι -carrageenases and κ -carrageenases from marine bacteria. *Can. J. Microbiol.*, **28**, 874-880.
- Berteau, O, McCort, I, Gouasdoue, N, Tissot, B, and Daniel, R (2002) Characterization of a new α -L-fucosidase isolated from the marine mollusk *Pecten maximus* that catalyzes the hydrolysis of α -L-fucose from algal fucoidan (*Ascophyllum nodosum*). *Glycobiology*, **12**, 273-282.
- 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. *Glycobiol.*, **13**, 29-40.
- Bilan, MI, Grachev AA, Ustuzhanina, NE, Shashkov, AS, Nifantiev, NE, and Usov, AI (2004) A highly regular fraction of a fucoidan from the brown seaweed *Fucus distichus* L.. *Carbohydr. Res.*, **339**, 511-517.
- Bilan, MI, Grachev, AA, Ustuzhanina, NE, Shashkov, AS, Nifantiev, NE, and Usov, AI (2002) Structure of a fucoidan from the brown seaweed *Fucus evanescens* C. Ag.. *Carbohydr. Res.*, **337**, 719-730.
- Bouarab, K, Potin, P, Correa, J, and Kloareg, B (1999) Sulfated oligosaccharides mediate the interaction between a marine red alga and its green algal pathogenic endophyte. *Plant Cell*, **11**, 1635-1650.
- Boucher, Y, Douady, CJ, Papke, RT, Walsh, DA, Boudreau, ME, Nesbo, CL, Case, RJ, and Doolittle, WF (2003) Lateral gene transfer and the origins of prokaryotic groups. *Annu Rev Genet.*, **37**, 283-328.
- Bourin, MC, and Lindahl, U. (1993) Glycosaminoglycans and the regulation of blood coagulation. *Biochem. J.*, **289**, 313-330.
- Bourne, Y, and Henrissat, B (2001) Glycosides hydrolases and glycosyltransferases: families and functional modules. *Curr. Opin. Struct. Biol.*, **11**, 593-600.
- Bourquelot, E and Bridel, M (1913) Synthèse des glucosides d'alcools à l'aide de l'émulsine et réversibilité des actions fermentaires. *Ann. de Chim. et de Phys.*, **8**, 145-217.
- Braam, J. (1999) If walls could talk. *Curr. Opin. Plant. Biol.*, **2**, 521-524.

-C-

- Caceres, PJ, Carlucci, MJ, Damonte, EB, Matsuhiro, B, and Zuniga, EA (2000) Carrageenans from Chilean samples of *Stenogramme interrupta* (Phylloporaceae): structural analysis and biological activity. *Phytochemistry*, **53**, 81-86.
- Carlucci, MJ, Pujol, CA, Ciancia, M, Nosedá, MD, Matulewicz, MC, Damonte, EB, and Cerezo, AS. (1997) Antiherpetic and anticoagulant properties of carrageenans from the red seaweed

Gigartina skottsbergii and their cyclized derivatives: correlation between structure and biological activity. *Int. J. Biol. Macromol.*, **20**, 97-105.

- Carlucci, MJ, Scolaro, LA, and Damonte, EB (1999) Inhibitory action of natural carrageenans on herpes simplex virus infection of mouse astrocytes. *Chemotherapy*, **45**, 429-436.
- Chabeau, P (1995) Purification et clonage de la λ -carrageenases de *Pseudoalteromonas carrageenovora*. Rapport de DEA de l'Université Technologique de Compiègne.
- Chaubet, F, Chevolut, L, Jozefonvicz, J, Durand, P, and Boisson-Vidal, C (2000). Relationships between chemical characteristics and anticoagulant activity of low molecular weight fucans from marine algae. In *Bioactive Carbohydrate Polymers* (B.S. Paulsen Ed.) Kluwer Academic Publisher. **6**, 59-83.
- Chevolut, L, Foucault, A, Chaubet, F, Kervarec, N, Siquin, C, Fisher, AM, and Boisson-Vidal, C (1999) Further data on the structure of brown seaweed fucans: relationships with anticoagulant activity. *Carbohydr. Res.*, **319**, 154-165.
- Chevolut, L, Mulloy, B, Ratiskol, J, Foucault, A, and Collic-Jouault, S (2001) A disaccharide repeat unit is the major structure in fucoidans from two species of brown algae. *Carbohydr. Res.*, **330**, 529-535.
- Chizhov, AO, Dell, A, Morris, HR, Haslam, SM, Mc Dowell, RA, Shashkov, AS, Nifant'ev, NE, Khatuntseva, EA and Usov AI (1999) A study of fucoidan from seaweed *Chorda filum*. *Carbohydr. Res.*, **320**, 108-119.
- Church, FC, Meade, JB, Treanor, RE, and Whinna, HC (1989) Antithrombin activity of fucoidan. The interaction of fucoidan with heparin cofactor II, antithrombin III, and thrombin. *J. Biol. Chem.*, **264**, 3618-3623.
- Ciancia, M, Nosedà, MD, Matulewicz, MC, and Cerzo AS (1993) Alkali-modification of carrageenans: mechanism and kinetics in the kappa/iota, mu/nu and lambda-series. *Carbohydr. Polym.*, **20**, 95-98.
- Clark, AH, and Ross-Murphy, SB (1987) Structural and mechanical properties of biopolymer gels. *Adv. Polym. Sci.*, **83**, 57-192.
- Coutinho, PM, and Henrissat, B (1999) Carbohydrate-active enzymes: an integrated database approach. In "Recent Advances in Carbohydrate Bioengineering", H.J. Gilbert, G. Davies, B. Henrissat and B. Svensson eds., The Royal Society of Chemistry, Cambridge, pp. 3-12.
- Craigie, JS (1990) Cell walls. In : *Biology of the red algae*, 221-258, Ed. KM Cole and RG Sheath, Cambridge University Press, Cambridge.
- Craigie, JS, and Wong, KF (1978) Carrageenan biosynthesis. *Proc. Int. Seaweed Symp.*, **9**, 595-609.
- Crout, DHG, and Vic, G (1998) Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. *Curr. Opin. Chem. Biol.*, **2**, 98-111.

-D-

- Daniel, R, Berteau, O, Chevolut, L, Varenne, A, Gareil, P, and Gouasdoue, N (2001) Regioselective desulfation of sulfated L-fucopyranoside by a new sulfoesterase from the marine mollusk *Pecten maximus*. *Eur. J. Biochem.*, **268**, 5617-5626.
- Daniel, R, Berteau, O, Jozefonvicz, J, and Goasdoue, N (1999) Degradation of algal (*Ascophyllum nodosum*) fucoidan by an enzymatic activity contained in digestive glands of the marine mollusk *Pecten maximus*. *Carbohydr. Res.*, **322**, 291-297.
- Davies, GJ, and Henrissat, B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure*, **3**, 853-859.
- De Reviere (2003) Biologie et phylogénie des algues. Vol. 1 et 2, Belin, Paris
- Descamps-Lepagnol, V, Klarzynski, O, Barbeyron, T, Kloareg, B, Fritig, B, Joubert, JM, Plesse, B, and Yvin, JC (1998) Fuco-oligosaccharides, enzyme pour leur préparation à partir des fucanes, bactérie productrice de l'enzyme et applications des fuco-oligosaccharides à la protection des plantes. Patent n° FR9811756, n° WOFR9902243.
- Descamps-Lepagnol, V (1998) Purification et caractérisation d'enzymes de dégradation de polysaccharides des algues brunes pour la production d'oligosaccharides actifs. Rapport de thèse de l'Université de Technologies de Compiègne.
- Dickens, NJ, Beatson, S, and Ponting, CP (2002) Cadherin-like domains in alpha-dystroglycan, alpha/epsilon-sarcoglycan and yeast and bacterial proteins. *Curr. Biol.*, **12**, 197-199.
- Dolan, TCS, and Rees, DA (1965) The carrageenan Part II : the positions of the glycosidic linkages and sulfate esters in λ -carrageenan. *J. Chem. Soc.*, **33**, 3534-3539.
- Doolittle, WF (1999) Phylogenetic Classification and the Universal Tree. *Science*, **284**, 2124-2128.
- Draget KI, Strand B, Hartmann M, Valla S, Smidsrod O, Skjak-Braek G. (2000) Ionic and acid gel formation of epimerised alginates; the effect of AlgE4. *Int. J. Biol. Macromol.*, **27**, 117-122.
- Draget, KI, Skjak-Braek, G, and Smidsrod, O. (1997) Alginate based new materials. *Int. J. Biol. Macromol.*, **21**, 47-55.
- Draget, KI, Smidsrod, O, and Skjak-Braek, G (2003) Alginate from algae. In : Biopolymer vol.6 : Polysaccharides II, polysaccharides from Eukaryotes, pp 215-248, ed. Sophie De Baets, Eric Vandamme and Alexander Steinbüchel, Wiley-VCH, Weinheim.
- Duarte, MER, Cardoso, MA, Nosedá, MD, and Cerezo, AS (2001) Structural studies on fucoidans from the brown seaweed *Sargassum stenophyllum*. *Carbohydr. Res.* **333**, 281-293.

-E-

- Elias, CJ, Coggins, C, Alvarez, F, Brache, V, Fraser, IS, Lacarra, M, Lahteenmaki, P, Massai, R, Mishell, DR, Phillips, DM, and Salvatierra, AM (1997) Colposcopic evaluation of a vaginal gel formulation of iota-carrageenan. *Contraception*, **56**, 387-389.

- Engel, A, Thoms, S, Riebesell, U, Rochelle-Newall, E, and Zondervan, I (2004) Polysaccharide aggregation as a potential sink of marine dissolved organic carbon. *Nature*, **428**, 929-932.

-F-

- Falshaw, R, and Furneaux, RH (1994) Carrageenan from the tetrasporic stage of *Gigartina decipiens* (Gigartinaceae, Rhodophyta). *Carbohydr. Res.*, **252**, 171-182.
- Falshaw, R, and Furneaux, RH (1995) Carrageenans from the tetrasporic stages of *Gigartina clavifera* and *Gigartina alveata* (Gigartinaceae, Rhodophyta). *Carbohydr. Res.*, **276**, 155-165.
- Falshaw, R, and Furneaux, RH (1998) Structural analysis of carrageenans from the tetrasporic stages of the red algae, *Gigartina lanceata* and *Gigartina chapmanii* (Gigartinaceae, Rhodophyta). *Carbohydr. Res.*, **307**, 325-331.
- Falshaw, R, Furneaux, RH, and Slim, GC (1999) Carbohydrate sulfates. In : Carbohydrates : structures, synthesis and dynamics, 107-149, ed. P. Finch, Kluwer Academic Publishers, Dordrecht.
- Farias, WR, Valente, AP, Pereira, MS, and Mourão PA (2000) Structure and anticoagulant activity of sulfated galactans. Isolation of a unique sulfated galactan from the red algae *Botryocladia occidentalis* and comparison of its anticoagulant action with that of sulfated galactans from invertebrates. *J. Biol. Chem.*, **275**, 29299-29307.
- Farkas, V (1979) Biosynthesis of cell walls of Fungi. *Microbiol. Rev.*, **43**, 117-144.
- Frei, E, and Preston, RD (1961) Variants in structural polysaccharides of algal cell walls. *Nature (Lond.)*, **192**, 939-943.
- Fulop, V, and Jones, DT (1999) Beta propellers: structural rigidity and functional diversity. *Curr. Opin. Struct. Biol.*, **9**, 715-721.
- Furukawa, S, Fujikawa, T, Koga, D, and Ide, A (1992 A) Production of fucoidan-degrading enzymes, fucoïdanase and fucoidan sulfatase by *Vibrio* sp. N-5. *Nippon Suisan Gakkaishi*, **58**, 1499-1503.
- Furukawa, S, Fujikawa, T, Koga, D, and Ide, A (1992 B) Purification and some properties of exo-type fucoïdanases from *Vibrio* sp. N-5. *Biosc. Biotechnol. Biochem.*, **56**, 1829-1834.

-G-

- Ginalski, K, Elofsson, A, Fischer, D, and Rychlewski, L (2003) 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics*, **19**, 1015-1018.
- Girond, S, Crance, JM, Van Cuyck-Gandre, H, Renaudet, J, and Deloince, R (1991) Antiviral activity of carrageenan on hepatitis A virus replication in cell culture. *Res. Virol.*, **224**, 261-270.
- Glockner, FO, Kube, M, Bauer, M, Teeling, H, Lombardot, T, Ludwig, W, Gade, D, Beck, A, Borzym, K, Heitmann, K, Rabus, R, Schlesner, H, Amann, R, and Reinhardt, R (2003) Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc Natl Acad Sci U S A*, **100**, 8298-8303.

- Gonzalez, JM, and Weiner, RM (2000) Phylogenetic characterization of marine bacterium strain 2-40, a degrader of complex polysaccharides. *Int. J. Syst. Evol. Microbiol.*, **8**, 831-834.
- Gordon-Mills, EM and McCandless, EL (1975) Carrageenans in the cell walls of *Chondrus crispus* Stack. III. Localization with fluorescent antibody. *Phycologia*, **17**, 95-104.
- Greer, CW (1984). A study of carrageenases from marine bacteria. PhD thesis, McGill University, Montreal. 208 pp.
- Greer, CW, and Yaphe, W (1984) Purification and properties of ι -carragénase from a marine bacterium. *Can. J. Microbiol.*, **30**, 1500-1506.
- Guruprasad, K, and Dhamayanthi, P (2004) Structural plasticity associated with the small beta Greek-propeller architecture. *Int. J. Biol. Macromol.*, **34**, 55-61.

-H-

- Haroun-Bouhedja, F, Ellouali, M, Sinquin, C, and Boisson-Vidal, C (2000) Relationship between sulfate groups and biological activities of fucans. *Thromb. Res.*, **100**, 453-459.
- Harvey, DJ (1999) Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. *Mass Spectrom. Rev.*, **18**, 349-451.
- Haug, A, Larsen, B, and Smidsrød, O (1974) Uronic acid sequence in alginate from different sources. *Carbohydr. Res.*, **32**, 217-225.
- Henrissat, B (1991) A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.*, **280**, 309-316.
- Henrissat, B, and Bairoch, A (1993) New families in the classification of glycosyl hydrolases based on amino- acid sequence similarities. *Biochem. J.*, **293**, 781-788.
- Henrissat, B, and Bairoch, A (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.*, **316**, 695-696.
- Henrissat, B, and Davies, GJ (1997) Structural and sequence-based classification of glycosyl hydrolases. *Curr. Opin. Struct. Biol.*, **7**, 637-644.
- Hoffman, R (1993) Carrageenans inhibit growth-factor binding. *Biochem. J.*, **289**, 331-334.
- Hoffman, R, burns, WW and paper, DH (1995) Selective inhibition in cell proliferation and DNA synthesis by the polysulfated carbohydrate λ -carrageenan. *Cancer. Chemother. Pharmacol.*, **36**, 325-334.
- Hoshino, T, Hayashi, T, Hayashi, K, Hamada, J, Lee, JB, and Sankawa, U (1998) An antivirally active sulfated polysaccharide from *Sargassum horneri* (TURNER) C. AGARDH. *Biol. Pharm. Bull.*, **21**, 730-734.
- Howard, MB, Ekborg, NA, Taylor, LE, Weiner, RM, and Hutcheson, SW (2004) Chitinase B of "*Microbulbifer degradans*" 2-40 contains two catalytic domains with different chitinolytic activities. *J. Bacteriol.*, **186**, 1297-1303.

- Howard, MB, Ekborg, NA, Taylor, LE, Weiner, RM, and Hutcheson SW. (2003) Genomic analysis and initial characterization of the chitinolytic system of *Microbulbifer degradans* strain 2-40. *J. Bacteriol.*, **185**, 11, 3352-3360.
- Hsu, HY, Hajjar, DP, Khan, KM, and Falcone, DJ (1998) Ligand binding to macrophage scavenger receptor-A induces urokinase-type plasminogen activator expression by a protein kinase-dependent signaling pathway. *J. Biol. Chem.*, **273**, 1240-1246.

-I-

- Imberty, A, and Perez, S (2000) Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem. Rev.*, **100**, 4567-4588.
- Ivanova, EP, Bakunina, IY, Nedashkovskaya, OI, Gorshkova, NM, Alexeeva, YV, Zelepuga, EA, Zvaygintseva, TN, Nicolau, DV, and Mikhailov, VV (2003) Ecophysiological variabilities in ectohydrolytic enzyme activities of some *Pseudoalteromonas* species, *P. citrea*, *P. issachenkonii*, and *P. nigrifaciens*. *Curr. Microbiol.*, **46**, 6-10.

-J-

- 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 paralog enzymes with different molecular organizations and catalytic behavior. *Biochem. J.*, **385**, 703-713.
- Janaswamy, S, and Chandrasekaran, R (2001) Three-dimensional structure of the sodium salt of iota-carrageenan. *Carbohydr. Res.*, **335**, 181-194.
- Janaswamy, S, and Chandrasekaran, R (2002) Effect of calcium ions on the organization of iota-carrageenan helices: an X-ray investigation. *Carbohydr. Res.*, **337**, 523-535
- Jawad, Z, and Paoli, M (2002) Novel sequences propel familiar folds. *Structure (Camb.)*, **10**, 447-454.
- Johansson, E, Hedbys, L, and Larsson, PO (1986) Synthesis of mannose oligosaccharides via reversal of the α -mannosidase reaction. *Biotechnol. Lett.*, **8**, 421-424.
- Johnston, KH, and McCandless, EL (1973) Enzymic hydrolysis of the potassium chloride soluble fraction of carrageenan: properties of " λ -carrageenase" from *Pseudomonas carrageenovora*, *Can. J. Microbiol.*, **19**, 779-788.

-K-

- Kara, S, Tamerler, C, Bermek, H, and Pekcan, O (2003) Cation effects on sol-gel and gel-sol phase transitions of kappa-carrageenan-water system. *Int. J. Biol. Macromol.*, **31**, 177-185.

- Kariya, Y, Mulloy, B, Imai, K, Tominaga, A, Kaneko, T, Asari, A, Suzuki, K, Masuda, H, Kyogashima, M, and Ishii, T (2004) Isolation and partial characterization of fucan sulfates from the body wall of sea cucumber *Stichopus japonicus* and their ability to inhibit osteoclastogenesis. *Carbohydr. Res.*, **339**, 1339-1346.
- Karplus, K, Barrett, C, and Hughey, R (1998) Hidden Markov Models for Detecting Remote Protein Homologies. *Bioinformatics*, **14**, 846-856.
- Kertesz, MA (1999) Riding the sulfur cycle-metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiol. Rev.*, **24**, 135-175.
- Kidby, DK, and Davidson, DJ (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.*, **55**, 321-325.
- Kitamura, K, Matsuo, M, and Yasui, T (1992) Enzymatic degradation of fucoidan by fucoidanase from the hepatopancreas of *Patinopecten yessoensis*. *Biosci. Biotech. Biochem.*, **56**, 490-494.
- Klarzynski, O, Descamps, V, Plesse, B, Yvin, JC, Kloareg, B, and Fritig, B (2003) Sulfated fucan oligosaccharides elicit defense responses in tobacco and local and systemic resistance against tobacco mosaic virus. *Mol. Plant Microb. Interact.*, **16**, 115-122.
- Kloareg, B, and Quatrano, RS (1988) Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanogr. Mar. Biol. Annu. Rev.*, **26**, 259-315.
- Kloareg, B, Demarty, M, and Mabeau, S (1986) Polyanionic characteristic of purified sulfated homofucans from brown algae. *Int. J. Biol. Macromol.*, **8**, 380-386.
- Knutsen, SH, Myslabodski, DE, and Grasdalen H (1994) A modified system of nomenclature for red algal galactans. *Bot. Mar.*, **37**, 163-169.
- Kobata, A (1979) Use of endo et exoglycosidases for structural studies of glycoconjugates. *Anal. Biochem.*, **100**, 1-14.
- Kobayashi, S (1999) Enzymatic polymerization: a new method of polymer synthesis. *J. Polym. Sci. Part. A*, **37**, 3041-3056.
- Kosak, M (1999) Initiation of translation in prokaryotes and eucaryotes. *Gene*, **234**, 187-208.
- Kreger, D. R. (1962) Physiology and biochemistry of algae, pp 315-335, ed. R. A. Lewin, Academic Press, New York.
- Kusaykin, MI, Burtseva, YV, Svetasheva, TG, Sova, VV, and Zvyagintseva, TN (2003) Distribution of O-Glycosylhydrolases in Marine Invertebrates. Enzymes of the Marine Mollusk *Littorina kurila* that Catalyze Fucoidan Transformation. *Biochemistry (Moscow)*, **68**, 317-324.
- Kylin, H (1913) Zur biochemie der Meersalgen. *Z. Physiol. Chem.*, **83**, 171-197.

-L-

- La Clair, JW, and Dawes, CJ (1976) An autoradiographic and histochemical localization of sulfated polysaccharides in *Euchema nudum*. *J. Phycol.*, **12**, 368-375.

- Larsen, B, Haug, A, and Painter, T (1966) Sulphated polysaccharides in brown algae. I. Isolation and preliminary characterization of three sulfated polysaccharides from *Ascophyllum nodosum* (L.) Le JOL.. *Acta Chem. Scand.*, **20**, 219-230.
- Larsen, B, Salem, DM, Sallam, MA, Mishrikey, MM, and Beltagy, AI. (2003) Characterization of the alginates from algae harvested at the Egyptian Red Sea coast. *Carbohydr. Res.*, **338**, 2325-2336.
- Lawrence, JG, and Roth, JR (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics*, **143**, 1843–1860.
- Le Gall, Y, Braud, JP, and Kloareg, B (1990) Protoplast production in *Chondrus crispus* gametophytes (Gigartinales, rhodophyta). *Plant Cell Reports*, **8**, 582-585.
- Le Questel, JY, Cros, S, Mackie, W, and Perez, S (1995) Computer modeling of sulfated carbohydrates: applications to carrageenans. *Int. J. Biol. Macromol.*, **17**, 161-175.
- Logeart, D, Prigent-Richard, S, Jozefonvicz, J, and Letourneur, D (1997) Fucans, sulfated polysaccharides extracted from brown seaweeds, inhibit vascular smooth muscle cell proliferation. I. Comparison with heparin for antiproliferative activity, binding and internalization. *Eur. J. Cell Biol.*, **74**, 376-384.

-M-

- Mackenzie, LF, Wang, Q, Warren, RAJ, and Withers, SG (1998) Glycosynthases: mutant glycosidases for oligosaccharides synthesis. *J. Am. Chem. Soc.*, **120**, 5583-5584.
- Mackie, W, and Preston, RD (1974) Cell wall and intercellular region polysaccharides. In : *Algal Physiology and Biochemistry*, 40-85, eds WDP. Stewart, Blackwell Scientific, Oxford.
- Marais, MF, and Joseleau, JP (2001) A fucoidan fraction from *Ascophyllum nodosum*. *Carbohydr. Res.*, **336**, 155-159.
- Matulewicz, MC, Ciancia, M, Nosedá, MD, and Cerezo, AS (1989) Carrageenan systems from tetrasporic and cystocarpic stages of *Gigartina skottsbergii*. *Phytochemistry*, **28**, 2937-2941.
- Matulewicz, MC, Ciancia, M, Nosedá, MD, and Cerezo, AS (1990) Methylation analysis of carrageenan from tetrasporic and cystocarpic stages of *Gigartina skottsbergii*. *Phytochemistry*, **29**, 3407-3410.
- McCandless, EL, and Craigie, JS (1979) Sulfated polysaccharides in red and brown algae. *Ann. Rev. Plant Physiol.*, **30**, 41-53.
- McCandless, EL, and Gretz, MR (1984) Biochemical and immunochemical analysis of carrageenans of the Gigartinales and Phylloporaceae. *Hydrobiologia*, **116/117**, 175-178.
- McCandless, EL, Craigie, JS, and Walter, JA (1979) Carrageenans in the gametophytic and sporophytic stages of *Chondrus crispus*. *Planta*, **112**, 201-212.
- McCandless, EL, West, JA, and Guiry, MD (1983) Carrageenans pattern in the Gigartinales. *Bioch. Syst. Ecol.*, **11**, 175-182.

- McClure, MO, Moore, JP, Blanc, DF, Scotting, P, Cook, GM, Keynes, RJ, Weber, JN, Davies, D, and Weiss, RA (1992) Investigations into the mechanism by which sulfated polysaccharides inhibit HIV infection in vitro. *AIDS Res. Hum. Retroviruses*, **8**, 19-26.
- McGuffin, LJ, Bryson, K, and Jones, DT (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404-405.
- McHugh, DJ (2003) A guide to the seaweed industry. *FAO Fisheries Technical Paper*, n°441, Food and Agriculture Organization, Rome.
- McLean, MW, and Williamson, FB (1979 A) Kappa-carrageenase from *Pseudomonas carrageenovora*. *Eur. J. Biochem.*, **93**, 553-558.
- McLean, MW, and Williamson, FB (1979 B) Glycosulphatase from *Pseudomonas carrageenovora*, purification and some properties. *Eur. J. Biochem.*, **101**, 497-505.
- McLean, MW, and Williamson, FB (1981) Neocarratetraose 4-O-monosulfate β -hydrolase from *Pseudomonas carrageenovora*. *Eur. J. Biochem.*, **113**, 447-456.
- McNeely, D (1959) Industrial Gums, pp117-125, eds RL Whistler, Academic Press, New-York.
- Medcalf, DG, Root, CF, Craney, CL, Mukhopadhyay, D, Miller, CK, and Hopewell, WD (1972). Chemical characterization of mucilaginous polysaccharides from Ulvaceae species native to Puget Sound. *Proc. Int. Seaweed Symp.*, **7**, 541-547.
- Melo, FR, Pereira, MS, Foguel, D, and Mourão PA (2004) Antithrombin-mediated anticoagulant activity of sulfated polysaccharides: different mechanisms for heparin and sulfated galactans. *J. Biol. Chem.*, **279**, 20824-20835.
- Menard, R, Alban, S, de Ruffray, P, Jamois, F, Franz, G, Fritig, B, Yvin, JC, and Kauffmann, S. (2004) Beta-1,3 glucan sulfate, but not beta-1,3 glucan, induces the salicylic acid signaling pathway in tobacco and Arabidopsis. *Plant Cell.*, **16**, 3020-3032.
- Mercier, L, Lafitte, C, Borderies, G, Briand, X, Esquerré-Tugayé, MT, and Fournier, J (2001). The algal polysaccharide carrageenans can act as an elicitor of plant defence. *New Phytol.*, **149**, 43-51.
- Michel, G, Chantalat, L, Duee, E, Barbeyron, T, Henrissat, B, Kloareg, B, and Dideberg, O (2001 A) The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure (Camb.)*, **9**, 513-525.
- Michel, G, Chantalat, L, Fanchon, E, Henrissat, B, Kloareg, B, and Dideberg, O (2001 B) The iota-carrageenase of *Alteromonas fortis*. A beta-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J. Biol. Chem.*, **276**, 40202-40209.
- Michel, G, Helbert, W, Kahn, R, Dideberg, O, and Kloareg, B (2003) The structural bases of the processive degradation of iota-carrageenan, a main cell wall polysaccharide of red algae. *J. Mol. Biol.*, **334**, 421-433.
- Molin, S, and Tolker-Nielsen, T (2003) Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Curr. Opin. Biotech.*, **14**, 255-261.

- Mori, J (1939) Carbohydrase acting on the mucilage from *Chondrus ocellatus* I. *J. Agr. Soc. Japan*, **15**, 1070-1074.
- Mori, J (1943) Carbohydrase acting on the mucilage from *Chondrus ocellatus* II. Purification, unit determination and distribution of enzyme. *J. Agr. Soc. Japan*, **19**, 740-742.
- Morigana, T, Araki, T, Ito, M, and Kitamikado, M (1981) A search for fucoidan-degrading bacteria in coastal sea environment of Japan. *Bull. Japanese Soc. Scientific Fisheries*, **47**, 621-625.
- Mourão, PA, and Bastos, IG (1987) Highly acidic glycans from sea cucumbers. Isolation and fractionation of fucose-rich sulfated polysaccharides from the body wall of *Ludwigothurea grisea*. *Eur. J. Biochem.*, **166**, 639-645.
- Mourão, PA, and Pereira MS (1999) Searching for alternatives to heparin: sulfated fucans from marine invertebrates. *Trends Cardiovasc. Med.*, **9**, 225-232.
- Mulloy, B, Mourão, PA, and Gray, E (2000) Structure/function of anticoagulant sulphated polysaccharides using NMR. *J. Biotech.*, **77**, 123-135.
- Mulloy, B, Ribeiro, AC, Alves, AP, Vieira, RP, and Mourão, PAS (1994) Sulfated fucans from echinoderms have a regular tetrasaccharide repeating unit defined by specific patterns of sulfation at the O-2 and O-4 positions. *J. Biol. Chem.*, **269**, 22113-22123.

-N-

- Nagano, N, Orengo, CA, and Thornton, JM (2001) The $(\beta/\alpha)_8$ glycosidases: sequences and structure analyses suggest distant evolutionary relationships. *Prot. Eng.*, **14**, 845-855.
- Nagano, N, Orengo, CA, and Thornton, JM (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.*, **321**, 741-765.
- Nagaoka, M, Shibata, H, Kimura-Takagi, I, Hashimoto, S, Kimura, K, Makino, T, Aiyama, R, Ueyama, S, and Yokokura, T (1999) Structural study of fucoidan from *Cladosiphon okamuranus* TOKIDA. *Glycoconj. J.*, **16**, 19-26.
- Nimchuk, Z, Eulgem, T, Holt, BF, and Dangl JL (2003) Recognition and response in the plant immune system. *Annu. Rev. Genet.*, **37**, 579-609.
- Nishino, T, Fukuda, A, Nagumo, T, Fujihara, M, and Kaji E (1999) Inhibition of the generation of thrombin and factor Xa by a fucoidan from the brown seaweed *Ecklonia kurome*. *Thromb. Res.*, **96**, 37-49.
- Nurizzo, D, Turkenburg, JP, Charnock, SJ, Roberts, SM, Dodson, EJ, McKie, VA, Taylor, EJ, Gilbert, HJ, and Davies, GJ (2002) *Cellvibrio japonicus* alpha-L-arabinanase 43A has a novel five-blade beta-propeller fold. *Nat. Struct. Biol.*, **9**, 665-668.

-O-

- O'Neill, AN (1955) Derivatives of 4-O- β -D-galactopyranosyl-3,6-anhydro-D-galactose from κ -carrageenan. *J. Am. Chem. Soc.*, **77**, 6324-6326.

- Ochman, H, Lawrence, JG, and Groisman, EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299-305.

-P-

- Painter, TJ (1983) Algal polysaccharides In : The polysaccharides, vol. 2, 195-285, ed. G. O. Aspinall, Academic Press, New York.
- Parish, CR, Coombe, DR, Jakobsen, KB, Bennet, FA, and Underwood, PA (1987) Evidences that sulfated polysaccharides inhibit tumor metastasis by blocking tumor-cell-derived heparanases. *Int. J. Cancer.*, **40**, 511-518.
- Parolis, H (1978) Sulfated polysaccharides of the Grateloupiaceae family. Part 10. The structure of the polysaccharide of *Pachymenia carnosa*. *Carbohydr. Res.*, **62**, 313-320.
- Passow, U (2002) Transparent exopolymer particules (TEP) in aquatic environments. *Prog. Oceanogr.*, **55**, 287-333.
- Patankar, MS, Oehninger, S, Barnett, T, Williams, RL, and Clark, GF (1993) A revised structure of fucoidan may explain some of its biological activities. *J. Biol. Chem.*, **268**, 21770-21776.
- Patier, P, Potin, P, Rochas, c, Kloareg, B, Yvin, JC, and Liénardt, Y (1995) Free and silica-bound oligo-kappa-carrageenans elicit laminarinase activity in *Rubus* cells and protoplasts. *Plant Sci.*, **110**, 27-35.
- Pavão, MS, Albano, RM, Lawson, AM, and Mourão, PA (1989) Structural heterogeneity among unique sulfated L-galactans from different species of ascidians (tunicates). *J. Biol. Chem.*, **264**, 9972-9979.
- Penman, A, and Rees, DA (1973) Carrageenans. IX. Methylation analysis of galactan sulphates from *Furcellaria fastigiata*, *Gigartina canaliculata*, *Gigartina chamissoi*, *Gigartina atropurpurea*, *Ahnfeltia durvillaei*, *Gymnogongrus furcellatus*, *Euclidean isiforme*, *Euclidean uncinatum*, *Aghardhiella tenera*, *Pachymenia hymantophora*, and *Gloiopeltis cervicornis*. Structure of xi-carrageenan. *J. Chem. Soc.*, **19**, 2182-2187.
- Percival, E (1979) The polysaccharides of green, red, and brown seaweeds : their basic structure, biosynthesis and function, *Br. Phycol. J.*, **14**, 103-117.
- Percival, E, and McDowell RH (1967) in *Chemistry and Enzymology of Marine Algal Polysaccharides*, Academic Press, London.
- Percival, E, and Young, M (1971) Characterisation of sucrose lactate and other oligosaccharides found in the Cladophorales. *Carbohydr Res.*, **20**, 217-223.
- Pereira, MS, Melo, FR, and Mourão, PA (2002 B) Is there a correlation between structure and anticoagulant action of sulfated galactans and sulfated fucans? *Glycobiology*, **12**, 573-580.
- Pereira, MS, Mulloy, B, and Mourão, PA (1999) Structure and anticoagulant activity of sulfated fucans. Comparison between the regular, repetitive, and linear fucans from echinoderms with the more heterogeneous and branched polymers from brown algae. *J. Biol. Chem.*, **274**, 7656-7667.

- Pereira, MS, Vilela-Silva, AC, Valente, AP, and Mourão, PA (2002 A) A 2-sulfated, 3-linked alpha-L-galactan is an anticoagulant polysaccharide. *Carbohydr. Res.*, **337**, 2231-2238.
- Perez, S, and Kouwijzer, ML (1999) "Shapes and Interactions of Polysaccharide Chains", in: *Carbohydrates : Structures, Dynamics and Syntheses*, 258-293, Ed. P. Finch, Kluwer.
- Perugino, G, Trincone, A, Rossi, M, and Moracci, M (2004) Oligosaccharide synthesis by glycosynthases. *Trends Biotechnol.*, **22**, 31-37.
- Petitou, M, Casu, B, and Lindahl, U (2003) 1976-1983, a critical period in the history of heparin: the discovery of the antithrombin binding site. *Biochimie*, **85**, 83-89.
- Ponce, NM, Pujol, CA, Damonte, EB, Flores, ML, and Stortz, CA (2003) Fucoidans from the brown seaweed *Adenocystis utricularis*: extraction methods, antiviral activity and structural studies. *Carbohydr. Res.*, **338**, 153-165.
- Potin (1992) Recherche, production, purification, et caractérisation de galactane hydrolases pour la préparation d'oligosaccharides des parois d'algues rouges. Thèse de doctorat de l'Université de Bretagne Occidentale.
- Potin, P, Bouarab, K, Kupper, F, and Kloareg, B (1999) Oligosaccharide recognition signals and defense reactions in marine plant-microbe interactions. *Curr. Opin. Microbiol.*, **2**, 276-283.
- Potin, P, Patier, P, Floc'h, JY, Yvin, JC, 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.
- Potin, P, Richard, C, Barbeyron, T, Henrissat, B, Gey, C, Petillot, Y, Forest, E, Dideberg, O, Rochas, C, and Kloareg, B (1995) Processing and hydrolytic mechanism of the cgkA-encoded kappa-carrageenase of *Alteromonas carrageenovora*. *Eur. J. Biochem.*, **228**, 971-975.
- Potin, P, Sanseau, A, Le Gall, Y, Rochas, C, and Kloareg, B (1991) Purification and characterization of a new kappa-carrageenase from a marine Cytophaga-like bacterium. *Eur. J. Biochem.*, **201**, 241-247.
- Preston, RD (1974) The physical biology of plant cell walls. Chapman & Hall, Londres.
- Preston, RD (1979) Polysaccharide conformation and cell wall function, *Ann. Rev. Plant Physiol.*, **30**, 55-78.

-R-

- Rappé, MS, and Giovannoni, SJ (2003) The uncultured microbial majority. *Annu. Rev. Microbiol.*, **57**, 369-394.
- Rees, DA (1969) Structure, conformation, and mechanism in the formation of polysaccharide gels and networks. *Adv. Carbohydr. Chem. Biochem.*, **24**, 267-332.
- Rees, DA (1975) Polysaccharide conformation. In *Carbohydrates*, Aspinell GO ed., MTP International Rev. Sci., **7**, 251-283.

- Rees, DA, and Welsh, EJ (1977) Secondary and tertiary structure of polysaccharides in solutions and gels. *Angew. Chem. Int. Ed. Engl.*, **16**, 214-224.
- Rees, DA, Morris, ER, Thom, D, and Madden, JK (1982) Shapes and interactions of carbohydrates chains. In : The polysaccharides, vol. 1, 195-290, ed. GO Aspinall, Academic Press, New York.
- Ribeiro, AC, Vieira, RP, Mourão, PA, and Mulloy, B (1994) A sulfated alpha-L-fucan from sea cucumber. *Carbohydr. Res.*, **255**, 225-240.
- Rigden, DJ, Jedrzejak, MJ, and de Mello, LV (2003) Identification and analysis of catalytic TIM barrel domains in seven further glycoside hydrolase families. *FEBS Lett.*, **544**, 103-111.
- Riou, D, Collic-Jouault, S, Pinczon du Sel, D, Bosch, S, Siavoshian, S, Le Bert, V, Tomasoni, C, Siquin, C, Durand, P, and Roussakis, C (1996) Antitumor and antiproliferative effects of a fucan extracted from *Ascophyllum nodosum* against a non-small-cell bronchopulmonary carcinoma line. *Anticancer Res.*, **16**, 1213-1218.
- Rochas, C, and Lahaye, M (1989) Solid state ¹³C-NMR spectroscopy of red seaweeds, agars and carrageenans. *Carbohydr. Polym.*, **10**, 189-204.
- Rörhig, H, Schmidt, J, Walden, R and Schell, J (1997) Cell signaling by oligosaccharides. *Trends Plant Sci.*, **2**, 111-115.
- Ross, AG (1953) Some typical analyses of red seaweeds. *J. Sci. Food Agric.*, **4**, 333-335.
- Rost, B (1996) PHD: predicting one-dimensional protein structure by profile based neural networks. *Meth. Enzymol.*, **266**, 525-539.
- Rychlewski, L, Fischer, D, and Elofsson, A (2003) LiveBench-6: large-scale automated evaluation of protein structure prediction servers. *Proteins*, **53**, 542-547.
- Rye, CS, and Withers, SG (2000) Glycosidase mechanisms. *Curr. Opinion Chem. Biol.*, **4**, 573-580.

-S-

- Sakai, T, Ishizuka, K, and Kato, I (2003 C) Isolation and characterization of a fucoidan-degrading marine bacterium. *Mar. Biotechnol.*, **5**, 409-416.
- Sakai, T, Kawai, T, and Kato, I (2004) Isolation and characterization of a fucoidan-degrading marine bacterial strain and its fucoidanase. *Mar. Biotechnol.*, **6**, 335-346.
- Sakai, T, Kimura, H, and Kato, I (2002) Marine strain of Flavobacteriaceae utilizes brown seaweed fucoidan. *Mar. Biotechnol.*, **4**, 399-405.
- Sakai, T, Kimura, H, Ishizuka, K, Shimanaka, K, Ikai, K, and Kato, I (2003 A) Structure of oligosaccharides derived from *Cladosiphon okamuranus* fucoidan by digestion with marine bacterial enzymes. *Mar. Biotechnol.*, **5**, 536-544.
- Sakai, T, Kimura, H, Kojima, K, Shimanaka, K, Ikai, K, and Kato, I (2003 B) Marine bacterial sulfated fucoglucuronomannan (SFGM) lyase digests brown algal SFGM into trisaccharides. *Mar. Biotechnol.*, **5**, 70-78.

- Santos, JA, Mulloy, B, Mourão, PA (1992) Structural diversity among sulfated alpha-L-galactans from ascidians (tunicates). Studies on the species *Ciona intestinalis* and *Herdmania monus*. *Eur. J. Biochem.*, **204**, 669-677.
- Sen, AK, Das, AK, Banerji, N, Siddhanta, AK, Mody, KH, Ramavat, BK, Chauhan, VD, Vedasiromoni, JR, and Ganguly, DK (1994) A new sulfated polysaccharide with potent blood anti-coagulant activity from the red seaweed *Grateloupia indica*. *Int. J. Biol. Macromol.*, **16**, 279-280.
- Simpson, AG, and Roger, AJ (2004) The real 'kingdoms' of eukaryotes. *Curr. Biol.*, **14**, 693-697.
- Smith, DB, and Cook, WH (1953) Fractionation of carrageenin. *Arch. Biochem. Biophys.*, **45**, 232-233.
- Somerville, C, Bauer, S, Brininstool, G, Facette, M, Hamann, T, Milne, J, Osborne, E, Paredez, A, Persson, S, Raab, T, Vorwerk, S, and Youngs, H (2004) Toward a systems approach to understanding plant cell walls. *Science*, **306**, 2206-2211.
- Stevenson, TT, and Furneaux, RH (1991) Chemical methods for the analysis of sulphated galactans from red algae. *Carbohydr. Res.*, **210**, 277-298.
- Stortz, CA, and Cerezo, SA (1993) The systems of carrageenans from cystocarpic and tetrasporic stages from *Iridaea undulosa*: fractionation with potassium chloride and methylation analysis of the fractions. *Carbohydr. Res.*, **242**, 217-227.
- Stortz, CA, Bacon, BE, Cherniak, R, and Cerezo, AS (1994) High-field NMR spectroscopy of cystocarpic and tetrasporic carrageenans from *Iridaea undulosa*. *Carbohydr. Res.*, **261**, 317-326.
- Sulzenbacher, G, Bignon C, Nishimura, T, Tarling, CA, Withers, SG, Henrissat, B, Bourne, Y (2004) Crystal structure of *Thermotoga maritima* alpha-L-fucosidase. Insights into the catalytic mechanism and the molecular basis for fucosidosis. *J. Biol. Chem.*, **279**, 13119-13128.

-T-

- Takayama M, Koyama N, Sakai T & Kato I (2002) Enzymes capable of degrading a sulfated-fucose-containing polysaccharide and their encoding genes. Patent n°US6489155 B1.
- Thanassi, NM, and Nakada, HI (1967) Enzymic degradation of fucoidan by enzymes from the hepatopancreas of Abalone, *Haliotis* species. *Arch. Biochem. Biophys.*, **118**, 172-177.
- Thornton, DCO (2004) Formation of transparent exopolymeric particles (TEP) from macroalgal detritus. *Mar. Ecol. Prog. Ser.*, **282**, 1-12.
- Thuault, M (1999) Synthèse d'une sonde pour le clonage du gène fucanase. Rapport de stage DUT.
- Tong, HK, Lee, KH, and Wong, HA (1980) The molecular weight and viscosity of the water-soluble polysaccharide(s) from *Eucaema spinosum*. *Carbohydr. Res.* **81**, 1-6.
- Turvey, JR, and Williams, EL (1970) The structure of some xylans from red algae. *Phytochemistry*, **9**, 2383-2388.

-V-

- Van de Velde, F, Knutsen, SH, Usov, AI, Rollema, HS, and Cerezo, AS (2002) ^1H and ^{13}C high resolution NMR spectroscopy of carrageenans: application in research and industry. *Trends Food Sci. Technol.*, **13**, 73–92.
- Van de Velde, F, Pereira, L, and Rollema, HS (2004) The revised NMR chemical shift data of carrageenans. *Carbohydr. Res.*; **339**, 2309-2313.
- Van de Velde, F, Rollema, HS, Grinberg, NV, Burova, TV, Grinberg, VY, and Tromp RH (2002) Coil-helix transition of iota-carrageenan as a function of chain regularity. *Biopolymers*, **65**, 299-312.
- Vassela, A, Davies, GJ, and Böhm, M (2002) Glycosidase mechanisms. *Curr. Opinion Chem. Biol.*, **6**, 619-629.
- Vasseur, E (1948) Chemical studies on the jelly coat of the sea-urchin egg. *Acta Chem. Scand.*, **2**, 900-913.
- Vilela-Silva, ACES, Alves, AP, Valente, AP, Vacquier, VD, and Mourão, PAS (1999) Structure of the sulfated α -L-fucan from the egg jelly coat of the sea urchin *Strongylocentrotus franciscanus* : patterns of preferential 2-O- and 4-O-sulfation determine sperm cell recognition. *Glycobiology*, **9**, 927-933.
- Vilela-Silva, ACES, Castro, MO, Valente, AP, Biermann, CH, and Mourão, PAS (2002) Sulfated fucans from the egg jellies of the closely related sea urchins *Strongylocentrotus droebachiensis* and *Strongylocentrotus pallidus* ensure species-specific fertilization. *J. Biol. Chem.*, **277**, 379-387.
- Voget, S, Leggewie, C, Uesbeck, A, Raasch, C, Jaeger, KE, and Streit, WR (2003) Prospecting for novel biocatalysts in a soil metagenome. *Appl. Env. Microbiol.*, **69**, 6235-6242.
- Vorwerk, S, Somerville, S, and Somerville, C (2004) The role of plant cell wall polysaccharide composition in disease resistance. *Trends Plant Sci.*, **9**, 203-209.

-W-

- Weigl, J, and Yaphe, W (1966) The enzymic hydrolysis of carrageenan by *Pseudomonas carrageenovora*: purification of a kappa-carrageenase. *Can. J. Microbiol.*, **12**, 939-947.
- Wierenga, RK (2001) The TIM-barrel fold/ a versatile framework for efficient enzymes. *FEBS lett.*, **492**, 193-198.
- William, SJ, and Withers, SG (2000) Glycosyl fluorides in enzymatic reactions. *Carbohydr. Res.*, **327**, 27-46.
- Witvrouw, M, and De Clercq, E (1997) Sulfated polysaccharides extracted from sea algae as potential antiviral drugs. *Gen. Pharmacol.*, **29**, 497-511.

- Wong, TY, Preston, LA, and Schiller, NL (2000) ALGINATE LYASE: review of major sources and enzyme characteristics, structure-function analysis, biological roles, and applications. *Annu. Rev. Microbiol.*, **54**, 289-340.

-X-

- Xiao, L, Yang, C, Patterson, PS, Udhayakumar, V, and Lal, AA (1996) Sulfated polyanions inhibit invasion of erythrocytes by plasmodial merozoites and cytoadherence of endothelial cells to parasitized erythrocytes. *Infect. Immun.*, **64**, 1373-1378.

-Y-

- Yaoi, K, Kondo, H, Noro, N, Suzuki, M, Tsuda, S, and Mitsuishi, Y (2004) Tandem repeat of a seven-bladed beta-propeller domain in oligoxyloglucan reducing-end-specific cellobiohydrolase. *Structure (Camb.)*, **12**, 1209-1217.
- Yaphe, W, and Baxter, B (1955) The enzymic hydrolysis of carrageenin. *Appl. Microbiol.*, **3**, 380-383.
- Yaphe, W, and Morgan, K (1959) Enzymic hydrolysis of fucoidan by *Pseudomonas atlantica* and *Pseudomonas Carrageenovora*. *Nature (Lond.)*, **183**, 761-762.

-Z-

- Zablackis, E, Huang, J, Muller, B, Darvill, AG, and Albersheim, P (1995) Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiol.*, **107**, 1129-1138
- Zhong, Z, Toukdarian, A, Helinski, D, Knauf, V, Sykes, S, Wilkinson, JE, O'Bryne, C, Shea, T, DeLoughery, C, and Caspi, R (2001) Sequence analysis of a 101-kilobase plasmid required for agar degradation by a *Microscilla* isolate. *Appl. Environ. Microbiol.*, **67**, 5771-5779.

ANNEXES

ANNEXE 1

Contig: *P. carrageenovora* (1-5665 nt)

Pourcentage: 32,3% A; 17,3% C; 20,8% G; 29,6% T

Traduction:- code universel

- 3 phases ouvertes de lecture (ORF), 2 colinéaires sur 2 cadres, 1 sur le brin opposé (en gras)

```
1      TTCCTAAAGGTTTCACGCTTGGCTCGATTGAAGACGTTTTAGCAATTCAAAAATTGATGA
1      P K G F T L G S I E D V L A I Q K L M

61     TAAAGCATAATGGTGAGACCGTTTTACGGTTGGAAATGTATGCTTCCGTCAAAACAATGGCA
20     I K H N G E T V Y G W K C M L P S N N G

121    TAATTACTCTAGCTCCGCTGCTTCATAAGCCAATTTATGCCACTAAGGCTTGCTTGTTAA
40     I I T L A P L L H K P I Y A T K A C L L

181    AGACAAAAATAACAAAGCTCTGATAGAGCCTGAGATTGCTTTTATATTAACAAAGACT
60     K T K N N K A L I E P E I A F I L N K D

241    TAAGTGC GTTGAAGGTATATACAAATGAACAAATAGATAATGCTGTAGGCAGTGCACATC
80     L S A L K V Y T N E Q I D N A V G S A H

301    TAGCATTAGAGCTTATTGAAAACAGGTTTAGCAAGCAATACGAAGCTACACAGTTTGAAA
100    L A L E L I E N R F S K Q Y E A T Q F E

361    AACTTGCCGACGGTTTATCTAATCAAGGTGTGTATATCGGCCAGAAAATAAGCCAGAAAA
120    K L A D G L S N Q G V Y I G P E I S Q K

421    AAGCATATGAGAGTAATAATATAACGCTTATAGTGGAGCAGGAAGGCGACAATAGCTTTA
140    K A Y E S N N I T L I V E Q E G D N S F

481    ATGTTGCAGGGCAGCATCCAGGCAAGTTACCTCAACTTCCACTTTATTGGGCCATCAATT
160    N V A G Q H P G K L P Q L P L Y W A I N

541    TTTTATCTGCAAATGGCTTTAAGTTAGAAAAAACCAAGTGTTTATAACCGGCTCATATA
180    F L S A N G F K L E K N Q V F I T G S Y

601    ATGGCCTTATGGAGATGGAGATAAATAAACAGATACAAATAACGTATGATGGGCTTGTA
200    N G L M E M E I N K Q I Q I T Y D G L G

661    AAATACACGTTAAATTTGAAACTTAAATTCTATCCCATAATAGTTATTTCGTTAAGTTTAA
220    K I H V K F E T *

721    TATGTTTGTAGTCTAAAGAGAGCTAGTTTAAATTAAGCTGGCTCAGTTTACTCTGTTTAA
781    GCTTTAGAGCCTTCAATAGCATGAGCTGTCAGTTTACCCGCCAAATCTCCAGCATTATAA
841    GCGTCCAAATTCTCCTCTGGATAAAAAATATCCTAAATGTAACATAATATAATTTATTTA
901    CTGATAACAATTGTAAAGATTAGCAAATTACAAATTATGTAAAGTAAAAGTTACAAGTGG
961    TATCGTTCTCCCTATTGTGTATATTTATTTTAGAGTGATTATCCTAATTCTTAATTATCT
1021   TCCTACATGCCAGTAAACAAGGCAGGTTCAAAGATGAATATAATAGTATATGTATTATAA
1081   ATAGAGGTTTCTTATGATTTTTAGGTTTAAATGTAAGTAAACCAATGAAAAAGGATTG
1141   TTTATGCTGTTTTTTTCTGAAAATAAATTTGCAGCTAATTTTTGACGCTTATTTTTTA
1201   AAAAGCATAAACTCAATATTTAGGAAATAGGTACAATCTAATTCCTCAACACGTTATAAA
```

1261 TAACTTAATGT**AAGGAAT**GTTTATGAAAATAAAAAATTCTATCTGCAATGGTAGCTAGCTCG
 421 M K I K I L S A M V A S S
 1321 TTATTAATTGGCTGCGTTATCCCTACCGTTAAAGCTTCTCAATCGGCTATTAAGTATT
 441 L L I G C V I P T V K A S Q S A I K S I
 1381 GAAACAAACCGAACAATTACTAAAGTTAGAACAGGAATGTTGAGTGGAGGCTCATCAATC
 461 E T N R T I T K V R T G M L S G G S S I
 1441 ATAACTACAAGCTATGAAGGGACTGTAGCTGCATATAAGTTTAAATGGAGAAAAACTGTGG
 481 I T T S Y E G T V A A Y K F N G E K L W
 1501 GAAAATGAACTCTCGGGTTTTATGAATCATGATATTTGGGTTCAAGATATTAATGGTGAT
 501 E N E L S G F M N H D I W V Q D I N G D
 1561 GGACTTGTAGAGATATTTGCTGCGAATGCCGATGGCAATGTTTACTGTATTAATAGTGAT
 521 G L V E I F A A N A D G N V Y C I N S D
 1621 GGTTCTTTAAAGTGGACGTTTGGTCTAAATGAAGTCCCTATGAACTCTGTAAGTGTAAATC
 541 G S L K W T F G L N E V P M N S V T V I
 1681 TCTGATGCAGATAAAAAAGTATGTTGTGGCAGGTGGTTACGATAAAAACTTGTATTACATA
 561 S D A D K K Y V V A G G Y D K N L Y Y I
 1741 TCGACTAATGGTGAACCTTTTAAAAACAATTGAATCAGGTACTTACTCAGAAGAAGGGGTG
 581 S T N G E L L K T I E S G T Y S E E G V
 1801 TTTGGGGATGGCGTCAAGCCTGAAGCTCGAACTCATACTGTAAATTTTCGTCCGTCCAGTA
 601 F G D G V K P E A R T H T V N F V R P V
 1861 AAATCTAGTGACGGCACTGAAAACTAGTAGTTCTAGGGACTAATAATAGCCTTCAGAGT
 621 K S S D G T E K L V V L G T N N S L Q S
 1921 TCTGGTTCGGTTCTATATTTTTGAACCATTTGCTGATTTACCCAGTGAAAAGAGTCGAATA
 641 S G R F Y I F E P F A D L P S E K S R I
 1981 TCTATTA AAAAGGGGATAGGAGACCTTCGCACTGTAGATTTTGATAATGATGGTAATGAC
 661 S I K K G I G D L R T V D F D N D G N D
 2041 GAACTTACGCTAGGTAATAGTGCACAGATAGGTGATGCTGCAATTTCTGTTATGAATTTG
 681 E L T L G N S A Q I G D A A I S V M N L
 2101 GATGATCTGTACAGAAGAAAAGCCAAATTAATGACATTGCACGTAGAATAGATCGCTTT
 701 D D L S Q K K S Q I N D I A R R I D R F
 2161 GGGTATAGGGTTGCTCAAACAGAAGTTGTTATGAATGAAGGCACTCCTACATACCTGACA
 721 G Y R V A Q T E V V M N E G T P T Y L T
 2221 CTTTTGGCTCTAGGATACTTCTAACACCAGAGTCATTTGACGTGAACGATTCTGAAATA
 741 L F G S R I L L T P E S F D V N D S E I
 2281 CTTGCAAATAAATACTCTTATTATGACATGTGGAAAGATAAAAAGTTCAAATAAGTTGGTT
 761 L A N K Y S Y Y D M W K D K S S N K L V
 2341 TTAGCTAGTGCACAAAGTGGTGGCAGCCAAGTGCATATTATTGATACTTCTAACCCAAGT
 781 L A S A Q S G G S Q V H I I D T S N P S
 2401 TGGAAGTCTGCTTACGAGGAATTAGAGCCTCAAGGTAAATTAGCTGCAATACAAGAAAAT
 801 W K S A Y E E L E P Q G K L A A I Q E N

2461 ACTAGAGCAATTGAAAGGCAACTATCTAACTTTCAAAGCCTACGCGAGAGCGTGCTCCT
 821 T R A I E R Q L S N F Q K P T R E R A P

2521 TTGCCCGTGTATTTTATTTTTCAGAAAGTAGAAATGAAATTCACAACAATCGAGCGAAGT
 841 L P V Y F I S E S R N E I P T T I E R S

2581 GAGTTTTTGTATGACTCACCAGTATTTTTTAAATTACTCAACTCTACCAAATGTAGAGAAC
 861 E F L Y D S P V F L N Y S T L P N V E N

2641 TGGGATAGAAGTGAAGTTCCTTGCTGATAATCCTAAATACAGAGATAAGAGAGACCGTCGT
 881 W D R S E V L A D N P K Y R D K R D R R

2701 AAAAACTATACACTATCATCTGAAGAAATGTTTAATAAGCTTTTCGGCTGGATATGATAAT
 901 K N Y T L S S E E M F N K L S A G Y D N

2761 AGCGATGGGATATCACAATGGGCAGGGCATGGAAATGATCCGTATATGATTAGTTTAGCT
 921 S D G I S Q W A G H G N D P Y M I S L A

2821 ACTATGAAAAGAATTATTAGTAGTGGCGATGGGAAAAAACGGTCAATATATATCCAGAG
 941 T M K R I I S S G D G K K T V N I Y P E

2881 ATAGAAGGACATGGAGATGCATTTAATAAAGTACTAAGTGATCATTTCTACCCTTTGGCT
 961 I E G H G D A F N K V L S D H F Y P L A

2941 GAGTTTAGTAGTGAGAATAATGCTAATTTGTTTATGCGCAATAAACATACATTTTGGCAA
 981 E F S S E N N A N L F M R N K H T F W Q

3001 TCAACTATATATGCTCCAGAGTGGTCGGAGTTACGTTCTGGACGTTTAGCTGATGCGTTT
 1001 S T I Y A P E W S E L R S G R L A D A F

3061 GTACCTGCAATGGAAGAAACAACAGATAAATCAATGGAAATGAGTGTGGCCGGGAGAATG
 1021 V P A M E E T T D K S M E M S V A G R M

3121 GGGTTATGGGCTGCAGGCTCTGTAGATAATTGGGGTGAAAGATACGCAAGAGACAATCCT
 1041 G L W A A G S V D N W G E R Y A R D N P

3181 AGTTTTGACCGTTTACGCCAACACTCTCATCAAATGGTCCCCAACCATGCACTAAGGCAA
 1061 S F D R L R Q H S H Q M V P N H A L R Q

3241 ATAATATATAAGATAGCGTCTGGTGCCAGATATATAAATAACTTCGGCTTCAATCAGGAG
 1081 I I Y K I A S G A R Y I N N F G F N Q E

3301 TATATGAGCCTTGCTTGGGAGTTAATTGGTAAAGGCGCTCTTTATGTACCAAAGCGTGAA
 1101 Y M S L A W E L I G K G A L Y V P K R E

3361 GAATTATTAAGCTTGTACCAGTTCACATAAGTATGAAAGAACCAGATCCAATCTACAGA
 1121 E L L S L S P V H I S M K E P D P I Y R

3421 GAAACATCAAATAATGTGAAGTGGACTACGTTCTATGATGAAGAAAAAGACTCTATTCCA
 1141 E T S N N V K W T T F Y D E E K D S I P

3481 TATGTGTTTAGTCGCTTAAATGGTACTTGGCCAGGTGCAAAAACCTTGCCTTGGGATTAC
 1161 Y V F S R L N G T W P G A K T L P W D Y

3541 TCTAATTATGCTGCCGATACTAAAGAAAGGAGATTAGATTTTATACCAAATTTCTCTAAA
 1181 S N Y A A D T K E R R L D F I P K F P K

3601 GGATTAGTTTTAATTACGCCTGTTTCAGCAAGGTAAATTTAAAGATGAAGGCACCGTTAGA
 1201 G L V L I T P V Q Q G K F K D E G T V R

3661 GGCACATTAGCTGATAACATGCACCCTATCTATAAAGATATAATGAAAGAGTATATTACT
1221 G T L A D N M H P I Y K D I M K E Y I T

3721 GACGGTAAAACTACTATAACCCTAATGGTGAGCAAGTAATGGCTGCTGATAGTGTTAGA
1241 D G K N Y Y N P N G E Q V M A A D S V R

3781 TACAGACAAATTAATAAAGATTGAAGAAAAATCGAATCTTTTACCAATGACTGTATCA
1261 Y R Q I K N K I E E K S N L L P M T V S

3841 GGAGAAGCCGCTTGGGTTGTTGCTCAATCTGCTGAAAAGCATTTCGCGACTGACGCTGGTT
1281 G E A A W V V A Q S A E K H L R L T L V

3901 GATAGTGGTTATTTAAATCCGAGTAACAAAGTTGCTAAGGTTAAATTTAACTCTGTAACA
1301 D S G Y L N P S N K V A K V K F N S V T

3961 CCTGTAGCGATAGTAGATGTATTGTCAGGTGAGACATTTTCGCCAGATTCAAATGGGGTT
1321 P V A I V D V L S G E T F S P D S N G V

4021 GTAGAAATACCGGTCTTAGCTGGTGCTTTTAGGTTTCATTGATGTGAAAATTACTGAAGAC
1341 V E I P V L A G A F R F I D V K I T E D

4081 CTTAGAAACATGCAAAGTTCAACATTGTAATTACGGCGTTTAATTTATAGTGCGGGATTG
1361 L R N M Q S S T L *

4141 AAGTTTATTAGTCTGCACTATTTAGTCTAAAGTAAGTTGCTCACAAAAAGCAGAAATTG
4201 CAAGGCGGCTACGAGCGAGAAGCAGACGGTGAGTAAACTCGTTTTTTGGACAAAAGCTCGT
4261 CAGAGTAAATAACCTAGAACCCCGCTTTTCAGATAACTTCTCGTATCTCAAACGTCAAAG
4321 ACCACCGTTGTCAGCGGTGGTCGGTTTTGCGCGGAAGTGGACAGTTTCGAGCATGGTCGAT
4381 TTTGGGAGTGAATTATTTTTTTGCTGCAATACTGTATATTAATCCACGAAAACACCATCAGC
4441 **AAACGATTTGTGTCACTTTCTTCTGAGCAGTTCCCAGAAAAGCAATAACAGCGGGGCTAA**
4501 **CCTGCTCTCTGCTTGGACAGATTAGATTTAACTCAACAGGTGGTGATTCAAAGCCTGGTA**
4561 **AAAGCCGTACAAGTTGCCCACTCCGAAGTTCTGAATGGACATCTATCTGGGAGCGATACG**
4621 **CTATCCCTTTACCCTAACAGCCCATTTACGAACAATGTCAGTATCATTGCTGACTCGAT**
4681 **TGCTATCCACTTTGACTTTGTATGAACCAGAATGACCAGCGTATTCCCAACTGTTCAATA**
4741 **AACGACCAGCTCGACGGTGCAATAAGCAATTATGCTCCTTTAAGCTGTTAGGATGCGTTA**
4801 **GCTCACCATGTTTAGATATATAGGATGGAGAGGCGCAGGTGATCCTGCTCATGGTGCGA**
4861 **TATGAAATGAAACCATCGATGAGTCTTCTAGCTTCCCATAAACGAAGTGCCATATCAACTT**
4921 **GGTCGAGAAAAAATCTGAAATTGAATCGCCAACCGTAAGGTCAATCGACAATAAAGGGT**
4981 **GTTTCATCCAGTAACTCTTCTACCCAGTGTAAAAGAGTATTGCGCCCTAAGTCAGATGAAA**
5041 **CGGACAACCTTAACTTACCACTCACCTTACCTTGTGTCTGATGTGCAGCTATACGACCTT**
5101 **GCTCCAAACTGGCTAATGCCTGGCGACAATGAAACAGAAATTGTTGCGCTTGAGGTGTGA**
5161 **TCATTAATTGCCTTGTGGTTCTAATAAGTAGCTGTATATCAAACCTTTTTTCCAGTCTTT**
5221 **TAAGCCAGAGCTGACTGCCGAGGAGTGACACTCAGTTGTTTTGCGGTTCCGTTATGC**
5281 **TGCTGTTTTCCACTATTCGTATGAATAGATTGAGATCTGATGTATTCAATTATTAATAAT**
5341 ATTTAAAAGTGTTTGAACCTATTGCTATCTTTTTAGTTTAGTAAAAATCGGGGACAATGTC
5401 TCTACCGTTATTAATTAACACCGTACATAAGTATTGGTGCCCCGAGAGGTAGATATGACC
5461 CAGCCATTACACACCGCGAAGTACTTACTGATCCTCCTTTAGAGCTCCCTTGTGGCGCAGTC
5521 ATCAAAAACCGTCTAATAAAGTCAGCAATGTCTGACTCACTCGCTGACGGTGAAGGGAAC
5581 CCTACTGAAGCTCAGTCTCGACTCTATGAACGATGGGCAGAGGGTGTATTGGGTTATCC
5641 GTCATTGGTGAAGTGCAGGTAGATC

ORF 1 : incomplet

Universal code

Total amino acid number: 227, MW=25254

Max ORF: 3-683

1 TTCCTAAAGGTTTCACGCTTGGCTCGATTGAAGACGTTTTAGCAATTCAAAAATTGATGA
1 P K G F T L G S I E D V L A I Q K L M

61 TAAAGCATAATGGTGAGACCGTTTTACGGTTGGAAATGTATGCTTCCGTCAAAACAATGGCA
20 I K H N G E T V Y G W K C M L P S N N G

121 TAATTACTCTAGCTCCGCTGCTTCATAAGCCAATTTATGCCACTAAGGCTTGCTTGTTAA
40 I I T L A P L L H K P I Y A T K A C L L

181 AGACAAAAATAACAAAGCTCTGATAGAGCCTGAGATTGCTTTTTATATTAACAAAGACT
60 K T K N N K A L I E P E I A F I L N K D

241 TAAGTGC GTTGAAGGTATATACAAATGAACAAATAGATAATGCTGTAGGCAGTCACATC
80 L S A L K V Y T N E Q I D N A V G S A H

301 TAGCATTAGAGCTTATTGAAAACAGGTTTAGCAAGCAATACGAAGCTACACAGTTTGAAA
100 L A L E L I E N R F S K Q Y E A T Q F E

361 AACTTGCCGACGGTTTTATCTAATCAAGGTGTGTATATCGGCCCCAGAAAATAAGCCAGAAAA
120 K L A D G L S N Q G V Y I G P E I S Q K

421 AAGCATATGAGAGTAATAATATAACGCTTATAGTGGAGCAGGAAGGCGACAATAGCTTTA
140 K A Y E S N N I T L I V E Q E G D N S F

481 ATGTTGCAGGGCAGCATCCAGGCAAGTTACCTCAACTTCCACTTTATTGGGCCATCAATT
160 N V A G Q H P G K L P Q L P L Y W A I N

541 TTTTATCTGCAAATGGCTTTAAGTTAGAAAAAACCAAGTGTTTATAACCGGCTCATATA
180 F L S A N G F K L E K N Q V F I T G S Y

601 ATGGCCTTATGGAGATGGAGATAAATAAACAGATACAAATAACGTATGATGGGCTTG GTA
200 N G L M E M E I N K Q I Q I T Y D G L G

661 AAATACACGTAAATTTGAAACTTAAATTCTATCCCATAATAGTTATTCGTTAAGTTTAA
220 K I H V K F E T

ORF 2 : CglA complet

Universal code

Total amino acid number: 942, MW=105501

Max ORF: 1282-4107

1261 TAACTTAATGT**AAGGAAT**GTTATGAAAATAAAAAATTCTATCTGCAATGGTAGCTAGCTCG
1 M K I K I L S A M V A S S

1321 TTATTAATTGGCTGCGTTATCCCTACCGTTAAAGCTTCTCAATCGGCTATTA AAAAGTATT
14 L L I G C V I P T V K A S Q S A I K S I

1381 GAAACAAACCGAACAATTACTAAAGTTAGAACAGGAATGTTGAGTGGAGGCTCATCAATC
34 E T N R T I T K V R T G M L S G G S S I

1441 ATAACTACAAGCTATGAAGGGACTGTAGCTGCATATAAGTTTAAATGGAGAAAAACTGTGG
54 I T T S Y E G T V A A Y K F N G E K L W

1501 GAAAATGAACTCTCGGGTTTTATGAATCATGATATTTGGGTTCAAGATATTAATGGTGAT
74 E N E L S G F M N H D I W V Q D I N G D

1561 GGACTTGTAGAGATATTTGCTGCGAATGCCGATGGCAATGTTTACTGTATTAATAGTGAT
94 G L V E I F A A N A D G N V Y C I N S D

1621 GGTTCTTTAAAGTGGACGTTTGGTCTAAATGAAGTCCCTATGAACTCTGTAAGTCTGTAATC
114 G S L K W T F G L N E V P M N S V T V I

1681 TCTGATGCAGATAAAAAGTATGTTGTGGCAGGTGGTTACGATAAAAACTTGTATTACATA
134 S D A D K K Y V V A G G Y D K N L Y Y I

1741 TCGACTAATGGTGAACCTTTTAAAAACAATTGAATCAGGTACTTACTCAGAAGAAGGGGTG
154 S T N G E L L K T I E S G T Y S E E G V

1801 TTTGGGGATGGCGTCAAGCCTGAAGCTCGAACTCATACTGTAAATTTTCGTCCGTCCAGTA
174 F G D G V K P E A R T H T V N F V R P V

1861 AAATCTAGTGACGGCACTGAAAACTAGTAGTTCTAGGGACTAATAATAGCCTTCAGAGT
194 K S S D G T E K L V V L G T N N S L Q S

1921 TCTGGTCGGTTCTATATTTTTGAACCATTTGCTGATTTACCCAGTGAAAAGAGTCAATA
214 S G R F Y I F E P F A D L P S E K S R I

1981 TCTATTA AAAAGGGGATAGGAGACCTTCGCACTGTAGATTTTGATAATGATGGTAATGAC
234 S I K K G I G D L R T V D F D N D G N D

2041 GAACTTACGCTAGGTAATAGTGCACAGATAGGTGATGCTGCAATTTCTGTTATGAATTTG
254 E L T L G N S A Q I G D A A I S V M N L

2101 GATGATCTGTCACAGAAGAAAAGCCAAATTAATGACATTGCACGTAGAATAGATCGCTTT
274 D D L S Q K K S Q I N D I A R R I D R F

2161 GGGTATAGGGTTGCTCAAACAGAAGTTGTTATGAATGAAGGCACTCCTACATACCTGACA
294 G Y R V A Q T E V V M N E G T P T Y L T

2221 CTTTTTGGCTCTAGGATACTTCTAACACCAGAGTCATTTGACGTGAACGATTCTGAAATA
314 L F G S R I L L T P E S F D V N D S E I

2281 CTTGCAAATAAATACTCTTATTATGACATGTGGAAAGATAAAAAGTTCAAATAAGTTGGTT
334 L A N K Y S Y Y D M W K D K S S N K L V

2341 TTAGCTAGTGCACAAAGTGGTGGCAGCCAAGTGCATATTATTGATACTTCTAACCCAAGT
354 L A S A Q S G G S Q V H I I D T S N P S

2401 TGGAAGTCTGCTTACGAGGAATTAGAGCCTCAAGGTAAATTAGCTGCAATACAAGAAAAT
374 W K S A Y E E L E P Q G K L A A I Q E N

2461 ACTAGAGCAATTGAAAGGCAACTATCTAACTTTCAAAGCCTACGCGAGAGCGTGCTCCT
394 T R A I E R Q L S N F Q K P T R E R A P

2521 TTGCCCGTGTATTTTATTTTTCAGAAAGTAGAAATGAAATTTCAAACAACAATCGAGCGAAGT
414 L P V Y F I S E S R N E I P T T I E R S

2581 GAGTTTTTGTATGACTCACCAGTATTTTTAAATTACTCAACTCTACCAAATGTAGAGAAC
434 E F L Y D S P V F L N Y S T L P N V E N

2641 TGGGATAGAAGTGAAGTTCTTGCTGATAATCCTAAATACAGAGATAAGAGAGACCGTCGT
454 W D R S E V L A D N P K Y R D K R D R R

2701 AAAA ACTATACTATCATCTGAAGAAATGTTTAATAAGCTTTTCGGCTGGATATGATAAT
474 K N Y T L S S E E M F N K L S A G Y D N

2761 AGCGATGGGATATCACAATGGGCAGGGCATGGAAATGATCCGTATATGATTAGTTTAGCT
494 S D G I S Q W A G H G N D P Y M I S L A

2821 ACTATGAAAAGAATTATTAGTAGTGGCGATGGGAAAAAACGGTCAATATATATATCCAGAG
 514 T M K R I I S S G D G K K T V N I Y P E

 2881 ATAGAAGGACATGGAGATGCATTTAATAAAGTACTAAGTGATCATTTCTACCCTTTGGCT
 534 I E G H G D A F N K V L S D H F Y P L A

 2941 GAGTTTAGTAGTGAGAATAATGCTAATTTGTTTATGCGCAATAAACATACATTTTGGCAA
 554 E F S S E N N A N L F M R N K H T F W Q

 3001 TCAACTATATATGCTCCAGAGTGGTTCGGAGTTACGTTCTGGACGTTTAGCTGATGCGTTT
 574 S T I Y A P E W S E L R S G R L A D A F

 3061 GTACCTGCAATGGAAGAAACAACAGATAAATCAATGGAAATGAGTGTGGCCGGGAGAATG
 594 V P A M E E T T D K S M E M S V A G R M

 3121 GGGTTATGGGCTGCAGGCTCTGTAGATAATTGGGGTGAAAGATACGCAAGAGACAATCCT
 614 G L W A A G S V D N W G E R Y A R D N P

 3181 AGTTTTGACCGTTTACGCCAACACTCTCATCAAATGGTCCCAACCATGCACTAAGGCAA
 634 S F D R L R Q H S H Q M V P N H A L R Q

 3241 ATAATATATAAGATAGCGTCTGGTGCCAGATATATAAATAACTTCGGCTTCAATCAGGAG
 654 I I Y K I A S G A R Y I N N F G F N Q E

 3301 TATATGAGCCTTGCTTGGGAGTTAATTGGTAAAGGCGCTCTTTATGTACCAAAGCGTGAA
 674 Y M S L A W E L I G K G A L Y V P K R E

 3361 GAATTATTAAGCTTGTCAACAGTTCACATAAGTATGAAAGAACCAGATCCAATCTACAGA
 694 E L L S L S P V H I S M K E P D P I Y R

 3421 GAAACATCAAATAATGTGAAGTGGACTACGTTCTATGATGAAGAAAAAGACTCTATTCCA
 714 E T S N N V K W T T F Y D E E K D S I P

 3481 TATGTGTTTAGTCGCTTAAATGGTACTTGGCCAGGTGCAAAAACCTTTGCCTTGGGATTAC
 734 Y V F S R L N G T W P G A K T L P W D Y

 3541 TCTAATTATGCTGCCGATACTAAAGAAAGGAGATTAGATTTTATACCAAATTTCTCTAAA
 754 S N Y A A D T K E R R L D F I P K F P K

 3601 GGATTAGTTTTAATTACGCCTGTTTCAGCAAGGTAAATTTAAAGATGAAGGCACCGTTAGA
 774 G L V L I T P V Q Q G K F K D E G T V R

 3661 GGCACATTAGCTGATAACATGCACCCTATCTATAAAGATATAATGAAAGAGTATATTACT
 794 G T L A D N M H P I Y K D I M K E Y I T

 3721 GACGGTAAAACTACTATAACCCTAATGGTGAGCAAGTAATGGCTGCTGATAGTGTTAGA
 814 D G K N Y Y N P N G E Q V M A A D S V R

 3781 TACAGACAAATTAATAAAGATTGAAGAAAAATCGAATCTTTTACCAATGACTGTATCA
 834 Y R Q I K N K I E E K S N L L P M T V S

 3841 GGAGAAGCCGCTTGGGTTGTTGCTCAATCTGCTGAAAAGCATTGCGACTGACGCTGGTT
 854 G E A A W V V A Q S A E K H L R L T L V

 3901 GATAGTGGTTATTTAAATCCGAGTAACAAAGTTGCTAAGGTTAAATTTAACTCTGTAACA
 874 D S G Y L N P S N K V A K V K F N S V T

 3961 CCTGTAGCGATAGTAGATGTATTGTCAGGTGAGACATTTTCGCCAGATTCAAATGGGGTT
 894 P V A I V D V L S G E T F S P D S N G V

4021 GTAGAAATACCGGTCTTAGCTGGTGCCTTTTAGGTTTCATTGATGTGAAAATTACTGAAGAC
914 V E I P V L A G A F R F I D V K I T E D

4081 CTTAGAAACATGCAAAGTTCAACATTGTAATTACGGCGTTTAATTTATAGTGCGGGATTG
934 L R N M Q S S T L

ORF 3 : complet

Universal code

Total amino acid number: 297, MW=33148

Max ORF : sequence reverse-complémentaire 4436-5327

1 ATGAATACATCAGATCTCAATCTATTCATACGAATAGTGGAACAGGCAGCATAACGGAA
1 M N T S D L N L F I R I V E T G S I T E

61 ACCGCAAAACAACACTGAGTGTCACTCCTGCGGCAGTCAGCTCTGCGCTTAAAAGACTGGAA
21 T A K Q L S V T P A A V S S A L K R L E

121 AAAAAGTTGGATATACAGCTACTTATTAGAACCACAAGGCAATTAAGGATCACACCTCAA
41 K K L D I Q L L I R T T R Q L R I T P Q

181 GGCGAACAAATTTCTGTTTCATTGTGCGCCAGGCATTAGCCAGTTTGGAGCAAGGTCGTATA
61 G E Q F L F H C R Q A L A S L E Q G R I

241 GCTGCACATCAGACACAAGGTAAGGTGAGTGGTAAGTTAAGGTTGTCCGTTTCATCTGAC
81 A A H Q T Q G K V S G K L R L S V S S D

301 TTAGGGCGCAATACTCTTTTACACTGGGTAGAAGAGTTACTGGATGAACACCCTTTATTG
101 L G R N T L L H W V E E L L D E H P L L

361 TCGATTGACCTTACGGTTGGCGATTCAATTTTCAGATTTTTTTCTCGACCAAGTTGATATG
121 S I D L T V G D S I S D F F L D Q V D M

421 GCACTTCGTTATGGGAAGCTAGAAGACTCATCGATGGTTTCATTTTCATATCGCCACCATG
141 A L R Y G K L E D S S M V S F H I A T M

481 AGCAGGATCACCTGCGCCTCTCCATCCTATATATCTAAACATGGTGAGCTAACGCATCCT
161 S R I T C A S P S Y I S K H G E L T H P

541 AACAGCTTAAAGGAGCATAAATTGCTTATTGCACCGTCGAGCTGGTCGTTTATTGAACAGT
181 N S L K E H N C L L H R R A G R L L N S

601 TGGGAATACGCTGGTCATTCTGGTTTCATACAAAGTCAAAGTGGATAGCAATCGAGTCAGC
201 W E Y A G H S G S Y K V K V D S N R V S

661 AATGATACTGACATTGTTTCGTAATGGGCTGTTAGTGGTAAAGGGATAGCGTATCGCTCC
221 N D T D I V R K W A V S G K G I A Y R S

721 CAGATAGATGTCCATTCAGAACTTCGGAGTGGGCAACTTGTACGGCTTTTACCAGGCTTT
241 Q I D V H S E L R S G Q L V R L L P G F

781 GAATCACACCTGTTGAGTTAAATCTAATCTGTCCAAGCAGAGAGCAGGTTAGCCCCGCT
261 E S P P V E L N L I C P S R E Q V S P A

841 GTTATTGCTTTTCGGGAAGTCTCAGGAAGAAAGTGACACAAATCGTTTCGTGA
281 V I A F R E L L R K K V T Q I V S *

ANNEXE 2

Contig: SW5 (1-7723 nt)

Pourcentage: 35,2% A; 15,5% C; 20,5% G; 28,8% T

Traduction:- code universel

- 5 phases ouvertes de lecture (ORF)
colinéaires sur les 3 cadres

LES SEQUENCES EN GRAS INDIQUE UNE SEQUENCE INCERTAINE CAR
OBTENUE EN EXPLOITANT L'ADN GENOMIQUE DE SW5 COMME MATRICE.

```
1      TGTGGATCGTTATTTAGTAACAGATACACAAAGGATTTTCATGGCCAGTAAAAGGAAAACAAAGTTG
1      V D R Y L V T D T Q R I S W P V K G K Q S C

67     TGTTATGAATGGAGATTGGCGATTGGTGAATGAACTGTATAACACAGCTACCGACCCAGG
23     V M N G D W R L V N G N E L Y N T A T D P G

133    ACAAAAAATAAACCTAGCTGATAAACATCCTGATAAAGTGGCCAATATGAATGATTTTTATAACAC
45     Q K I N L A D K H P D K V A N M N D F Y N T

199    ATGGTGGGATAGTGTATAAAAGAACTAAATACTCTACCATCGATTTAGGAGTTGGTGAAGAAGA
67     W W D S V I K E T K Y S T I D L G V G E E D

265    TGTTATTACCTGTCATGATGCGAGAACAATTGATTATTTCCCGCCATGGAATCAAGAAATGATAAG
89     V I T C H D A R T I D Y F P P W N Q E M I R

331    AAAAGGAGAGCCTATGAAACCAGCACCATTCTTCGTGAATTTTCGTAAAACCGGGTAACTATACATT
111    K G E P M K P A P F F V N F V K P G N Y T F

397    TAAGTTAAGACGTTGGCCTGCTGAAAGTGGCATGTCGTTAGGAGCTGAAATTTTAGATGAAATACC
133    K L R R W P A E S G M S L G A E I L D E I P

463    GGCAACTCCCACAACCGATGGTAGAAGTATTGGAAAAGCCATGAAGTTTAAAAAAGCATTTTTTAAA
155    A T P T T D G R S I G K A M K F K K A F L K

529    AATAGGAGACCAAGAAGTTTCTGTTGAGGCCAATAATGCGAGTCAATCTGCAGATATAGATTTTCGT
177    I G D Q E V S V E A N N A S Q S A D I D F V

595    CGTTGAGCAAGGAAAACCGAGTTATTGGCTTGGTTTGGAGATGGAAAACGGCACATTAACAAATGC
199    V E Q G K T E L L A W F E M E N G T L T N A

661    TTTTTATGTGAATGTTGAAAAAAGCAAAGAATAGCATAGAGTTAAAGTAAATGGTAACATTTAAAT
221    F Y V N V E K S K E *

727    TATGAAAAGGTTTTTAATATATGTTTTATCCATTGGTTTTGCAATAGGGCTCAAAGCACAGAGTTC
1     M K R F L I Y V L S I G F A I G L K A Q S S

793    TAGCCAAAAAGTGGTATCTAAAAATGAATTTTCAAAACTGGATAGCAATAAAAAACGGCACGCTTGA
23     S Q K V V S K N E F S K L D S N K N G T L E

859    AAGAGATGAATTATCACGGATTTGGACTCAAATAAAACAGCATGATACCGATAATAATGGAAAAGAT
45     R D E L S R I W T Q I K Q H D T D N N G K I

925    TACATTAGAAGAATTTTCAAAATTTGAAATTCACATTTAGAACTAAAGGCGATATAATTTTAAA
67     T L E E F S K F E I P H L E T K G D I I L N
```

991 TATTAATACAAAACAACAAAAGAAGAAGATTTGTATTTAGATGTTTATTACCCCAATAAAAACC
 89 I K Y K T T K E E D L Y L D V Y Y P T I K P

1057 AGAAACTAAAAAATACCCTATCATGTTATATACACATGGTGGGGGCTGGTTTAAACGGGAGTAAGGA
 111 E T K K Y P I M L Y T H G G G W F N G S K E

1123 AAACATCATGAAATCTCCAGTTAAAGAACCTTTTTTAGAACTCGTAGAACAAGGTTTTGCAGTAGT
 133 N I M K S P V K E P F L E L V E Q G F A V V

1189 ATCAATCAATTACCGCCTCACTAGGCAGAAGTCTGTTTTAATGCGGGATTGTGTTATTGATGCCAT
 155 S I N Y R L T R Q K S V L M R D C V I D A M

1255 GGATGCTTTAAGATATATTTCAAAAAATAGCGATGCTTTGAAATTGGACGCAAATCAGGTATATGT
 177 D A L R Y I S K N S D A L K L D A N Q V Y V

1321 TTTAGGAGATTCGGCAGGTGGGCACATAGCGCAAATGATTACGTTGGCAAACCTAACGATTTTAA
 199 L G D S A G G H I A Q M I T L A N P N D F K

1387 AGGAGATGAAAACCTTTACGGAAATCCATACAAAGTAATTGCTGGCGTTTTTCATGGTACGGACCATC
 221 G D E N L Y G N P Y K V I A G V S W Y G P S

1453 TGATTTACCATAAAAAAATTATTTGAAACCGATGATCCCACAAAAGAAGCAGATAGATTCAGTTC
 243 D F T I K K L F E T D D P T K E A D R F S S

1519 TAGAATAACAAAAGAAGAATCAGATCCCGAAAAAATACAGGTTATGTATAGAGAGATGAGCCCCAT
 265 R I T K E E S D P E K I Q V M Y R E M S P I

1585 ATTTTATTTAACCAAAAAAGTCCGCCATTATTTATGATGGCAGCAGATAATGATACTACAATACC
 287 F Y L T K K S P P L F M M A A D N D T T I P

1651 AGTAGGGCATGCTTACCATATGAAAGAGAAGGCAGATAGTATAGGTGCCAATGTTGACTTGTTTCAT
 309 V G H A Y H M K E K A D S I G A N V D L F I

1717 TGTA AAAAATGCTGGACATAATTGGCGAAAAGCTGGAGGAAACATCAACCCA ACTCTTGAAGAAAT
 331 V K N A G H N W R K A G G N I N P T L E E I

1783 CACAAAAAACAGTAGATTTTTTCTTGAGGTACAAATAACTATTTTTTTATTAAAAAGGTAAAAATGA
 353 T K K T V D F F L R Y K *

1849 GAGGATATGATTTGCTAGTGAAAATATTTCAAGGAACATATTATTTATTAGAAAAATCAAAGTGCT

1915 TAAAACATTGTTCTTAAATATGAATTGAATATCAA ACTAACTATATACTAACTAAATAAATCTTT

1981 TTTATTATGAAAAAACACTACCAACTAAAAAGTCTAATTTATGGTTTCCTTATGGCTTGTTTAAATA
 1 M K K T L P T K K S N L W F L M A C L I

2047 ATAACCTTCCATAAGGTAGAGGCACAAGTACCAGATCCAAACCAAGGATTACGTGCCGAATGGATG
 21 I T F H K V E A Q V P D P N Q G L R A E W M

2113 CGAGGCGCATTGGGCATGCTTTGGTTGCCCGAGCGTACTTTCAACGGTAACATTGAAGGCATAAGA
 43 R G A L G M L W L P E R T F N G N I E G I R

2179 ATTGATGACTTTTTAACCCAAATTAAGATATTAGAACGGTTGATTATGTTCAACTTCCATTAACA
 65 I D D F L T Q I K D I R T V D Y V Q L P L T

2245 AGCCCAAATATTTTTTCTCCA ACTCATGTAGCTCCACATCCTATCATTGAAAGTCTATGGCAAGGT
 87 S P N I F S P T H V A P H P I I E S L W Q G

2311 GATACAGATGCTAATGGAGATCCAATTAATTTGGTGGCTCCACGGGAATCAGTTGATGATCCACTT
 109 D T D A N G D P I N L V A P R E S V D D P L

2377 TTAAGTTGGCTCAAGGCCCTTAGGGCTGCGGGTTTAAGAACAGAGATTTATGTGAATTCCTATAAT
 131 L S W L K A L R A A G L R T E I Y V N S Y N

2443 TTATTAGCAAGGATTCCCGAAGATACGCAAGCAGATTACCCAGACGTTTCAGCACGCTGGATGGAG
 153 L L A R I P E D T Q A D Y P D V S A R W M E

2509 TGGTGTGATACTAACACTGAGGCTCAAGCTTTTATTAATAGCCAAACTTATCATGAAGGTAATGGC
 175 W C D T N T E A Q A F I N S Q T Y H E G N G

2575 AGACGTAAATACATGTTCTGTTATGCAGAATTTATCCTTAAAGAATATGCCCAGCGTTATGGAGAT
 197 R R K Y M F C Y A E F I L K E Y A Q R Y G D

2641 TTAATTGATGCGTGGTGTGTTTGGATTTCGGCAGATAATGTTATGGAAGATGAGTGTGGAGACGATCCT
 219 L I D A W C F D S A D N V M E D E C G D D P

2707 GCTTCCGAAGATGTAAACGATCAACGAATTTATCAGGCCTTTGCAGATGCATGTCATGCAGGGAAT
 241 A S E D V N D Q R I Y Q A F A D A C H A G N

2773 CCGAACGCGGCCATTGCATTTAATAACAGCGTTGGTGATAGAGAGGGAAACCCATTCACATCAGCA
 263 P N A A I A F N N S V G D R E G N P F T S A

2839 ACTCTTTTTGATGATTATACGTTTGGCCATCCTTTTTGGCGGCGCAGGTAACATGGTGGTTCAGAA
 285 T L F D D Y T F G H P F G G A G N M V V P E

2905 GCTTTATACACTTACAATCATGATTTAGTTGTGTTTCATGCAAACCAACAATGGTTATGCTTTTAGA
 307 A L Y T Y N H D L V V F M Q T N N G Y A F R

2971 GATGATACAAGAACTTGAACGACAATGTAGTTGCGCATTTCTTTTCTTAAACAGAGTACCACATCG
 329 D D T R T W N D N V V A H F F P K Q S T T S

3037 TGGAATGCAGGAAATACACCTTGTCTTACAGATGAGCAATTCGTAGAATGGACTAGTACAGGCATT
 351 W N A G N T P C L T D E Q F V E W T S T G I

3103 GTAAATGGAGGTGGCATCACTTGGGGAACACCGCTTGTTAGAACAATTTGGAAAATGCTCCAGTG
 373 V N G G G I T W G T P L V R T N L E N A P V

3169 CTTACTTTACAACCTTATGCGCTTAATCAGTTTCGAACTAACAGATACTTATTTGAAAAGAATTTCAA
 395 L T L Q P Y A L N Q F E L T D T Y L K E F Q

3235 TCACCAGGAAAGCCTAATTGGTCAAGACAATATAACAATCTTGCCAGCCATTTACCCAGGTCAGCCT
 417 S P G K P N W S R Q Y T I L P A I Y P G Q P

3301 TATTCTCATAATTTGGTTGAAGGCGTAGATTTTTGGGATCCAGAAGGAGTCGGTATAACCGGGCTT
 439 Y S H N L V E G V D F W D P E G V G I T G L

3367 ACAGCGTCTGGAACCTTTACCAGCATGGTTAACAATTAGTCAAACAGCTACAGGAACTTGGACGTTA
 461 T A S G T L P A W L T I S Q T A T G T W T L

3433 AGTGGAACTCCGCTGTATCTGAGGCTTCTAATTATAACCTTTGAACTTATGGCTCAGGATAGTGAT
 483 S G T P P V S E A S N Y T F E L M A Q D S D

3499 GGAGTAACAAACAGAGAAGTGAAGTTAGAGGTGATTTACATCCTGCAGGTTTTACAAATCCTGGA
 505 G V T N R E V K L E V I S H P A G F T N P G

3565 GATGGTACTCCAGTATGGTTTTCGAATCCCATGGTACTTGCGAAAGCAACAGCATTAAAGGATTAT
 527 D G T P V W F S N P M V L A K A T A L K D Y

3631 GGAAGCCTTTTGAACCTTGGTGTGATTTTTATGATTTTGAAGGAGATGTTCTTACTATAACTAAA
 549 G S L L K L G V D F Y D F E G D V L T I T K

3697 ACATCAGGACCTGATTGGCTTGTTTTAACTCAGAATTCAGATGATACTTGGCGTTTAAAGTGGGATG
 571 T S G P D W L V L T Q N S D D T W R L S G M

3763 CCAACAGCTGCTGATGCAGGTGAAAATTCTTTTACATTTAATGTTAGCGATGGTATTCTTTCATCA
 593 P T A A D A G E N S F T F N V S D G I L S S

3829 GATACAGAGATTAATAATTACAGTTGATCATGTTGCAGGTTTTACTAATTTGGGTAATGGAGCACCT
 615 D T E I K I T V D H V A G F T N L G N G A P

3895 GTTTGGTCTTCACCTATACTTAATCTTACCGATGGAAAAGGATCATTTCGCCTATAAATTATACTTTG
 637 V W S S P I L N L T D G K G S F A Y N Y T L

3961 CAATTAGGAACCGATTATTATGATTTTTGAAGGTGATGCACTTACTATTACTAAAAACATCAGGACCT
 659 Q L G T D Y Y D F E G D A L T I T K T S G P

4027 GATTGGCTTACTATTCAACAAACCGATGCAAATAGTTGGAAATTAAGTGGAACTCCAATAAATTCA
 681 D W L T I Q Q T D A N S W K L S G T P I N S

4093 GATGCGGGTGAAAATTCTTTTACGTTCAATCTCAGCGATGATACCAATTCAACAACCTGCTGAAAAT
 703 D A G E N S F T F N L S D D T N S T T A E I

4159 CTGATCAACGTAATTGCAACTATTATTAGTGATGGAAATGTTGAAATTAAGCAACTGCCAATACA
 725 L I N V I A T I I S D G N V E I K A T A N T

4225 ACTTATGGTATTAATACTGTTGCTAGTATGTATTCCGCAGTACAAACGGCTCCAGATGGACTAGCG
 747 T Y G I N T V A S M Y S A V Q T A P D G L A

4291 ACTTATAGAATTTCTATCGACGTTACACCTCCAACTGATAAAGGTATTTATTCTGGTTCTTCAGGA
 769 T Y R I S I D V T P P T D K G I Y S G S S G

4357 GGTATAACGACTACGACGTCATGGGGAATCGGTGATGGTACAGACGCTATTCAAAATACTATATTC
 791 G I T T T T S W G I G D G T D A I Q N T I F

4423 CGAGGGAGTGATAATGAATGGACAGAAAGTATTAACAATATTAATAATGGTTGATTTTAATGCAAAT
 813 R G S D N E W T E S I N N I K M V D F N A N

4489 GGAGGATCTCTTACGACGGATAATGTTACCATGTTTTTTAAATCTATTTCTATAGGGAATTCACAA
 835 G G S L T T D N V T M F F K S I S I G N S Q

4555 TCTGTAAATGATTTTGTATCGCTCAAAGTTGGAGGAGTAATATCAAATCCAGGGAGATCAGCAAAC
 857 S V N D F V S L K V G G V I S N P G R S A N

4621 CAATACGAAACCATCGATTTAACCTTAGCTACATCAGTGAGTAATATTGCAAATTTTGCTATTGGA
 879 Q Y E T I D L T L A T S V S N I A N F A I G

4687 ACAGGTAATGATTCCGATACTAATAAATGGTCTGTTGAAGGTATAACAGTATTTCGTAGACTTTGCA
 901 T G N D S D T N K W S V E G I T V F V D F A

4753 GGCACGCTTTCCGTTACTAATCCAATACAGGATGATGTTGACTCTTTCAAGCTATATCCTAACCCA
 923 G T L S V T N P I Q D D V D S F K L Y P N P

4819 GCAAAAGATAGAATTTTCATAAACAAACAACCTGTTACTGTTTCAGATTTTTGATGTAACCTGGTAAG
 945 A K D R I F I N K Q P V T V Q I F D V T G K

4885 CTTGTAAAAGTTGATTTTAAAGGAAAAAATGAACTTGATATTTCAACTTTAAAGCAAGGATTATAT
 967 L V K V D F K G K N E L D I S T L K Q G L Y

4951 ATTTTGAAAATACAAACCGCTGAAGGGAATATGTTGTTTCGAAAATTCATTAAGAAATAAATTATC
 989 I L K I Q T A E G N M L F E K F I K K *

5017 TAATAAATAGATAATAACTGTATTTATAAATTAAAATGATGATTGATAAAAAATAATTCTAGTTG
1 M M I D K K I I L V

5083 TATGTTGCCTCATGGCATCCATAGGTATTTATGCGCAAACACAATCAAAAAAGTCAAAAAACCAATA
11 V C C L M A S I G I Y A Q T Q S K S Q K P N

5149 TCATTTTCATTTTTGCTGACGATTGGGGTTATGGCGATTTAAGTGCTCATGGTAGTACTTGGGTAA
33 I I F I F A D D W G Y G D L S A H G S T W V

5215 AAACACCTAACATTGATAAAATGATTGGCGATGGCATGGATTTTTCCAATTTCACTGTAAACAGCC
55 K T P N I D K M I G D G M D F S N F T V N S

5281 CAGTGTGTTACCTAGTAGAGTTGCTGTTATGACAGGGCAATTTCTGCTAGACAATCTATTCATC
77 P V C S P S R V A V M T G Q F P A R Q S I H

5347 AACATTTTCAAGGATGGAAAGCGCATGAAAACAGAGGCATGCCAGATTGGATGGATCCAGAAGGTA
99 Q H F Q G W K A H E N R G M P D W M D P E G

5413 TGTCGTTTTCCAAGAGAATTTCAAAAAGCAGGTTATGTGACTGCGCATTTTGAAAAATGGCACTTAG
121 M S F P R E F Q K A G Y V T A H F G K W H L

5479 GGGCGGCTCCAGATGTGCCTAAAGAAAGCGCTTACGGTTACGATGAGTATGCAACGTTTAATGGTT
143 G A A P D V P K E S A Y G Y D E Y A T F N G

5545 CTAAAAACATTGATATTCCTAAGGCAGGTTCTATAGGTGTCGATTATGCCGAAGATTTTCATCAAAA
165 S K N I D I P K A G S I G V D Y A E D F I K

5611 GAAATAAGACAAACCGTTCTTCATCAATTTATGGTTACATGAGGCGCATACGGCACATTATCCTT
187 R N K D K P F F I N L W L H E A H T A H Y P

5677 TAGAAAAGTTTATGGATAAGTTTGGTAATTTGGATGAACAAAAACAAGTATATGCTTCTATAATAG
209 L E K F M D K F G N L D E Q K Q V Y A S I I

5743 CTGAAGGTGATGAAGCTGTTGGAAGAATTGTAAATTTATTAGAAGAATTAGGTCTTGATGAAAATA
231 A E G D E A V G R I V N L L E E L G L D E N

5809 CTTTAGTGGTGTCTCGACAGATAATGGTCCAGAATGGGAAGGTACAGAGAAAAGATAAACTTCATA
253 T L V V F S T D N G P E W E G T E K D K L H

5875 AACCAGAGAAGGGTGTGACGATGATAATGGACCAGTAGGTCTTGAAAAATATTATTCGGTAGGTG
275 K P E K G D D D D N G P V G L G K Y Y S V G

5941 AAACAGGAGGTTTAAAAGGTCAAAAACGCTCTTTGTTTGTGCGGGAATTCGTGTGCCTTTTGTG
297 E T G G L K G Q K R S L F A G G I R V P F V

6007 CTAAATGGCCTAAAGTTATTCCTGAAGGCGTTAAAAACAATAGTGCTGTAATAACCGCAGTAGATT
319 A K W P K V I P E G V K N N S A V I T A V D

6073 TATTACCTACTTTTTTATAAGCTGCAAATATTCCAATGCCAAAAGGGTATCAACCAGACGGTGAAA
341 L L P T F L E A A N I P M P K G Y Q P D G E

6139 GTATGATGTCTGCTTTTTAAAGGCAACATATTTGAAAGAAGTAAACCCATTTTTTGGGAATGGAAAG
363 S M M S A F K G N I F E R S K P I F W E W K

6205 GTGGAGACAATTTTTGTTTATACGTGGCCAACTTTAGGAGTTAGAGATGGAGATTATAAATTGGTGG
385 G G D N F V Y T W P T L G V R D G D Y K L V

6271 TAGATGTTACGGGAGAAAAATATGAATTGTTTGTATATCAAAAACGATTGGAAAGAAGAACAATC
407 V D V T G E K Y E L F D I K N D W K E E H N

6337 TTGCCGATACCAATCCTAAAAAGGCACAAGAATTGTTAGACTTAGTGTATGAATGGAAAAAACAC
429 L A D T N P K K A Q E L L D L V Y E W K K T

6403 TTCCAGAAAAACCTAGGGAAAGTTGTTTGTCTTCAGCAAGAAACAAGGTAAAAGCCTGTAGATAAAG
451 L P E K P R E S C L S S A R N K V K P V D K

6469 TAAAGAAGGAAAAAAGTAAAGAAAATAAAAATGAAGACGATTAAAAGATTATAGGCATAAAATGAA
473 V K K E K S K E N K N E D D *

6535 ATTTCCATAACAAAATTTCAATATCCAAGATAATCACTATGTGGAAGTTACATGTAACCAATAAAA

6601 AAAATAAATAGTATATATATAGACAGATGAAAAAACTATTTCAACTAGTATTCATAGCAACAGTTTTT
1 M K K L F Q L V F I A T V F

6667 CTAAGTTGTAAAACAACCTAAGACGCTCACTTATGAACAAATGAATGTTTCAAAGAACAGCCTAAA
15 L S C K T T K T L T Y E Q M N V S K E Q P K

6733 GAAGCGTTCAATGATGCTAAATATGGTATGTTTATCCATTGGGGTTTGTATGCCATTCCAGGAGGT
37 E A F N D A K Y G M F I H W G L Y A I P G G

6799 ATTTGGAAAGGGAAAAAGATGGAGGAATTAAGGCCCTAAAGTAGCAGAGTGGATTCAATTTGGT
59 I W K G K K M E E L K G P K V A E W I Q F G

6865 GCCGAAATACCAAGGGACGAATATGCACAATTAGCAACGCAATTTAACCCTCAATTTGATGCC
81 A E I P R D E Y A Q L A T Q F N P T Q F D A

6931 GATGCCATTGCAAAATTAGCAAAAGATGCCGGATGAAATATCTAGTAATCACATCTAAGCATCAT
103 D A I A K L A K D A G M K Y L V I T S K H H

6997 GATGGCTTTGCTATGTACGATTCTAAAGTAAGCGACTACGATATTGTAGATGCAACACCCTATAAA
125 D G F A M Y D S K V S D Y D I V D A T P Y K

7063 AAAGATGTGGTGAAACAACCTGTATGATGCCTGTAAAAACAAGGTATCGATTTTGGTTTATATTAC
147 K D V V K Q L Y D A C K K Q G I D F G L Y Y

7129 TCACACAATATTGATTGGATGGATGGGAATGATTGCGGTCTTTCCGAATATTTAGCCTCAGGAGGT
169 S H N I D W M D G N D C G L S E Y L A S G G

7195 GAAGAACACGATAGGGTAAAGCGTAAAGCAGGTGTTAACTTGGGATCCAAGCCCAAATACCTTT
191 E E H D R V K R K A G V N T W D P S P N T F

7261 TCAGAATACTTGAAAAATAAAGCGTATCCACAGGTAAAGGAAATATTAACAAAGTTTCCTGATTG
213 S E Y L E N K A Y P Q V K E I L T K F P D L

7327 AAAACGTTGTGGTACGATTACCCGCATTATGTAACGCCAGAACAAGTTTAAAGTTTATAAAAAATA
235 K T L W Y D Y P H Y V T P E Q S L K F Y K I

7393 GCTTACGATTTACAACCAACATGTTAGTGAATTCTAGAGTTGGAAATGAATTAGGAGATTTGAT
257 A Y D L Q P N M L V N S R V G N E L G D F D

7459 ATTCAGGAGACAACAAAATCCCTCAAGATCATTAGCGATTACTAAACCTTGGCAAACCGTAGGT
279 I P G D N K I P Q D H L A I T K P W Q T V G

7525 ACCACAAACAATTCTTGGGGATACAATTCTTATGATAACGATTGGAAATCTGTAAAAGAACTATTG
301 T T N N S W G Y N S Y D N D W K S V K E L L

7591 TTCTGGTTAACCGAAATTGTAAGTAAAGGGGGGAATTATATGTTGAATATAGGGCCAGATGCCACA
323 F W L T E I V S K G G N Y M L N I G P D A T

7657 GGTCATGTGCCTTCTGAATCGGTAAATAATCTATTAGAAGTTGGAAAGTGGTTGGAAGTTAATGGT
345 G H V P S E S V N N L L E V G K W L E V N G

ORF 1 : SulB, partiel

Universal code

Total amino acid number: 230, MW=26057

Max ORF: 2-691

```
1      TGTGGATCGTTATTTAGTAACAGATACACAAAGGATTTTCATGGCCAGTAAAAGGAAAACA
1      V D R Y L V T D T Q R I S W P V K G K Q

61     AAGTTGTGTTATGAATGGAGATTGGCGATTGGTTAATGGAAATGAACTGTATAACACAGC
21     S C V M N G D W R L V N G N E L Y N T A

121    TACCGACCCAGGACAAAAAATAAACCTAGCTGATAAACATCCTGATAAAGTGGCCAATAT
41     T D P G Q K I N L A D K H P D K V A N M

181    GAATGATTTTTATAACACATGGTGGGATAGTGTATAAAAAGAACTAAATACTCTACCAT
61     N D F Y N T W W D S V I K E T K Y S T I

241    CGATTTAGGAGTTGGTGAAGAAGATGTTATTACCTGTCATGATGCGAGAACAATTGATTA
81     D L G V G E E D V I T C H D A R T I D Y

301    TTTCCC GCCATGGAATCAAGAAATGATAAGAAAAGGAGAGCCTATGAAACCAGCACCATT
101    F P P W N Q E M I R K G E P M K P A P F

361    CTTCGTGAATTTTCGTAAAACCGGGTAACTATACATTTAAGTTAAGACGTTGGCCTGCTGA
121    F V N F V K P G N Y T F K L R R W P A E

421    AAGTGGCATGTCGTTAGGAGCTGAAATTTTAGATGAAATACCGGCAACTCCCACAACCGA
141    S G M S L G A E I L D E I P A T P T T D

481    TGGTAGAAGTATTGGAAAAGCCATGAAGTTTAAAAAAGCATTTTTAAAAATAGGAGACCA
161    G R S I G K A M K F K K A F L K I G D Q

541    AGAAGTTTCTGTTGAGGCCAATAATGCGAGTCAATCTGCAGATATAGATTTTCGTCGTTGA
181    E V S V E A N N A S Q S A D I D F V V E

601    GCAAGGAAAAACCGAGTTATTGGCTTGGTTTGGAGATGGAAAACGGCACATTAACAAATGC
201    Q G K T E L L A W F E M E N G T L T N A

661    TTTTTATGTGAATGTTGAAAAAAGCAAAGAATAGCATAGAGTTAAAGTAAATGGTAACAT
221    F Y V N V E K S K E
```

ORF 2 : EstA, complet

Universal code

Total amino acid number: 364, MW=41444

Max ORF: 728-1819

```
721    TAAAATTATGAAAAGGTTTTTAAATATATGTTTTATCCATTGGTTTTGCAATAGGGCTCAA
1      M K R F L I Y V L S I G F A I G L K

781    AGCACAGAGTTCTAGCCAAAAAGTGGTATCTAAAAATGAATTTTCAAACTGGATAGCAA
19     A Q S S S Q K V V S K N E F S K L D S N

841    TAAAACGGCAGCCTTGAAGAGATGAATTATCACGGATTTGGACTCAAATAAAACAGCA
39     K N G T L E R D E L S R I W T Q I K Q H
```

901 TGATACCGATAATAATGGAAAGATTACATTAGAAGAATTTTCAAAATTTGAAATTCCACA
 59 D T D N N G K I T L E E F S K F E I P H
 961 TTTAGAAACTAAAGGCGATATAATTTTAAATATTTAAATACAAAACAACAAAAGAAGA
 79 L E T K G D I I L N I K Y K T T K E E D
 1021 TTTGTATTTAGATGTTTATTACCCCAACAATAAAACCAGAACTAAAAAATACCCTATCAT
 99 L Y L D V Y Y P T I K P E T K K Y P I M
 1081 GTTATATACACATGGTGGGGGCTGGTTTAAACGGGAGTAAGGAAAACATCATGAAATCTCC
 119 L Y T H G G G W F N G S K E N I M K S P
 1141 AGTTAAAGAACCTTTTTTAGAACTCGTAGAACAAGGTTTTGCAGTAGTATCAATCAATTA
 139 V K E P F L E L V E Q G F A V V S I N Y
 1201 CCGCCTCACTAGGCAGAAGTCTGTTTTAATGCGGGATTGTGTTATTGATGCCATGGATGC
 159 R L T R Q K S V L M R D C V I D A M D A
 1261 TTTAAGATATATTTCAAAAAATAGCGATGCTTTGAAATTGGACGCAAATCAGGTATATGT
 179 L R Y I S K N S D A L K L D A N Q V Y V
 1321 TTTAGGAGATTTCGGCAGGTGGGCACATAGCGCAAATGATTACGTTGGCAAACCCTAACGA
 199 L G D S A G G H I A Q M I T L A N P N D
 1381 TTTTAAAGGAGATGAAAACCTTTACGGAAATCCATACAAAGTAATTGCTGGCGTTTCATG
 219 F K G D E N L Y G N P Y K V I A G V S W
 1441 GTACGGACCATCTGATTTCCACATAAAAAAATTATTTGAAACCGATGATCCCACAAAAGA
 239 Y G P S D F T I K K L F E T D D P T K E
 1501 AGCAGATAGATTTCAGTTCTAGAATAACAAAAGAAGAATCAGATCCCGAAAAAATACAGGT
 259 A D R F S S R I T K E E S D P E K I Q V
 1561 TATGTATAGAGAGATGAGCCCCATATTTTATTTAACCACAAAAAAGTCCGCCATTATTTAT
 279 M Y R E M S P I F Y L T K K S P P L F M
 1621 GATGGCAGCAGATAATGATACTACAATACCAGTAGGGCATGCTTACCATATGAAAGAGAA
 299 M A A D N D T T I P V G H A Y H M K E K
 1681 GGCAGATAGTATAGGTGCCAATGTTGACTTGTTTCATTGTAAAAAATGCTGGACATAATTG
 319 A D S I G A N V D L F I V K N A G H N W
 1741 GCGAAAAGCTGGAGGAAACATCAACCCAACCTTGAAGAAATCACAAAAAAAACAGTAGA
 339 R K A G G N I N P T L E E I T K K T V D
 1801 TTTTTTCTTGAGGTACAAATAACTATTTTTTTATTTAAAAGGTAAAATGAGAGGATATGATT
 359 F F L R Y K

ORF 3 : FcnA, complet

Universal code

Total amino acid number: 1007, MW=110285

Max ORF: 1987-5007

1981 TTTATTATGAAAAAACAACACTACCAACTAAAAAGTCTAATTTATGGTTCCTTATGGCTTGT
 1 M K K T L P T K K S N L W F L M A C
 2041 TTAATAATAACCTTCCATAAGGTAGAGGCACAAGTACCAGATCCAAACCAAGGATTACGT
 19 L I I T F H K V E A Q V P D P N Q G L R

2101 GCCGAATGGATGCGAGGCGCATTGGGCATGCTTTGGTTGCCCGAGCGTACTTTCAACGGT
 39 A E W M R G A L G M L W L P E R T F N G

2161 AACATTGAAGGCATAAGAATTGATGACTTTTTAAACCCAAATTAAGATATTAGAACGGTT
 59 N I E G I R I D D F L T Q I K D I R T V

2221 GATTATGTTCAACTTCCATTAACAAGCCCAAATATTTTTCTCCAACATCATGTAGCTCCA
 79 D Y V Q L P L T S P N I F S P T H V A P

2281 CATCCTATCATTGAAAGTCTATGGCAAGGTGATACAGATGCTAATGGAGATCCAATTAAT
 99 H P I I E S L W Q G D T D A N G D P I N

2341 TTGGTGGCTCCACGGGAATCAGTTGATGATCCACTTTTAAGTTGGCTCAAGGCCCTTAGG
 119 L V A P R E S V D D P L L S W L K A L R

2401 GCTGCGGGTTTAAGAACAGAGATTTATGTGAATTCCTATAATTTATTAGCAAGGATTCCC
 139 A A G L R T E I Y V N S Y N L L A R I P

2461 GAAGATACGCAAGCAGATTACCCAGACGTTTCAGCACGCTGGATGGAGTGGTGTGATACT
 159 E D T Q A D Y P D V S A R W M E W C D T

2521 AACACTGAGGCTCAAGCTTTTATTAATAGCCAAACTTATCATGAAGGTAATGGCAGACGT
 179 N T E A Q A F I N S Q T Y H E G N G R R

2581 AAATACATGTTCTGTTATGCAGAATTTATCCTTAAAGAATATGCCAGCGTTATGGAGAT
 199 K Y M F C Y A E F I L K E Y A Q R Y G D

2641 TTAATTGATGCGTGGTGTGTTTTGATTCCGGCAGATAATGTTATGGAAGATGAGTGTGGAGAC
 219 L I D A W C F D S A D N V M E D E C G D

2701 GATCCTGCTTCCGAAGATGTAAACGATCAACGAATTTATCAGGCCTTTGCAGATGCATGT
 239 D P A S E D V N D Q R I Y Q A F A D A C

2761 CATGCAGGGAATCCGAACGCGGCCATTGCATTTAATAACAGCGTTGGTGATAGAGAGGGA
 259 H A G N P N A A I A F N N S V G D R E G

2821 AACCCATTCACATCAGCAACTCTTTTTGATGATTATACGTTTGGCCATCCTTTTTGGCGGC
 279 N P F T S A T L F D D Y T F G H P F G G

2881 GCAGGTAACATGGTGGTTCAGAAGCTTTATACACTTACAATCATGATTTAGTTGTGTTC
 299 A G N M V V P E A L Y T Y N H D L V V F

2941 ATGCAAACCAACAATGGTTATGCTTTTAGAGATGATACAAGAAGCTTGAACGACAATGTA
 319 M Q T N N G Y A F R D D T R T W N D N V

3001 GTTGCGCATTTCTTTCCCTAAACAGAGTACCACATCGTGGAATGCAGGAAAATACACCTTGT
 339 V A H F F P K Q S T T S W N A G N T P C

3061 CTTACAGATGAGCAATTCGTAGAATGGACTAGTACAGGCATTGTAAATGGAGGTGGCATC
 359 L T D E Q F V E W T S T G I V N G G G I

3121 ACTTGGGGAACACCGCTTGTTAGAACAAATTTGGAAAATGCTCCAGTGCTTACTTTACAA
 379 T W G T P L V R T N L E N A P V L T L Q

3181 CCTTATGCGCTTAATCAGTTCGAACTAACAGATACTTATTTGAAAGAATTTCAATCACCA
 399 P Y A L N Q F E L T D T Y L K E F Q S P

3241 GGAAAGCCTAATTGGTCAAGACAATATACAATCTTGCCAGCCATTTACCCAGGTCAGCCT
 419 G K P N W S R Q Y T I L P A I Y P G Q P

3301 TATTCTCATAATTTGGTTGAAGGCGTAGATTTTTGGGATCCAGAAGGAGTCGGTATAACC
439 Y S H N L V E G V D F W D P E G V G I T

3361 GGGCTTACAGCGTCTGGAACCTTACCAGCATGGTTAAACAATTAGTCAAACAGCTACAGGA
459 G L T A S G T L P A W L T I S Q T A T G

3421 ACTTGACGTTAAGTGGAACCTCCGCCTGTATCTGAGGCTTCTAATTATACCTTTGAACTT
479 T W T L S G T P P V S E A S N Y T F E L

3481 ATGGCTCAGGATAGTGATGGAGTAACAAACAGAGAAGTGAAGTTAGAGGTGATTTACACAT
499 M A Q D S D G V T N R E V K L E V I S H

3541 CCTGCAGGTTTTACAAATCCTGGAGATGGTACTCCAGTATGGTTTTCGAATCCCATGGTA
519 P A G F T N P G D G T P V W F S N P M V

3601 CTTGCGAAAGCAACAGCATTAAAGGATTATGGAAGCCTTTTTGAACTTGGTGTGATTTT
539 L A K A T A L K D Y G S L L K L G V D F

3661 TATGATTTTGAAGGAGATGTTCTTACTATAACTAAAACATCAGGACCTGATTGGCTTGT
559 Y D F E G D V L T I T K T S G P D W L V

3721 TTAACTCAGAATTCAGATGATACTTGGCGTTTAAAGTGGGATGCCAACAGCTGCTGATGCA
579 L T Q N S D D T W R L S G M P T A A D A

3781 GGTGAAAATTCTTTTACATTTAATGTTAGCGATGGTATTCTTTCATCAGATACAGAGATT
599 G E N S F T F N V S D G I L S S D T E I

3841 AAAATTACAGTTGATCATGTTGCAGGTTTTACTAATTTGGGTAATGGAGCACCTGTTTGG
619 K I T V D H V A G F T N L G N G A P V W

3901 TCTTCACCTATACTTAATCTTACCGATGGAAAAGGATCATTTCGCTATAAATTATACTTTG
639 S S P I L N L T D G K G S F A Y N Y T L

3961 CAATTAGGAACCGATTATTATGATTTTGAAGGTGATGCACTTACTATTACTAAAACATCA
659 Q L G T D Y Y D F E G D A L T I T K T S

4021 GGACCTGATTGGCTTACTATTCAACAAACCGATGCAAATAGTTGGAAATTAAGTGGAACT
679 G P D W L T I Q Q T D A N S W K L S G T

4081 CCAATAAATTCAGATGCGGGTGAAAATTCTTTTACGTTCAATCTCAGCGATGATACCAAT
699 P I N S D A G E N S F T F N L S D D T N

4141 TCAACAACCTGCTGAAATTTCTGATCAACGTAATTGCAACTATTATTAGTGATGGAAATGTT
719 S T T A E I L I N V I A T I I S D G N V

4201 GAAATTAAGCAACTGCCAATACTTATGGTATTAATACTGTTGCTAGTATGTATTCC
739 E I K A T A N T T Y G I N T V A S M Y S

4261 GCAGTACAAACGGCTCCAGATGGACTAGCGACTTATAGAATTTCTATCGACGTTACACCT
759 A V Q T A P D G L A T Y R I S I D V T P

4321 CCAACTGATAAAGGTATTTATTCTGGTTCTTCAGGAGGTATAACGACTACGACGTCATGG
779 P T D K G I Y S G S S G G I T T T T S W

4381 GGAATCGGTGATGGTACAGACGCTATTCAAAAATACTATATTTCCGAGGGAGTGATAATGAA
799 G I G D G T D A I Q N T I F R G S D N E

4441 TGGACAGAAAGTATTAACAATATTAATAATGGTTGATTTTAAATGCAAATGGAGGATCTCTT
819 W T E S I N N I K M V D F N A N G G S L

4501 ACGACGGATAATGTTACCATGTTTTTTAAATCTATTTCTATAGGGAATTCACAATCTGTA
839 T T D N V T M F F K S I S I G N S Q S V

4561 AATGATTTTGTATCGCTCAAAGTTGGAGGAGTAATATCAAATCCAGGGAGATCAGCAAAC
859 N D F V S L K V G G V I S N P G R S A N

4621 CAATACGAAACCATCGATTTAACCTTAGCTACATCAGTGAGTAATATTGCAAATTTTGCT
879 Q Y E T I D L T L A T S V S N I A N F A

4681 ATTGGAACAGGTAATGATTCCGATACTAATAAATGGTCTGTTGAAGGTATAACAGTATTC
899 I G T G N D S D T N K W S V E G I T V F

4741 GTAGACTTTGCAGGCACGCTTTCCGTTACTAATCCAATACAGGATGATGTTGACTCTTTC
919 V D F A G T L S V T N P I Q D D V D S F

4801 AAGCTATATCCTAACCCAGCAAAAGATAGAATTTTCATAAACAAACAACCTGTTACTGTT
939 K L Y P N P A K D R I F I N K Q P V T V

4861 CAGATTTTTGATGTAAGTGGTAAGCTTGTAAGTTGATTTTAAAGGAAAAAATGAACTT
959 Q I F D V T G K L V K V D F K G K N E L

4921 GATATTTCAACTTTAAAGCAAGGATTATATATTTTGAAAATACAAACCGCTGAAGGGAAT
979 D I S T L K Q G L Y I L K I Q T A E G N

4981 ATGTTGTTTCGAAAAATTCATTAAGAAATAAATTATCTAATAAATAGATAATAACTGTATT
999 M L F E K F I K K

ORF 4 : Sula, complet

Universal code

Total amino acid number: 486, MW=54620

Max ORF: 5052-6509

5041 TATAAATTAATAATGATGATTGATAAAAAATAATTCTAGTTGTATGTTGCCTCATGGCAT
1 M M I D K K I I L V V C C L M A

5101 CCATAGGTATTTATGCGCAAACACAATCAAAAAGTCAAAAACCCAATATCATTTTCATTT
17 S I G I Y A Q T Q S K S Q K P N I I F I

5161 TTGCTGACGATTGGGGTTATGGCGATTTAAGTGCTCATGGTAGTACTTGGGTAAAAACAC
37 F A D D W G Y G D L S A H G S T W V K T

5221 CTAACATTGATAAAATGATTGGCGATGGCATGGATTTTTCCAATTTCACTGTAAACAGCC
57 P N I D K M I G D G M D F S N F T V N S

5281 CAGTGTGTTACCTAGTAGAGTTGCTGTTATGACAGGGCAATTTCTGCTAGACAATCTA
77 P V C S P S R V A V M T G Q F P A R Q S

5341 TTCATCAACATTTTCAAGGATGGAAAGCGCATGAAAACAGAGGCATGCCAGATTGGATGG
97 I H Q H F Q G W K A H E N R G M P D W M

5401 ATCCAGAAGGTATGTCGTTTTCCAAGAGAATTTCAAAAAGCAGGTTATGTGACTGCGCATT
117 D P E G M S F P R E F Q K A G Y V T A H

5461 TTGGAAAATGGCACTTAGGGGCGGCTCCAGATGTGCCTAAAGAAAGCGCTTACGGTTACG
137 F G K W H L G A A P D V P K E S A Y G Y

5521 ATGAGTATGCAACGTTTAATGGTTCTAAAAACATTGATATTCCTAAGGCAGGTTCTATAG
157 D E Y A T F N G S K N I D I P K A G S I

5581 GTGTCGATTATGCCGAAGATTTTCATCAAAAAGAAATAAAGACAAACCGTTCTTCATCAATT
 177 G V D Y A E D F I K R N K D K P F F I N

5641 TATGGTTACATGAGGCGCATACGGCACATTATCCTTTAGAAAAGTTTATGGATAAGTTTG
 197 L W L H E A H T A H Y P L E K F M D K F

5701 GTAATTTGGATGAACAAAAACAAGTATATGCTTCTATAATAGCTGAAGGTGATGAAGCTG
 217 G N L D E Q K Q V Y A S I I A E G D E A

5761 TTGGAAGAATTGTAAATTTATTAGAAGAATTAGGTCTTGATGAAAAATACTTTAGTGGTGT
 237 V G R I V N L L E E L G L D E N T L V V

5821 TCTCGACAGATAATGGTCCAGAATGGGAAGGTACAGAGAAAGATAAACTTCATAAACCAG
 257 F S T D N G P E W E G T E K D K L H K P

5881 AGAAGGGTGATGACGATGATAATGGACCAGTAGGTCTTGAAAAATATTATTCGGTAGGTG
 277 E K G D D D D N G P V G L G K Y Y S V G

5941 AAACAGGAGGTTTAAAAGGTCAAAAACGCTCTTTGTTTGCTGGCGGAATTCGTGTGCCTT
 297 E T G G L K G Q K R S L F A G G I R V P

6001 TTGTTGCTAAATGGCCTAAAGTTATTCCTGAAGGCGTTAAAAACAATAGTGCTGTAATAA
 317 F V A K W P K V I P E G V K N N S A V I

6061 CCGCAGTAGATTTATTACCTACTTTTTTTAGAAGCTGCAAATATTCCAATGCCAAAAGGT
 337 T A V D L L P T F L E A A N I P M P K G

6121 ATCAACCAGACGGTGAAAGTATGATGTCTGCTTTTAAAGGCAACATATTTGAAAAGAAGTA
 357 Y Q P D G E S M M S A F K G N I F E R S

6181 AACCCATTTTTTGGGAATGGAAAGGTGGAGACAATTTTTGTTTATACGTGGCCAACTTTAG
 377 K P I F W E W K G G D N F V Y T W P T L

6241 GAGTTAGAGATGGAGATTATAAATTGGTGGTAGATGTTACGGGAGAAAAATATGAATTGT
 397 G V R D G D Y K L V V D V T G E K Y E L

6301 TTGATATCAAAAACGATTGGAAAGAAGAACAATCTTGCCGATACCAATCCTAAAAAGG
 417 F D I K N D W K E E H N L A D T N P K K

6361 CACAAGAATTGTTAGACTTAGTGTATGAATGGAAAAAACACTTCCAGAAAAACCTAGGG
 437 A Q E L L D L V Y E W K K T L P E K P R

6421 AAAGTTGTTTGTCTTCAGCAAGAAACAAGGTAAAGCCTGTAGATAAAGTAAAGAAGGAAA
 457 E S C L S S A R N K V K P V D K V K K E

6481 AAAGTAAAGAAAATAAAAAATGAAGACGATTAAGAGATTATAGGCATAAAAATGAAATTTCC
 477 K S K E N K N E D D

ORF 4 : FcsA, partiel

LES SEQUENCES EN GRAS INDIQUE UNE SEQUENCE INCERTAINE CAR OBTENUE SUR L'ADN GENOMIQUE DE SW5

Universal code

Total amino acid number: 366, MW=41690

Max ORF: 6625-7722

6601 AAAATAAATAGTATATATAGACAGATGAAAAAACTATTTCAACTAGTATTCATAGCAACA
1 M K K L F Q L V F I A T

6661 GTTTTTCTAAGTTGTAAAACAACCTAAGACGCTCACTTATGAACAAATGAATGTTTCAAAA
13 V F L S C K T T K T L T Y E Q M N V S K

6721 GAACAGCCTAAAGAAGCGTTCAATGATGCTAAATATGGTATGTTTATCCATTGGGGTTG
33 E Q P K E A F N D A K Y G M F I H W G L

6781 TATGCCATTCCAGGAGGTATTTGGAAAGGGAAAAAGATGGAGGAATTTAAAGCCCTAAA
53 Y A I P G G I W K G K K M E E L K G P K

6841 GTAGCAGAGTGGATTCAATTTGGTGCCGAAATACCAAGGGACGAATATGCACAATTAGCA
73 V A E W I Q F G A E I P R D E Y A Q L A

6901 ACGCAATTTAACCCAACCTCAATTTGATGCCGATGCCATTGCAAAATTAGCAAAAGATGCC
93 T Q F N P T Q F D A D A I A K L A K D A

6961 GGGATGAAATATCTAGTAATCACATCTAAGCATCATGATGGCTTTGCTATGTACGATTCT
113 G M K Y L V I T S K H H D G F A M Y D S

7021 AAAGTAAGCGACTACGATATTGTAGATGCAACACCCTATAAAAAAGATGTGGTGAAACAA
133 K V S D Y D I V D A T P Y K K D V V K Q

7081 CTGTATGATGCCTGTAAAAACAAGGTATCGATTTTGGTTTATATTACTCACACAATATT
153 L Y D A C K K Q G I D F G L Y Y S H N I

7141 GATTGGATGGATGGGAATGATTGCGGTCTTTTCGGAATATTTAGCCTCAGGAGGTGAAGAA
173 D W M D G N D C G L S E Y L A S G G E E

7201 CACGATAGGGTAAAGCGTAAAGCAGGTGTTAACACTTGGGATCCAAGCCCAAATACCTTT
193 H D R V K R K A G V N T W D P S P N T F

7261 TCAGAATACTTGGAAAATAAAGCGTATCCACAGGTAAAGGAAATATTAACAAAGTTTCCT
213 S E Y L E N K A Y P Q V K E I L T K F P

7321 GATTTGAAAACGTTGTGGTACGATTACCCGATTATGTAACGCCAGAACAAGTTTAAAG
233 D L K T L W Y D Y P H Y V T P E Q S L K

7381 TTTTATAAAATAGCTTACGATTTACAACCAAACATGTTAGTGAATTCTAGAGTTGGAAAT
253 F Y K I A Y D L Q P N M L V N S R V G N

7441 GAATTAGGAGATTTTCGATATTCCAGGAGACAACAAAATCCCTCAAGATCATTAGCGATT
273 E L G D F D I P G D N K I P Q D H L A I

7501 **ACTAAACCTTGGCAAACCGTAGGTACCACAAACAATTCTTGGGGATACAATTCTTATGAT**
293 **T K P W Q T V G T T N N S W G Y N S Y D**

7561 **AACGATTGGAAATCTGTAAAAGAACTATTGTTCTGGTTAACCGAAATTGTAAGTAAAGGG**
313 **N D W K S V K E L L F W L T E I V S K G**

7621 **GGGAATTATATGTTGAATATAGGGCCAGATGCCACAGGTCATGTGCCTTCTGAATCGGTA**
333 **G N Y M L N I G P D A T G H V P S E S V**

7681 **AATAATCTATTAGAAGTTGGAAAGTGGTTGGAAGTTAATGGTG**
353 **N N L L E V G K W L E V N G**

RESUME EN FRANCAIS :

Ce travail visait à caractériser et produire deux biocatalyseurs originaux hydrolysant d'une part un polysaccharide pariétal d'algue rouge (le λ -carraghénane), et d'autre part un polysaccharide d'algue brune (le fucoïdane). Ces deux endo-hydrolases extracellulaires sont produites par deux bactéries marines saprophytes, *Pseudoalteromonas carrageenovora*, (γ -Protéobactéries), et SW5 (Bacteroidetes). Suite à la purification des enzymes sauvages, les gènes codant pour ces deux enzymes ont été clonés puis séquencés. L'activité recombinante obtenue par surexpression dans *E. coli* a permis de valider les séquences obtenues. L'analyse de ces dernières montre l'architecture modulaire des deux protéines. Cependant, le domaine catalytique de la λ -carraghénase n'a pas pu être identifié. Il s'agit donc d'un membre orphelin au sein de la classification des glycoside hydrolases, distinct des carraghénases préalablement caractérisées. A l'inverse, les similitudes partagées par la séquence de la fucoïdanase avec deux autres fucoïdanases putatives, a permis de définir une nouvelle famille de glycoside hydrolases. En outre, il est proposé que son gène fasse partie d'un cluster fonctionnel où figurent deux sulfatase, une estérase et une fucosidase.

MOTS-CLES :

macroalgues marines, carraghénanes, fucoïdanes, oligosaccharides sulfatés, bactéries, glycoside hydrolases, sulfatases, surexpression, analyse de séquence