



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National des Sciences Appliquées de Toulouse (INSA de Toulouse)

Présentée et soutenue par :

Romain IRAGUE

le mercredi 1 juin 2011

Titre :

Ingénierie des glucane-saccharases de la famille 70 des glycoside-hydrolases
pour la synthèse de nouveaux biopolymères

École doctorale et discipline ou spécialité :

ED SEVAB : Ingénieries microbienne et enzymatique

Unité de recherche :

LISBP (UMR CNRS 5504, UMR INRA 792), INSA de Toulouse

Directeur(s) de Thèse :

Pr. Magali Remaud-Siméon
Dr. Gabrielle Potocki-Véronèse

Jury :

Pr. Birte Svensson (Université technique du Danemark), rapporteur
Pr. Michel Dion (Université de Nantes), rapporteur
Dr. Alain Buléon (directeur de recherche, INRA de Nantes), examinateur
Dr. Gabrielle Potocki-Véronèse (chargée de recherche, LISBP, INSA de Toulouse)
Pr. Magali Remaud-Siméon (LISBP, INSA de Toulouse)
Pr. Pierre Monsan (LISBP, INSA de Toulouse), président de jury

NOM : IRAGUE

Prénom : Romain

Titre : Ingénierie des glucane-saccharases de la famille 70 des glycoside-hydrolases pour la synthèse de nouveaux biopolymères.

Spécialité: Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries, Filière : Ingénieries microbienne et enzymatique

Année : 2011

Lieu : INSA de Toulouse

RESUME :

Les glucane-saccharases sont des transglucosidases de la famille 70 des glycoside-hydrolases. A partir de saccharose, ces enzymes catalysent la synthèse d' α -glucanes, polymères de haute masse molaire formés d'unités glucosyle. Elles sont aussi capables de synthétiser des oligosaccharides ou glucoconjugués par réaction de transglucosylation sur des accepteurs exogènes de natures variées. De par la diversité de leurs spécificités, tant au niveau des liaisons osidiques synthétisées [$\alpha(1\rightarrow2)$; $\alpha(1\rightarrow3)$; $\alpha(1\rightarrow4)$ ou $\alpha(1\rightarrow6)$] que de l'organisation de ces liaisons au sein des produits formés, ces biocatalyseurs peuvent être mis à profit pour produire des hydrates de carbones d'intérêt pour les secteurs de l'alimentation, de la santé et de l'environnement. L'objectif de ces travaux de thèse était de générer par ingénierie enzymatique de nouvelles glucane-saccharases capables de synthétiser des α -glucanes et des gluco-oligosaccharides de structures et propriétés innovantes, afin d'élargir le panel d'applications de ces molécules. Sur le plan fondamental, l'enjeu était aussi d'améliorer la compréhension des relations entre structure et spécificité des glucane-saccharases.

Pour atteindre nos objectifs, nous avons utilisé une stratégie d'ingénierie combinatoire de la dextrane-saccharase DSR-S de *Leuconostoc mesenteroides* NRRL B-512F, qui catalyse la synthèse d'un dextrane formé à 95 % de liaisons $\alpha(1\rightarrow6)$ et 5 % $\alpha(1\rightarrow3)$. Le travail a tout d'abord consisté à développer une méthode de criblage multi-étapes et à haut-débit, pour isoler et trier les variants de spécificités de liaisons variées. La stratégie comprend une première étape de sélection *in vivo* sur milieu solide suivie d'un criblage par RMN 1D ^1H automatisé au format microplaque. Il s'agit de la première méthode de criblage à haut-débit de la spécificité de liaison des transglucosidases et glycosyltransferases. Cette stratégie a ensuite été appliquée au criblage de deux banques de variants de la DSR-S, totalisant plus de 36 000 clones obtenus par mutagenèse de saturation et recombinaison simultanée de 8 résidus du domaine catalytique. Au total, 82 mutants capables de synthétiser entre 2 et 8 fois plus de liaisons $\alpha(1\rightarrow3)$ par rapport à l'enzyme parentale ont été isolés. La caractérisation des propriétés catalytiques de sept mutants représentatifs de la diversité de spécificité générée a permis d'identifier un nouveau motif peptidique $^{460}\text{DYVHT}^{464}$ impliqué dans la spécificité de DSR-S. Enfin, la caractérisation de la structure et des propriétés rhéologiques et mécaniques des dextrans synthétisés par ces 7 mutants a mis en évidence les différences de taille et de conformation de ces macromolécules en solution et révélé l'aptitude du polymère synthétisé par le variant H463R/T464V/S512T à former des films bio-sourcés innovants, dont les propriétés mécaniques sont remarquables en comparaison de celles d'autres biopolymères extraits de plantes ou produits par fermentation.

MOTS CLES :

Glucane-saccharase, spécificité de liaison, ingénierie combinatoire, criblage RMN à haut-débit, α -glucanes, propriétés physico-chimiques, biofilms.

Ecole Doctorale : SEVAB (Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries)

Laboratoire : Ingénierie des Systèmes Biologiques et des Procédés (UMR CNRS 5504, UMR INRA 792) de l'INSA de Toulouse

Last name : IRAGUE

First name : Romain

Title : Engineering glucansucrases from family 70 of glycoside-hydrolases for the synthesis of novel polysaccharides

Speciality: Ecological, Veterinary, Agronomic Sciences and Bioengineering, Filière : Enzymatic and microbial engineering

Year : 2011

Place : INSA, Toulouse

SUMMARY :

Glucansucrases are transglucosidases from glycoside hydrolase family 70. From sucrose, these enzymes catalyse the synthesis of α -glucans, high molecular weight polymers formed of glucosyl units. Glucansucrases are also able to synthesize oligosaccharides or glucoconjugates by transglucosylation reaction onto various exogenous acceptors. Due to the large specificity, in terms of osidic linkages synthesized [$\alpha(1\rightarrow2)$; $\alpha(1\rightarrow3)$; $\alpha(1\rightarrow4)$ or $\alpha(1\rightarrow6)$] and their organization within the formed products, these biocatalysts can be used to produce carbohydrates of interest in food, health and environment fields. The aim of this research work was to generate, by enzyme engineering, new glucansucrases able to synthesize α -glucans and gluco-oligosaccharides with novel structures and properties, to enlarge applications of these molecules. At fundamental level, the work was to improve the understanding of the relationships between structure and specificity of glucansucrases. To reach our objectives, we used a combinatorial engineering strategy on the dextransucrase DSR-S of *Leuconostoc mesenteroides* NRRL B-512F, which catalyses the synthesis of a dextran formed by 95 % of $\alpha(1\rightarrow6)$ and 5 % $\alpha(1\rightarrow3)$ linkages. The work consisted first in the development of a multi-step high throughput screening methodology to sort out and isolate variants with altered linkage specificities. The strategy includes a first stage of *in vivo* selection on solid medium followed by an automated 1D ^1H NMR-based screening using microplates. To date, this is the first high-throughput screening method for linkage specificity determination of transglucosidases and glycosyltransferases. This strategy was applied to the screening of two DSR-S variants libraries, totalizing more than 36,000 clones obtained by saturation mutagenesis and simultaneous recombination of 8 residues from the catalytic domain. Eighty two mutants able of synthesize 2 to 8 times more $\alpha(1\rightarrow3)$ linkages compared to the parental enzyme were isolated. The characterization of the catalytic properties of 7 representative mutants enabled the identification of a new peptide motif $^{460}\text{DYVHT}^{464}$, involved in DSR-S specificity. Finally, the characterization of structural, rheological and mechanical properties of dextrans synthesized by these 7 mutants highlighted the differences in size and conformation of these macromolecules in solution and revealed the capacity of the polymer synthesized by H463R/T464V/S512T variant to form biofilms, whose mechanical properties are remarkable in comparison to those of other biopolymers extracted from plants or produced by fermentation.

Keywords :

Glucansucrase, linkage specificity, combinatorial engineering, high throughput NMR-based screening, α -glucans, physico-chemical properties, biofilms.

Doctorate school : SEVAB (Ecological, Veterinary and Agronomic Sciences and Bioengineering)

Laboratory : Laboratory of Biosystems and Chemical Engineering (UMR CNRS 5504, UMR INRA 792) de l'INSA de Toulouse

PUBLICATIONS

NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity.

Iraque, R., Massou, S., Moulis, C., Saurel, O., Milon, A., Monsan, P., Remaud-Siméon, M., Portais, J.C., Potocki-Véronèse, G. 2011. *Analytical Chemistry*. **83**, 1202-1206.

Combinatorial engineering of dextransucrase specificity.

Iraque, R., Roncalli-Tarquis, L., André, I., Moulis, C., Morel, S., Monsan, P., Potocki-Véronèse, G., Remaud-Siméon, M. 2011. *En préparation*.

Characterization of bio-engineered α -D-glucans synthesized by dextransucrase mutants

Iraque, R., Rolland-Sabaté, A., Roncalli-Tarquis, L., Doublier, J.L., Moulis, C., Monsan, P., Remaud-Siméon, M., Potocki-Véronèse, G., Buléon, M. 2011. *En préparation*.

COMMUNICATIONS ORALES

Ingénierie des glucane-saccharases de la famille 70 des glycoside-hydrolases pour la synthèse de nouveaux biopolymères.

Iraque R., Lima J., Moulis C., Brison Y., Emond S., Massou S., André I., Monsan P., Rémaud-Siméon M., Potocki-Véronèse G.

Colloque du GDR 2825 « Evolution dirigée des protéines », octobre 2009, Aigues-mortes, France.

Ingénierie combinatoire des glucane-saccharases de la famille 70 des glycoside-hydrolases et synthèse de nouveaux biopolymères.

Iraque R., Moulis C., Massou S., André I., Monsan P., Rémaud-Siméon M., Potocki-Véronèse G.

Séminaire jeunes chercheurs CEPIA, août 2010, Le Hohwald, France.

COMMUNICATIONS PAR AFFICHES

Directed evolution to enlarge glucansucrase potentialities

Emond S., Iraque R., Brison Y., André I., Potacki-Véronèse G., Monsan P., Remaud-Siméon M.

Workshop "Directed evolution approaches in structural biology", EMBL Grenoble Outstation, Janvier 2008, France.

Glucansucrase molecular engineering for novel oligosaccharide and polysaccharide synthesis.

Brison Y., Iraque R., Moulis C., Champion E., Emond S., Pizzut-Serin S., André I., Potocki-Véronèse G., Morel S., Monsan P., Rémaud-Siméon M.

Biopolymers, Biomaterials, Green Chemistry Workshop, décembre 2008, La Baule, France.

Engineering glucansucrases from family 70 of glycoside-hydrolases for the synthesis of novel polysaccharides

Iraque R., Lima J., Moulis C., Brison Y., Emond S., Massou S., André I., Monsan P., Rémaud-Siméon M., Potocki-Véronèse G.

Congrès annuel de la SFBBM, août 2009, Nancy, France.

Ingénierie des glucane-saccharases de la famille 70 des glycoside-hydrolases pour la synthèse de nouveaux biopolymères

Iraque R., Lima J., Moulis C., Brison Y., Emond S., Massou S., André I., Monsan P., Rémaud-Siméon M., Potocki-Véronèse G.

Journée de l'école doctorale SEVAB, novembre 2009, Auzeville. France.

Glucansucrase engineering to produce functional biopolymers and oligosaccharides

Iraque R., Moulis C., Emond S., Brison Y., André I., Monsan P., Rémaud-Siméon M., Potocki-Véronèse G.

Fourth Symposium on Biocatalysis for the Food and Drink Industries, avril 2010, Toulouse, France.

NMR-based high-throughput screening of carbohydrate-active enzyme specificity

Iraque R., Massou S., Moulis C., Saurel O., Milon A., Monsan P., Remaud-Siméon M., Portais J.C., Potocki-Véronèse G.

9th Carbohydrate Bioengineering Meeting, mai 2011, Lisboa, Portugal.

MERCI

Je voudrais remercier en premier lieu mes deux encadrantes : Magali et Gabrielle. Merci de votre disponibilité, votre gentillesse, votre dynamisme et la confiance que vous m'avez accordé. Chacune à votre manière vous m'avez énormément apporté tant au niveau professionnel qu'au niveau personnel. Je voulais vous dire un grand MERCI pour la patience dont vous avez su faire preuve, j'ai conscience que ça n'a pas toujours été facile. Je voudrais également vous dire que se fut un réel plaisir, que dis-je, un bonheur pour moi de travailler sous votre direction.

Je voudrais également remercier l'ensemble des membres du jury pour avoir accepté de juger mon travail de thèse. Merci Pierre de m'avoir ouvert les portes du laboratoire et de m'avoir fait l'honneur de présider mon jury de thèse. Un très très grand MERCI, Alain, pour votre accueil lors de mes deux séjours dans le grand nord. Merci, Birte Svensson et Michel Dion, pour avoir accepté d'évaluer ce manuscrit, pour l'intérêt que vous y avez porté ainsi que pour vos conseils avisés lors de la soutenance.

Les résultats obtenus au cours de ces (presque) quatre années de thèse sont le fruit de diverses collaborations qui ont été pour moi extrêmement enrichissantes. Je tenais donc à remercier Stéphane Massou. Merci beaucoup Stéphane pour ta gentillesse, ton aide plus que précieuse, ton énorme investissement dans la mise au point du criblage RMN. Sans toi, tes idées et ton professionnalisme, je serais sûrement encore en train de passer des tubes au spectro pour cribler mes banques (Ouffffffff). Merci également pour tes remarques et tes conseils. Un grand MERCI Jean Charles Portais pour votre aide et vos remarques lors de la rédaction de mon premier papier. Je remercie également Olivier Saurel et Alain Millon pour m'avoir permis de toucher à la « Flow-NMR », ainsi que pour la relecture du papier. Et pardon à toi « préparateur d'échantillon » qui semble ne pas avoir apprécié les 4000 variants criblés !!!!

Merci à toutes les personnes rencontrées à Nantes. Merci beaucoup Agnès pour ton investissement et ton aide dans la caractérisation des polymères par AFFFF-MALLS. Merci également à Jean-louis Doublier pour m'avoir initié à la rhéologie des polymères, pour votre aide et remarques avisées. Merci à Marion, Laurent, Denis, Jane, Georges et tous ceux que j'ai croisé lors de mes périples Nantais. Alain, je n'oublierai jamais ce superbe week-end passé dans le golfe du Morbihan, la balade en vélo et surtout en voilier, cap vers l'île d'Houat.

Bien évidemment une bonne thèse ne peut se faire sans un environnement adéquat. C'est pourquoi je tenais à remercier l'ensemble des membres passés et présent de l'EAD1, stagiaires, doctorants, post-doctorants et permanents. Merci à Stéphane, Vincent, Elise, Yoann, David, Isabelle, Sandrine, Claire M, Laurence, Nelly, Peco, Florence, Alain, Michael, le gang des Sophie, Nicolas, Cédric, Claire D, Florent, Juliette, Martial, Yannick, Etienne, Olivier, Letian, Mawine, Sandra, Beau, Tuck, Faten, Fred, Manue, Régis, Elisabeth, Javier, Ann, Mélusine, Nathalie C, Hélène, Géraldine, Marie-laure, Marlène, Virginie, ainsi que les petits nouveaux et les petits moins nouveaux que je n'ai pas cité.

Un grand MERCI général à l'ensemble du LISBP.

Merci à Joana et Marta (The Portuguese Connection) pour leurs contributions à ces travaux.

Je remercie ma famille, mes parents, mon frère, les « Cousines » et Delphine pour leur soutien.

Enfin, je tiens à remercier mes amis : Gillou et Elodie, Kikou, Le SarrazinCrew, YoYo, Pat(ibulaire) et Amcha. Merci pour tous les moments de vie en dehors de la thèse. Une pensée à Gillou et Pat(ibulaire) pour les différents concerts de MétAAAAAAL partagés avec vous dans les salles obscures toulousaines ou bien au HELLFEST. Je remercie l'ensemble des assos toulousaines qui se bougent pour faire vivre la scène Métal toulousaine, française et internationale. Et puis merci à tous les groupes qui m'ont accompagné tout au long de la thèse : Between the Buried and Me, Cynic, Eryn Non Dae, Genghis Thron, Gojira, Hacride, Intronaut, Mastodon, Psykup, The dillinger escape plan, The ocean, Gorod, Neurosis, Magrudergrind, Moonspell, Napalm death, Nile, Obscura, Opeth, Trepalium, Ulcerate, Zubrowska, DHG, Manimal, Amenra, Celeste, Cellscape, Cult of luna, Dirge, Knut, Unexpected, Septic flesh, Giant squid, Nesseria, I pilot deamon, Nolentia, Time to burn, Hollow corp., Cephalic carnage, Kruger, Kylesa, Made out of baby, Plebeian grandstand, Revok, Rwake, Selenites, Diablo swing orchestra, Ephel duath, Fantomas, Deftones, Comity, Flying pooh, System of a down, Bumblefoot, The mars volta, Death in vegas, Ez3kiel, Igorrr, Le peuple de l'herbe, Massive attack, Portishead, Radius system...

Sommaire

INTRODUCTION	18
CHAPITRE 1 : Introduction bibliographique	22
Partie 1 : Les glucane-saccharases	24
I. LES GLUCANE-SACCHARASES	24
I.1 Classification des glucane-saccharases	26
I.2 Réactions catalysées et diversité des produits synthétisés par les glucane-saccharases	27
I.3 Relations structure/fonction des glucane-saccharases	31
I.3.1 Organisation structurale des glucane-saccharases de la famille GH13	31
I.3.1.A Repliement global	31
I.3.1.B Architecture du site actif et fixation des substrats glucosidiques	31
I.3.2 Mécanisme catalytique des glucane-saccharases de la famille GH13	33
I.3.3 Organisation structurale des glucane-saccharases de la famille GH70	35
I.3.3.A Structure globale et fonction	35
I.3.3.B Sous-sites de fixation au niveau du site actif	38
I.3.4 Mécanisme catalytique des glucane-saccharases de la famille GH70	41
I.3.5 Mécanisme de polymérisation des glucane-saccharases de la famille GH70	41
I.3.6 Mécanisme de réaction d'accepteur des glucane-saccharases de la famille GH70	42
I.3.7 Déterminants de la spécificité des glucane-saccharases de la famille GH70	42
Partie 2 : Ingénierie d'enzymes	50
II. INGENIERIE D'ENZYMES	50
II.1 Ingénierie par évolution dirigée : vue générale	52
II.1.1 Principe	52
II.1.2 Générer la diversité par mutagenèse aléatoire	53
II.1.3 Générer la diversité par recombinaison aléatoire	54
II.1.4 Rationaliser l'aléatoire : mutagenèse et recombinaison semi-aléatoire	56
II.1.4.A Mutagenèse de saturation	56
II.1.4.B Mutagenèse de saturation couplée à la recombinaison	57
II.1.5 Complémentarité des approches rationnelles et aléatoires	59
II.2. Trier la diversité : Sélection et criblage	59
II.2.1 Techniques de sélection <i>in vivo</i>	60
II.2.2 Techniques de sélection <i>in vitro</i>	61
II.2.2.A La technologie du « display »	61
II.2.2.B La compartimentation <i>in vitro</i>	63
II.2.3 Criblage à haut-débit	65
II.3 Ingénierie de la spécificité de transglycosidases vis-à-vis de leurs substrats et produits	67
Partie 3 : α -glucanes produits à partir de saccharose	74
III.1 Caractéristiques structurales	76
III.1.1 Structure des α -glucanes	76
III.1.2 Tailles des α -glucanes	80
III.2 Propriétés et caractéristiques physico-chimiques	84
III.2.1 Propriétés rhéologiques	84
III.2.2 Propriété thermique : <i>Transition vitreuse</i>	87
III.3 Applications des α -glucanes	89
III.3.1 Applications des dextrans	89

III.3.2 Applications de l'alternane	92
III.3.3 Applications des mutanes	92
III.3.4 Applications des amylopolysaccharides	93
Objectifs de la thèse	96
Références bibliographiques	98
CHAPITRE 2 : Développement d'une stratégie multi-étapes de sélection et criblage haut-débit de la spécificité des glucane-saccharases	120
Abstract	124
Results and discussion	125
Materials and Methods	133
Acknowledgment	136
References	138
CHAPITRE 3 : Ingénierie combinatoire de la spécificité de la dextrane-saccharase DSR-S	142
Abstract	146
Introduction	147
Results and discussion	149
Selection of amino acid targets	149
Analysis of the initial library	150
Selection and screening of active glucansucrase variants	152
Variants with altered linkage specificity	152
Functional and kinetic characterization of the selected variants	153
3D- modelling of DSR-S catalytic domain	155
Catalytic site description and docking of substrate and putative products	156
Conclusions	159
Materials and methods	160
Acknowledgments	166
Supporting Information	167
Tables section	167
Figure section	170
References	172
CHAPITRE 4 : Caractérisation physico-chimique d' α -glucanes produits par des mutants de dextrane-saccharase	176
Abstract	180
Introduction	181
Materials and methods	182
Results and discussion	186
α -glucan size and conformation in solution	186
Flow behaviour	190
Viscoelastic behaviour	191
Moisture contents and glass transition	192
Mechanical behaviour of the films	194

Conclusion	197
Acknowledgments	197
References	198
CONCLUSION.....	204
Développement d'une stratégie multi-étape pour le criblage de la spécificité des glucane- saccharases	207
Ingénierie des glucane-saccharases de la famille GH70 par family shuffling	209
Ingénierie semi-rationnelle de la dextrane-saccharase DSR-S	210
Caractérisation des propriétés physico-chimiques des nouveaux dextrans.....	212
Tables des illustrations	215
FIGURES	216
TABLEAUX	222

INTRODUCTION

Ce projet de thèse, financé par les départements CEPIA de l'INRA et ST2I du CNRS, s'inscrit dans le cadre du programme interdisciplinaire 'Chimie Pour le Développement Durable', et en particulier du Réseau de Recherche associé 'Intégrer une approche sociétale renouvelée de la chimie: les ressources renouvelables comme matières premières et sources de nouveaux produits et matériaux'. Ce réseau scientifique vise à stimuler le développement de procédés éco-compatibles de conversion de la biomasse végétale en synthons, nouveaux matériaux non persistants, et autres molécules à haute valeur ajoutée pour les secteurs de l'alimentaire, de la santé, de la chimie fine et de l'environnement.

Dans ce contexte, les glucane-saccharases d'origines bactériennes étudiées par l'équipe de Catalyse et Ingénierie Moléculaire EnzymatiqueS (CIMES) du Laboratoire d'Ingénierie des Systèmes et des Procédés (LISBP) de l'INSA de Toulouse sont particulièrement attractives. Ces transglucosidases (E.C.2.4.1.4, E.C.2.4.1.5 et 2.4.1.140) classées dans les familles 13 et 70 des glycoside-hydrolases produisent à partir de saccharose, une agro-ressource peu coûteuse, des α -glucanes de haute masse molaire. Elles peuvent également synthétiser des oligosaccharides ou glucoconjugués par réaction de transglucosylation sur des accepteurs exogènes de natures variées. De plus, elles présentent des spécificités de produits différentes, tant au niveau des liaisons osidiques synthétisées [$\alpha(1\rightarrow2)$; $\alpha(1\rightarrow3)$; $\alpha(1\rightarrow4)$ ou $\alpha(1\rightarrow6)$] que de l'organisation de ces liaisons au sein des produits formés. Cette grande diversité fonctionnelle peut être mise à profit pour produire diverses molécules à moyenne et haute valeur ajoutée de type polysaccharides, oligosaccharides ou glucoconjugués de principes actifs.

Pourtant, parmi les α -glucanes synthétisés par les glucane-saccharases, seuls les amylopolysaccharides produits par l'amylosaccharase de *Neisseria polysaccharea* et les dextrans produits par l'enzyme DSR-S de la souche *Leuconostoc mesenteroides* NRRL B-512F ont été caractérisés sur le plan de leurs propriétés physico-chimiques, ce qui explique le champ d'applications encore limité de ces polymères et des enzymes qui les synthétisent. Les dextrans de *Ln. mesenteroides* NRRL B-512F sont en effet les uniques polymères synthétisés par une glucane-saccharase produits à l'échelle industrielle. Ils trouvent de nombreuses applications dans les secteurs pharmaceutiques (en tant que substituts de plasma sanguin et anticoagulants), de la chimie analytique (en tant que supports chromatographiques) et de l'agroalimentaire (comme texturants pour la panification).

Lors de cette thèse, nous avons donc entrepris l'ingénierie combinatoire des glucane-saccharases de la famille 70 des glycoside-hydrolases pour générer de nouvelles enzymes de synthèse d' α -glucanes et

de gluco-oligosaccharides de structures et de propriétés originales. L'enjeu était double. Il s'agissait d'une part d'identifier les acides aminés impliqués dans la spécificité des glucane-saccharases. D'autre part, les travaux visaient à élargir le panel d' α -glucanes disponibles, et de caractériser leurs propriétés physico-chimiques afin d'évaluer leur potentiel applicatif, notamment en tant que biomatériaux.

Le premier chapitre de ce manuscrit est consacré à une étude bibliographique permettant de situer le contexte scientifique de l'étude. Il est divisé en deux parties dédiées respectivement à l'état de l'art sur les glucane-saccharases et l'ingénierie d'enzymes et se termine par un dernier volet consacré à la structure et aux propriétés des α -glucanes synthétisés à partir de saccharose.

Les résultats obtenus sont ensuite présentés sous la forme de trois articles dans les chapitres II, III et IV. Le chapitre II présente le développement d'une stratégie de criblage de la spécificité de liaison des glucane-saccharases. Le chapitre III est dédié à l'application de cette stratégie à l'ingénierie de la dextrane-saccharase DSR-S, afin d'isoler des mutants de spécificités modifiées et d'identifier de nouveaux résidus impliqués dans le contrôle de la spécificité de la dextrane-saccharase de *Ln. mesenteroides* NRRL B-512F. Enfin, le chapitre IV est consacré à la caractérisation des propriétés physico-chimiques des α -glucanes synthétisés par sept de ces variants enzymatiques.

CHAPITRE 1 : Introduction bibliographique

Partie 1 : Les glucane-saccharases

I. LES GLUCANE-SACCHARASES

Les glucane-saccharases sont des enzymes bactériennes qui catalysent la synthèse de glucanes (homopolymères d'unités α -D-glucopyranosyle) à partir d'un substrat unique, le saccharose. Ces enzymes sont également capables de catalyser le transfert de l'unité D-glucopyranosyle du saccharose vers un accepteur exogène. Cette réaction, nommée réaction d'accepteur, aboutit à la synthèse d'oligosaccharides (quand l'accepteur est un monosaccharide ou lui-même un oligosaccharide) ou de glucoconjugués (quand l'accepteur n'est pas de nature osidique). Contrairement aux glycosyltransférases de type Leloir qui utilisent des sucres activés pour catalyser la synthèse de liaisons osidiques, les glucane-saccharases ne requièrent que l'énergie du clivage de la liaison osidique du saccharose ($\sim 27,6 \text{ kJ.mol}^{-1}$) (André et al., 2010). Elles présentent une étonnante diversité de spécificités, tant au niveau du type de liaisons osidiques formées [$\alpha(1\rightarrow2)$; $\alpha(1\rightarrow3)$; $\alpha(1\rightarrow4)$ ou $\alpha(1\rightarrow6)$] que de l'organisation de ces liaisons au sein du polysaccharide synthétisé. Ainsi, elles sont capables de synthétiser une large gamme de biopolymères, de tailles et de structures différentes. Ces enzymes sont donc des outils à fort potentiel pour la bioconversion et la valorisation du saccharose, une agro-ressource abondante et peu coûteuse (155 millions de tonnes pour l'année 2010 pour un prix de 437 dollars la tonne, <http://agritrade.cta.int/index.php/fr>).

I.1 Classification des glucane-saccharases

Les glucane-saccharases sont des transglucosidases classées, selon l'Enzyme Commission, dans la famille des glycosyltransférases (E.C.2.4.x.y). Cette classification est fondée sur le type de réactions catalysées et sur la spécificité de substrats mais ne tient pas toujours compte du mécanisme d'action catalytique notamment au niveau moléculaire. En outre, elle ne fait pas référence aux liens de parenté structuraux ou mécanistiques pouvant exister entre enzymes. La classification CAZy (URL : <http://www.cazy.org/>), proposée par Henrissat *et al.*, répartit quant à elle les modules catalytiques des enzymes actives sur les sucres et les modules de liaison aux glucides selon leurs homologies de séquences (Cantarel *et al.*, 2009).

Ainsi, les enzymes actives sur les sucres sont réparties en cinq catégories dans la classification Cazy:

- Les **glycoside-hydrolases** (GHs), regroupant les enzymes capables de catalyser des réactions d'hydrolyse ou de transglycosylation (Davies and Henrissat, 1995; Henrissat, 1991; Henrissat and Bairoch, 1993, 1996; Henrissat and Davies, 1997).

- Les **glycosyltransferases** (GTs), regroupant les enzymes utilisant des sucres activés pour catalyser le transfert d'unités glucosyle sur des accepteurs (Campbell *et al.*, 1997; Coutinho *et al.*, 2003).
- Les **polysaccharide-lyases** (PLs), catalysant le clivage non-hydrolytique de liaisons glycosidiques (Coutinho and Henrissat, 1999).
- Les **carbohydrate-esterases** (CEs), catalysant les O- ou N-acétylations de sucres substitués (Coutinho and Henrissat, 1999)
- Les **carbohydrate-binding modules** (CBMs), modules de liaison aux glucides, le plus souvent associés à un ou plusieurs modules catalytiques (Boraston *et al.*, 2004).

Selon la classification CAZy, la plupart des glucane-saccharases appartiennent à la famille 70 des glycoside-hydrolases (abrégée GH70). Cette famille comprend des enzymes de masse moléculaire élevée variant de 130 à 313 kDa produites par des bactéries lactiques GRAM positif appartenant aux genres *Leuconostoc*, *Lactobacillus*, *Streptococcus*, *Weisella* et *Exiguobacterium* (André *et al.*, 2010; Van Hijum *et al.*, 2006).

Seules les amylosaccharases, qui catalysent la synthèse d' α -glucanes composés de liaisons $\alpha(1\rightarrow4)$ à partir de saccharose, sont classées dans la famille GH13 (famille principalement composée d'enzymes de modification de l'amidon). Avec la famille GH77 regroupant des 4- α -glucanotransferases, les enzymes des familles GH13 et GH70 constituent le clan GH-H, également connu sous le nom de « superfamille des α -amylases ». Les enzymes du clan GH-H partagent une organisation structurale en tonneau catalytique de type $(\beta/\alpha)_8$, un mécanisme de rétention de la configuration anomérique sous forme α et une machinerie catalytique impliquant une diade catalytique Asp-Glu.

1.2 Réactions catalysées et diversité des produits synthétisés par les glucane-saccharases

Les glucane-saccharases des familles GH13 et GH70 catalysent le transfert d'unités glucosyle provenant du clivage du saccharose, vers une molécule acceptrice présente dans le milieu réactionnel. La première étape de cette réaction est la formation d'un complexe covalent β -D-glucosyl-enzyme (Jensen *et al.*, 2004; Mooser and Iwaoka, 1989), accompagnée de la libération concomitante d'une molécule de fructose (Figure 1). Dans un second temps, le résidu glucosyle est transféré sur un accepteur, la nature du produit formé dépendant de la destination de l'unité glucosyle.

Dans le cas où seul le saccharose est présent dans le milieu, le produit majoritairement obtenu est un α -glucane de haut poids moléculaire résultant de la condensation des unités glucosyle (réaction de

polymérisation). L'eau et le fructose libéré peuvent aussi jouer le rôle de molécules acceptrices ce qui aboutit, respectivement, à l'hydrolyse et à la synthèse de saccharose (échange isotopique) et/ou d'isomères de saccharose (le leucrose étant le plus souvent décrit).

Lorsqu'un accepteur exogène présentant un ou plusieurs groupes hydroxyle est ajouté dans le milieu réactionnel, la réaction de synthèse de polymère peut être détournée si l'accepteur est reconnu par l'enzyme. Il en résulte une production plus ou moins abondante de glucooligosaccharides (si l'accepteur est de nature osidique) ou de glucocojugués (si l'accepteur n'est pas un glucide) (Monsan *et al.*, 2010). La spécificité de liaison des glucane-saccharases est conservée entre la réaction de polymérisation et la synthèse d'oligosaccharides par réaction d'accepteur (Remaud-Simeon *et al.*, 2000).

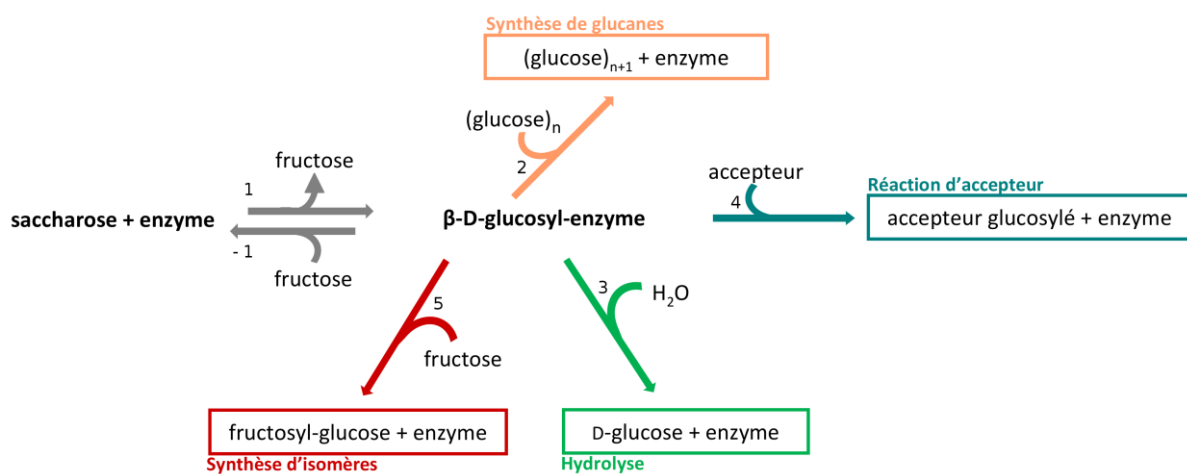


Figure 1 : Réactions catalysées par les glucane-saccharases. +1 : formation du complexe β -D-glucosyl-enzyme ; -1 : échange isotopique ; 2 : réaction de polymérisation ; 3 : hydrolyse ; 4 : réaction d'accepteur ; 5 : formation d'isomères de saccharose.

Les glucane-saccharases présentent des spécificités variées et catalysent la formation de polysaccharides variant par la nature, le nombre et l'organisation des liaisons osidiques selon l'enzyme considérée. On distingue ainsi :

× Les **dextrane-saccharases** (GH70 ; E.C.2.4.1.5). Produites principalement par des bactéries du genre *Leuconostoc* et, dans une moindre mesure, par des *Streptocoques* et *Lactobacilles*, elles synthétisent des α -glucanes, communément appelés dextranes (figure 2), dont la chaîne principale contient plus de 50 % de liaisons de type $\alpha(1\rightarrow6)$. Des ramifications de type $\alpha(1\rightarrow2)$, $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$ sont également présentes à des taux qui fluctuent selon l'enzyme considérée.

× Les **mutane-saccharases** (GH70 ; E.C.2.4.1.5). Produites par des bactéries du genre *Streptococcus*, elles synthétisent des mutanes (figure 2), glucanes constitués majoritairement de liaisons $\alpha(1\rightarrow3)$ consécutives dans la chaîne principale.

× Les **reuterane-saccharases** (GH70 ; E.C. 2.4.1.5) de *Lactobacillus* spp.. Elles produisent un polymère appelé reuterane (figure 2), formé à plus de 50 % d'unités glucosyle liées en $\alpha(1\rightarrow4)$. On retrouve également dans ce polymère des liaisons de type $\alpha(1\rightarrow6)$.

× Les **alternane-saccharases** (GH70 ; E.C. 2.4.1.140). Elles synthétisent un glucane atypique formé de liaisons $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow6)$ alternées dans la chaîne principale, appelé alternane (figure 2). On les retrouve chez 3 souches de *Leuconostoc citreum*, à savoir NRRL-1355, NRRL B-1498 et NRRL B-1501.

× Les **amylosaccharases** (GH13 ; 2.4.1.4). Elles synthétisent un polymère linéaire de type amylose, composé de liaisons $\alpha(1\rightarrow4)$. Les bactéries productrices sont du genre *Neisseria*, *Deinococcus* et *Alteromonas*.

A ce jour, 112 glucane-saccharases de la famille GH70 et 7 glucane-saccharases de la famille GH13 sont recensées dans la base de données CAZy. Quarante et une d'entre elles ont été caractérisées sous forme recombinante. Elles sont listées dans le tableau 1.

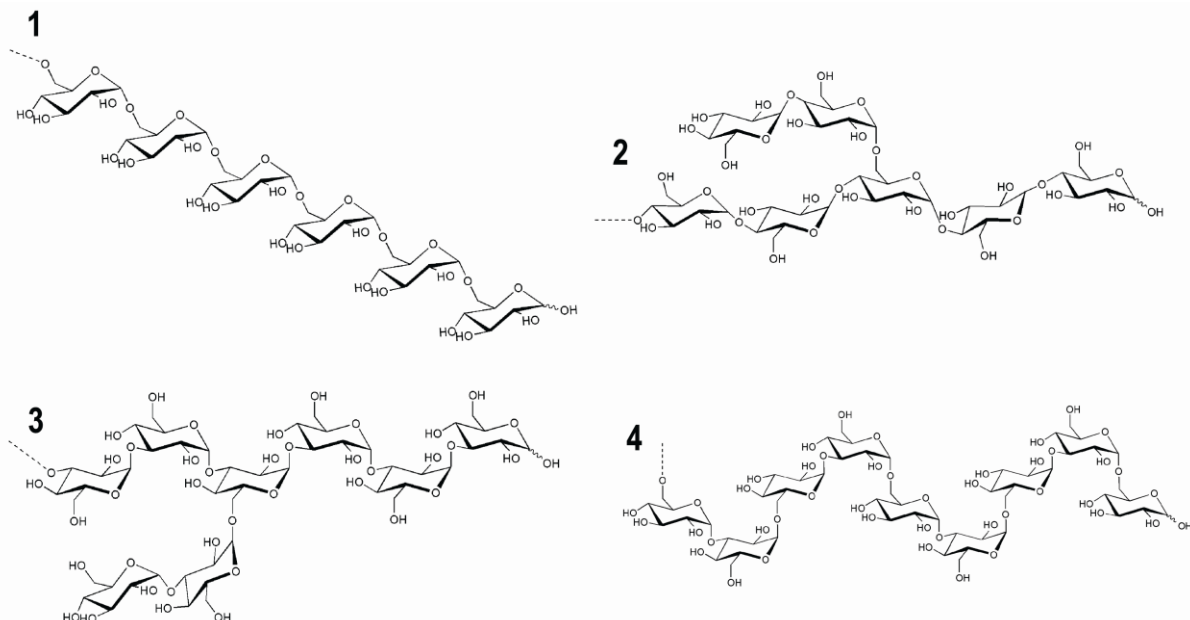


Figure 2 : Représentation schématique des différentes structures des glucanes produits par les glucane-saccharases GH70. 1 : dextrane (résidus de la chaîne principale liés majoritairement en $\alpha(1\rightarrow6)$) ; 2 : reuterane (résidus de la chaîne principale liés en $\alpha(1\rightarrow4)$ et $\alpha(1\rightarrow6)$) avec des ramifications en $\alpha(1\rightarrow6)$) ; 3 : mutane (résidus de la chaîne principale liés majoritairement en $\alpha(1\rightarrow3)$) et 4 : alternane (résidus de la chaîne principale liés alternativement en $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow6)$).

Tableau 1 : Liste des glucane-saccharases caractérisées en janvier 2011.

Organisme	Enzyme	Taille (aa)	Spécificité de liaison osidique				Référence
			$\alpha(1\rightarrow6)$	$\alpha(1\rightarrow4)$	$\alpha(1\rightarrow3)$	$\alpha(1\rightarrow2)$	
GH family 13							
<i>Neisseria polysaccharea</i>	Amylosaccharase	636		100			(Potocki De Montalk <i>et al.</i> , 1999)
<i>Deinococcus radiodurans</i>	Amylosaccharase	644		100			(Pizzut-Serin <i>et al.</i> , 2005)
<i>Deinococcus geothermalis</i>	Amylosaccharase	650		100			(Emond <i>et al.</i> , 2008b)
<i>Alteromonas addita</i>	Amylosaccharase	649		100			(Ha <i>et al.</i> , 2009)
<i>Alteromonas macleodii</i>	Amylosaccharase	649		100			(Ha <i>et al.</i> , 2009)
GH family 70							
Dextrane-saccharases							
<i>Leuconostoc mesenteroides</i> B-512F	DSR-S	1527	95		5		(Monchois <i>et al.</i> , 1997)
<i>Leuconostoc mesenteroides</i> B-512F	DSR-T	1015			nd*		(Funane <i>et al.</i> , 2000)
<i>Leuconostoc mesenteroides</i> B-512F	DSR-T5	1499	60		40		(Funane <i>et al.</i> , 2000)
<i>Leuconostoc mesenteroides</i> B-1299CB4	DSR-BCB4	1465	100				(Kang <i>et al.</i> , 2008)
<i>Leuconostoc citreum</i> B-1299	DSR-E	2835	68		3	29	(Dols <i>et al.</i> , 1997a)
<i>Leuconostoc citreum</i> B-1299	DSR-A	1290	85		15		(Monchois <i>et al.</i> , 1996a)
<i>Leuconostoc citreum</i> B-1299	DSR-B	1508	95				(Monchois <i>et al.</i> , 1996b)
<i>Leuconostoc citreum</i> B-1355	DSR-C	1477	96				(Argüello-Morales <i>et al.</i> , 2000)
<i>Leuconostoc mesenteroides</i> Lcc4	DSR-D	1527	Maj.*				(Neubauer <i>et al.</i> , 2003)
<i>Leuconostoc mesenteroides</i> IBT-QP	DSR-P	1454			nd		(Olvera <i>et al.</i> , 2007)
<i>Leuconostoc mesenteroides</i> L039	DSR-X	1485			nd		(Yalin <i>et al.</i> , 2008)
<i>Leuconostoc mesenteroides</i> 0326	DexYG	1527			nd		(Zhang <i>et al.</i> , 2008)
<i>Leuconostoc citreum</i> HJP4	LcDS	1477			nd		(Yi <i>et al.</i> , 2009)
<i>Streptococcus mutans</i> GS 5	GTF-D	1430	70		30		(Shimamura <i>et al.</i> , 1994)
<i>Streptococcus oralis</i> ATCC10557	GTF-R	1575	86				(Fujiwara <i>et al.</i> , 2000)
<i>Streptococcus downei</i> Mfe 28	GTF-S	1365	90				(Gilmore <i>et al.</i> , 1990)
<i>Streptococcus sobrinus</i> OMZ176	GTF-T	1468	79		21		(Hanada <i>et al.</i> , 1993)
<i>Streptococcus sobrinus</i> B13N	GTF-U	1554	69		31		(Hanada <i>et al.</i> , 2002)
<i>Streptococcus salivarius</i> ATCC25975	GTF-K	1599	100				(Simpson <i>et al.</i> , 1995)
<i>Streptococcus salivarius</i> ATCC25975	GTF-M	1577	95				(Simpson <i>et al.</i> , 1995)
<i>Streptococcus gordonii</i> CH1	GTF-G	1577	75		23		(Vickerman <i>et al.</i> , 1997)
<i>Lactobacillus reuteri</i> 180	GTF-180	1772	61		39		(Kralj <i>et al.</i> , 2004a)
<i>Lactobacillus sakei</i> Kg15	GTF-Kg15	1561	90		10		(Kralj <i>et al.</i> , 2004a)
<i>Lactobacillus fermentum</i> Kg3	GTF-Kg3	1595	92		8		(Kralj <i>et al.</i> , 2004a)
<i>Lactobacillus parabuchneri</i> 33	GTF-33	1463	81		19		(Kralj <i>et al.</i> , 2004a)
<i>Weissella cibaria</i>	rDSRWC	1472	Maj.				(Kang <i>et al.</i> , 2009)
Mutane-saccharases							
<i>Streptococcus salivarius</i> ATCC25975	GTF-L	1449	50		50		(Simpson <i>et al.</i> , 1995)
<i>Streptococcus salivarius</i> ATCC25975	GTF-J	1517			90		(Simpson <i>et al.</i> , 1995)
<i>Streptococcus mutans</i> GS 5	GTF-B	1475	12		88		(Shiroza <i>et al.</i> , 1987)
<i>Streptococcus mutans</i> GS 5	GTF-C	1375	15		86		(Ueda <i>et al.</i> , 1988)
<i>Streptococcus downei</i> Mfe 28	GTF-I	1597	12		88		(Russell <i>et al.</i> , 1987)
<i>Streptococcus sobrinus</i> 6715	GTF-la	1592			Maj.		(Konishi <i>et al.</i> , 1999)
<i>Lactobacillus reuteri</i> ML1	GTF-ML1	1772	35		65		(Kralj <i>et al.</i> , 2004a)
Reuterane-saccharases							
<i>Lactobacillus reuteri</i> 121	GTF-A	1781	43	57			(Kralj <i>et al.</i> , 2002)
<i>Lactobacillus reuteri</i> BioGaiia	GTF-O	1781	22	78			(Kralj <i>et al.</i> , 2005a)
Alternane-saccharase							
<i>Leuconostoc citreum</i> B-1355	ASR	2057	45		55		(Argüello-Morales <i>et al.</i> , 2000)

*Maj. :majoritaire ; nd : non déterminé.

I.3 Relations structure/fonction des glucane-saccharases

I.3.1 Organisation structurale des glucane-saccharases de la famille GH13

I.3.1.A Repliement global

L'amylosaccharase de *Neisseria polysaccharea* (ASNp) est une enzyme de 636 acides aminés. L'isolement, le clonage du gène codant pour cette enzyme (Potocki De Montalk et al., 1999) et sa purification à l'homogénéité ont permis la cristallisation et la résolution de sa structure à 1,4 Å, en 2001 (Skov et al., 2001). Elle est restée, jusqu' 'en juillet 2010, la seule structure tridimensionnelle de glucane-saccharase disponible dans la Protein data Bank (PDB). La structure de l'ASNp est organisée en 5 domaines appelés N, A, B, B' et C (Figure 3). Le domaine A, correspondant au domaine catalytique, adopte un repliement de type tonneau (β/α)₈ comme l'ensemble des enzymes de la famille GH13. Le domaine A est composé des résidus 98 à 184, 261 à 395 et 461 à 550. Certaines boucles du tonneau (β/α)₈ sont suffisamment longues pour constituer des domaines à part entière. C'est le cas de la boucle 3 (située entre le brin β 3 et l'hélice α 3) comprenant les résidus 185 à 260 qui constituent le domaine B. La boucle 7, comprenant les résidus 395 à 460 et définissant le domaine B' est quant à elle, spécifique de l'ASNp et n'a pas d'équivalent chez les α -amylases. Le domaine N-terminal de cette enzyme (résidus 1 à 90) est constitué de 6 hélices α dont 2 ont un rôle de soutien des hélices α 3 et α 4 du tonneau, contribuant ainsi au maintien de l'intégrité structurale. Enfin, le domaine C (résidus 555 à 636) est formé d'une succession de feuillets β antiparallèles formant une structure de clé grecque.

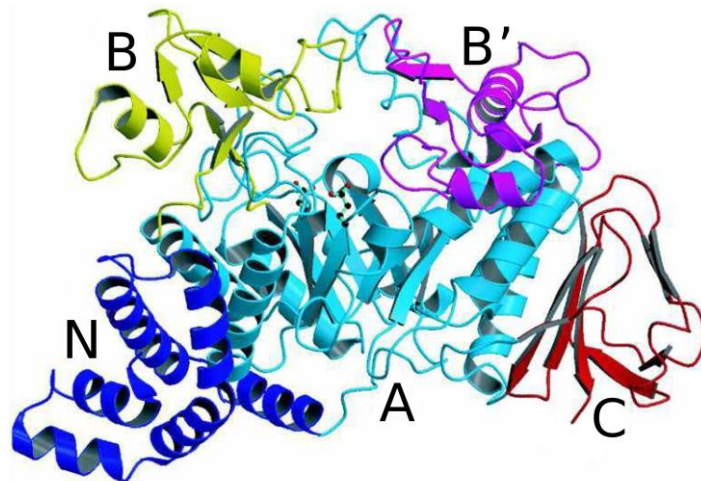


Figure 3 : Structure tridimensionnelle de l'amylosaccharase de *N. polysaccharea* (code PDB : 1G5A). Légende des domaines: N (bleu sombre), A (bleu clair), B (jaune), B' (magenta) et C (rouge). Les résidus catalytiques Asp286 et Glu328, respectivement localisés à la fin des feuillets β -4 et β -5 dans le domaine A, sont également indiqués.

I.3.1.B Architecture du site actif et fixation des substrats glucosidiques

La diade catalytique de l'ASNp, comprenant le nucléophile Asp286 et l'acide/base Glu328, se situe au fond d'une poche étroite recouverte par les domaines B et B'. L'analyse de la structure cristallographique de l'enzyme en complexe avec du glucose, du saccharose et du maltoheptaose a permis d'identifier plusieurs sous-sites de fixation des accepteurs (Mirza *et al.*, 2001; Skov *et al.*, 2002). Une nomenclature a été proposée pour les sous-sites de fixation des sucres dans le site actif des glycoside-hydrolases (Davies *et al.*, 1997). Les sous-sites sont numérotés de - n à + n, - n correspondant à la partie non-réductrice et + n à la partie réductrice. La rupture de la liaison osidique s'opère entre les positions -1 et +1. Le nombre et la nature des sous-sites est variable et est caractéristique de chaque enzyme. Dans le cas de l'amylosaccharase, la présence d'un pont salin entre les résidus Asp144 et Arg509 limite les sous-sites de la partie glycone au sous-site -1 (figure 4A) (Mirza *et al.*, 2001).

L'analyse structurale en complexe avec le maltoheptaose a permis d'identifier 3 sites de fixation des oligosaccharides, nommés OB1, OB2 et OB3 (OB pour Oligosaccharide Binding site). Seul OB1 se situe dans le site actif (figure 4B) (Skov *et al.*, 2002). Au total, 7 sous-sites allant de -1 à +6 ont été définis (figure 4). L'arrimage des substrats glycosidiques (saccharose et maltoheptaose) fait intervenir des interactions, principalement des liaisons hydrogène, au niveau de chaque résidu osidique impliquant un ou plusieurs acides aminés. Un phénomène de double stacking intervient également au niveau de l'unité glucose de la partie glycone.

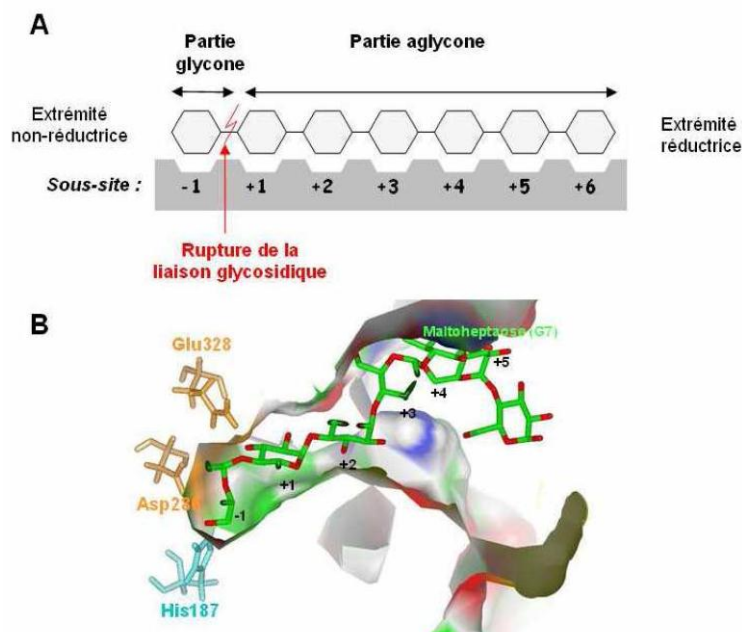


Figure 4 : Représentation des sous-sites de fixation du maltoheptaose au niveau du site catalytique de l'ASNp. A : Représentation selon la nomenclature proposée par Davies et al. (Davies and Henrissat, 1995). B : Représentation de l'ASNp en complexe avec le maltoheptaose dans le site actif (Champion, 2008).

1.3.2 Mécanisme catalytique des glucane-saccharases de la famille GH13

Les enzymes de la famille GH13 partagent toutes le même mécanisme de clivage de la liaison osidique s'opérant par rétention de la configuration anomérique sous forme α . Il fait intervenir trois résidus essentiels, comprenant deux acides aspartiques et un acide glutamique (Matsuura *et al.*, 1984; Svensson, 1994), strictement conservés au sein de cette famille. Le mécanisme a bien été étudié pour l'ASNp (figure 5).

La première étape (glycosylation) correspond au clivage de la liaison osidique du saccharose. Le groupement acide de l'acide glutamique 328 protone l'oxygène de la liaison osidique du saccharose. Simultanément, le carbone anomérique de l'unité glucosyle du saccharose subit une attaque nucléophile par l'acide aspartique 286, aboutissant à la formation d'un intermédiaire covalent β -D-glucosyl-enzyme (Jensen *et al.*, 2004) et au relargage du fructose. Au cours de la seconde étape (déglycosylation), l'acide glutamique 328 déprotoné joue le rôle de base et active une molécule acceptrice qui attaque l'intermédiaire covalent β -D-glucosyl-enzyme. Si la molécule acceptrice est une molécule d'eau, la réaction aboutit à l'hydrolyse de la molécule de saccharose. Si c'est un sucre ou une autre molécule hydroxylée, la réaction de transglucosylation aboutit à la formation d'un nouvel oligosaccharide. Les deux étapes sont accompagnées de la formation d'états de transitions de type oxocarbanium (Sinnott, 1990). Le second acide aspartique, Asp393, intervient quant à lui dans la stabilisation de l'intermédiaire covalent (Mosi *et al.*, 1997; Uitdehaag *et al.*, 1999; Yoshioka *et al.*, 1997). La caractérisation biochimique détaillée de l'ASNp réalisée par Albenne *et al.* décrit avec précision les étapes menant à la formation du polymère de type amylose. Au début de la réaction, l'unité glucosyle du complexe covalent β -D-glucosyl-enzyme est transférée sur une molécule d'eau. Le glucose ainsi libéré joue le rôle de molécule acceptrice pour être glucosylé et former du maltose qui est à son tour glucosylé pour former du maltotriose. La polymérisation se poursuit pour aboutir à la formation de maltooligosaccharides de tailles croissantes jusqu'à l'obtention d'une taille moyenne et d'une concentration critique responsables de la précipitation de l'amylose (Albenne *et al.*, 2004; Potocki-Veronese *et al.*, 2005). Le mécanisme de polymérisation est donc non-processif, ou multi-chaines. De plus, l'étude de la structure du variant inactif Glu238Gln de l'ASNp en complexe avec le maltoheptaose démontre clairement que l'extrémité non-réductrice de la chaîne oligosaccharidique est située à proximité des résidus catalytiques, confirmant ainsi que le transfert des unités glucosyle s'opère par l'extrémité non-réductrice.

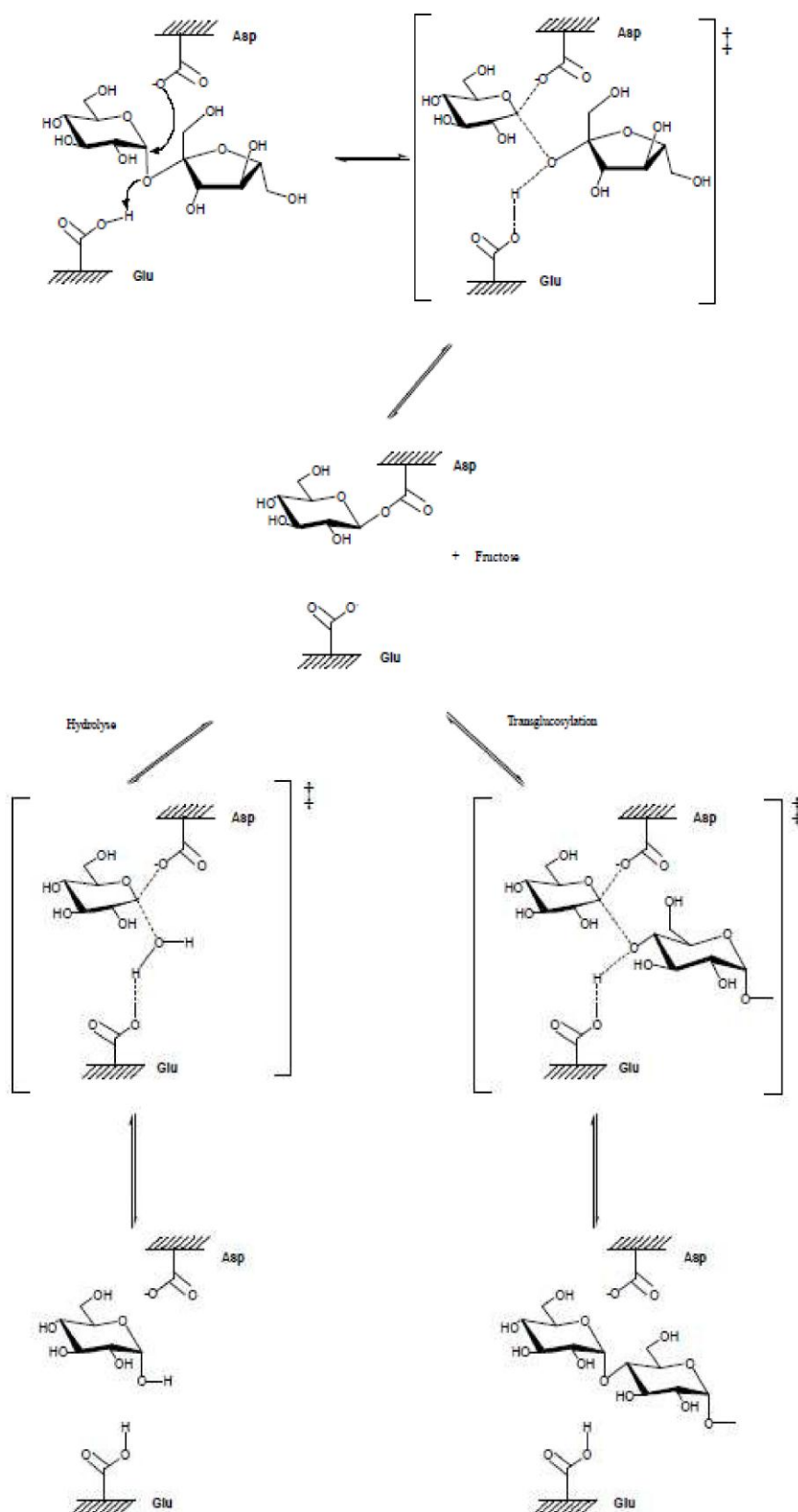


Figure 5 : Mécanisme catalytique avec rétention de la configuration anomérique du substrat, identifié pour l'ASNp (Skov *et al.*, 2001).

1.3.3 Organisation structurale des glucane-saccharases de la famille GH70

Jusqu'en juillet 2010, aucune structure tridimensionnelle d'une glucane-saccharase de la famille GH70 n'était disponible dans la PDB. La principale raison est la difficulté d'obtention d'une seule forme d'enzyme soluble, purifiée à homogénéité. et cristallisable. En effet, les glucane-saccharases natives présentent une masse molaire élevée supérieure à 150 kDa, possèdent probablement des régions flexibles, et sont souvent produites sous formes dégradées. Ce phénomène est probablement dû, chez *L. mesenteroides*, à la présence de protéases extracellulaires responsables de la dégradation de ces enzymes (Arguello-Morales *et al.*, 2005; Sanchez-Gonzalez *et al.*, 1999). En système hétérologue, chez *E. coli*, la production de formes dégradées est également observée dans la plupart des cas (Bozonnet *et al.*, 2002; Joucla *et al.*, 2006; Monchois *et al.*, 1997; Moulis *et al.*, 2006a) de même que la présence de corps d'inclusions (Brison, 2010; Joucla *et al.*, 2006). Seule la troncature rationnelle des glucane-saccharases de la famille GH70 a permis d'améliorer l'expression et la purification de ces protéines sous forme soluble (Brison, 2010; Ito *et al.*, 2010; Joucla *et al.*, 2006; Moulis *et al.*, 2006a; Pijning *et al.*, 2008; Vujicic-Zagar *et al.*, 2010).

Ces efforts ont finalement permis d'obtenir ces dernières années des cristaux de glucane-saccharases de la famille 70. La première structure cristallographique a été rendue publique très récemment. Il s'agit de celle de la glucane-saccharase GTF180 tronquée de son extrémité N-terminale, dont les coordonnées ont été relâchées en juillet 2010 (code PDB : 3KLK) (Pijning *et al.*, 2008; Vujicic-Zagar *et al.*, 2010). La dextrane saccharase GTF180 produite par *Lb. reuteri* 180 est une enzyme spécifique de la synthèse d'un glucane composé de 61 % liaisons de type $\alpha(1\rightarrow6)$ et 39 % de liaisons $\alpha(1\rightarrow3)$ (Van Leeuwen *et al.*, 2008c). La structure tridimensionnelle d'une mutane-saccharase de la famille 70, la GTF-SI, vient également d'être publiée et les coordonnées ont été rendues disponibles en avril 2011 (code PDB : 3AIE) (Ito *et al.*, 2010; Ito *et al.*, 2011). Enfin, une troisième structure d'enzyme de la famille 70 a également été résolue et sera prochainement publiée (Brison, 2010). La forme cristallisée et analysée par diffraction des rayons X est celle de ΔN_{123} -GBD-CD2, qui provient de la troncature de DSR-E, seule glucane-saccharase dotée de deux domaines catalytiques classés dans la famille 70.

1.3.3.A Structure globale et fonction

La structure de la GTF180- ΔN s'organise en cinq domaines distincts dénommés A, B, C, IV et V (figure 6). La même organisation a été observée pour la GTF-SI, à l'exception du domaine V qui a été tronqué pour permettre la cristallisation de l'enzyme, et qui n'est donc pas présent dans la dernière structure

proposée. Sur ces cinq domaines, seuls les domaines A, B, et C présentent une homologie de structure significative avec les domaines A, B et C des enzymes de la famille GH13. Le repliement global des protéines est en forme de « U », structure très originale résultant de l'enchevêtrement de deux segments polypeptidiques provenant des parties N- et C-terminales de l'enzyme. Seul le domaine C est formé d'un unique segment peptidique qui forme la base du U. En accord avec les prédictions de structures secondaires (MacGregor *et al.*, 1996), le domaine catalytique constitué des domaines A et B est organisé en tonneau $(\beta/\alpha)_8$ et présente une permutation circulaire par rapport à la structure du domaine catalytique des glucane-saccharases de la famille GH13 (Figure 7), le domaine C étant inséré entre l'hélice $\alpha 8$ et le feuillet $\beta 1$ du tonneau. La boucle localisée entre le feuillet $\beta 7$ et l'hélice $\alpha 7$, formée de deux hélices H1 et H2 inexistantes chez les GH13, contribue à la formation de la gorge catalytique. Vujicic-Zagar *et al.* ont proposé un modèle évolutif de « permutation par duplication » (Peisajovich *et al.*, 2006) permettant d'expliquer cette permutation circulaire (Vujicic-Zagar *et al.*, 2010). Le domaine catalytique des GH70 renferme quatre segments très conservés avec la famille des α -amylases (Svensson, 1994), nommés motifs I,II,III et IV (figure 8). Ils sont situés au niveau des feuillets $\beta 8$ (motif I), $\beta 1$ (motif II), $\beta 2$ (motif III) et $\beta 4$ (motif IV). Ces segments contiennent la diade catalytique comprenant le résidu nucléophile (motif II) et l'acide/base (motif III) ainsi que le probable stabilisateur de l'intermédiaire covalent (motif IV). Ces motifs sont également impliqués dans la spécificité de liaison osidique (cf. I.3.7).

Le domaine IV ne présente aucune homologie structurale avec d'autres domaines protéiques répertoriés dans la PDB. Pour le moment sa fonction n'est pas connue.

Le domaine V adopte une conformation de type β -solénoïde triangulaire. Il participe à la fixation du glucane (GBD) (Monchois *et al.*, 1999; Suwannarangsee *et al.*, 2007; Van Hijum *et al.*, 2006) et est impliqué dans la réaction de polymérisation, comme démontré pour l'enzyme DSR-S de *Ln. mesenteroides* B-512F (Moulis *et al.*, 2006b). En effet, il est composé d'unités répétées, classées selon leurs séquences en motifs A, B, C, D et N qui constituent des zones d'ancrage du polymère en cours d'élongation (Banas *et al.*, 1990; Ferretti *et al.*, 1987; Giffard *et al.*, 1991; Monchois *et al.*, 1998; Russell, 1990). Ces éléments peuvent être regroupés en un motif commun appelé YG (Giffard and Jacques, 1994). Le GBD peut-être également impliqué dans la fixation des glucane-saccharases à la surface bactérienne comme c'est le cas pour l'enzyme DSR-P produite par *Ln. mesenteroides* IBT-QP (Olvera *et al.*, 2007).

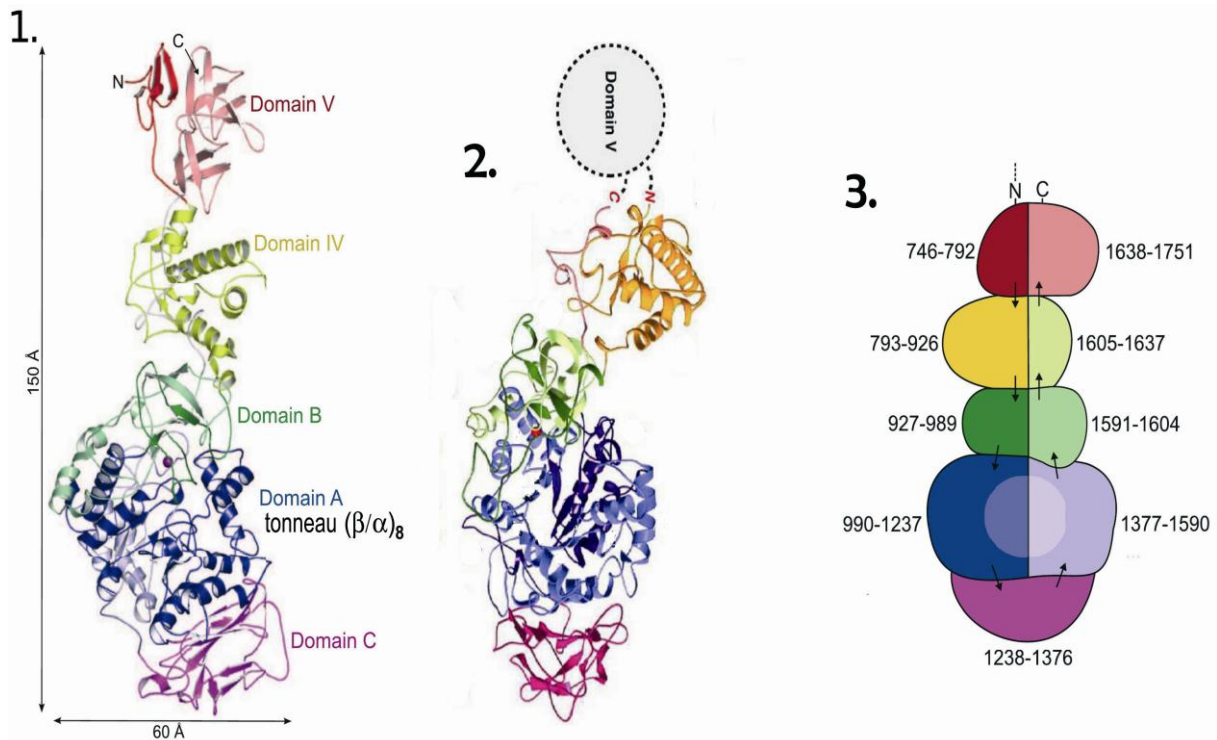


Figure 6 : Structure globale des glucane-saccharases GTF180-ΔN et GTF-SI. 1. Structure cristalline de la GTF180-ΔN (code PDB : 3KLK). La sphère rouge représente l'ion calcium présent entre les domaines A et B; 2. Structure cristalline de la GTF-SI (code PDB : 3AIE). 3. Représentation schématique de la conformation en U de la GTF180-ΔN. Adapté de (Ito *et al.*, 2011; Vujicic-Zagar *et al.*, 2010).

Trois formes cristallines de l'enzyme GTF180 ont été observées lors de sa cristallisation. Si aucun changement n'a été constaté pour les domaines A, B, C et IV, une forme cristalline présente l'enzyme avec le domaine V rabattu sur la gorge catalytique. Ceci pourrait correspondre à une conformation « fermée » de l'enzyme, limitant partiellement l'accès au site catalytique.

Très récemment, les travaux menés par Brison *et al.* ont abouti à la résolution de la structure d'une autre transglucosidase de la famille GH70 (Brison, 2010). Il s'agit du variant tronqué ΔN₁₂₃-GBD-CD2, variant de la transglucosidase GBD-CD2 (Fabre *et al.*, 2005), elle-même issue de la troncature rationnelle de la dextrane-saccharase DSR-E produite par *Ln. mesenteroides* B-1299 (récemment reclassée *Ln. citreum*, non publié) et synthétisant un polymère original présentant des ramifications liées en α(1→2) (Bozonnet *et al.*, 2002; Fabre *et al.*, 2005). Le repliement adopté par ΔN₁₂₃-GBD-CD2 est très similaire au repliement des enzymes GTF180-ΔN et GTF-SI.

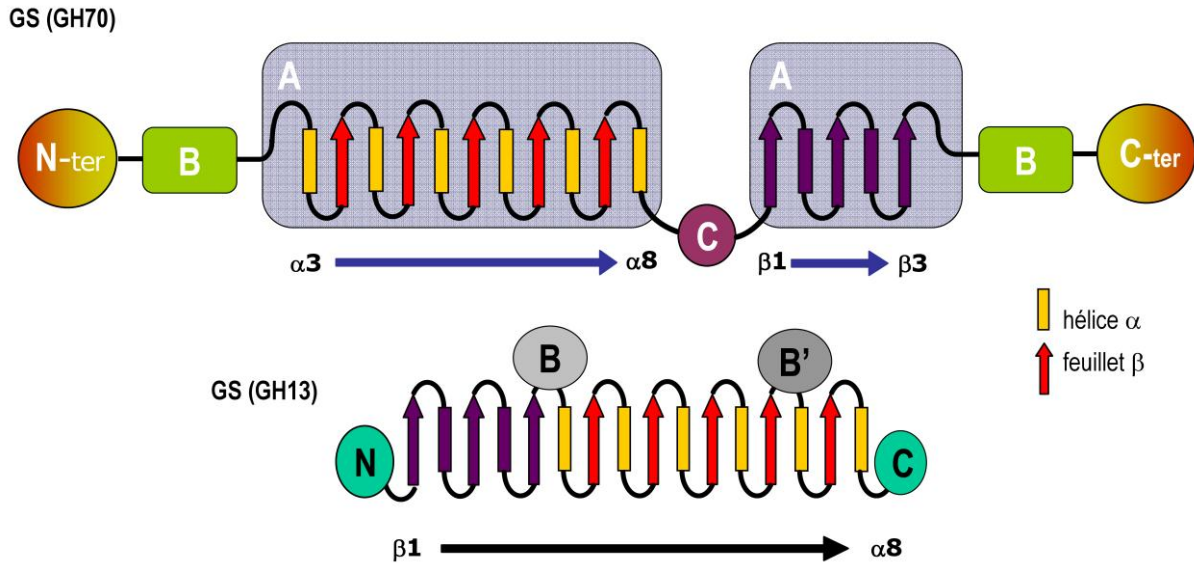


Figure 7 : Représentation schématique de la structure secondaire des glucane-saccharases des familles GH13 et GH70. Le domaine catalytique des amylosaccharases (GH13) est organisé en tonneau (β/α)₈ correspondant à une succession alternée de huit feuillets β et hélices α . En ce qui concerne les glucane-saccharases de la famille GH70, le tonneau commence avec l'hélice α 3 et se termine avec le feuillet β 3. De ce fait, le segment regroupant les éléments β 1 α 1 β 2 α 2 β 3 est situé à l'extrémité C-terminale du domaine catalytique. A : domaine A ; B et B' : domaines B et B' ; C : domaine C ; N-ter : Domaine N-terminal (comprenant les domaines IV et V chez les GH70) ; C-ter : Domaine C-terminal (comprenant les domaines IV et V chez les GH70).

Motif I		Motif III	
GTF180	1503 ADWVPDQ	GTF180	1059 INILEDWG
DSR-S	1022 ADWVPDQ	DSR-S	585 LSILEDWS
CD1 (DSR-E)	1013 NDWVPDQ	CD1 (DSR-E)	561 ISILEDWD
CD2 (DSR-E)	2688 ADVVDNQ	CD2 (DSR-E)	2244 ISLVEAGL
Amylosaccharase	181 VDFIFNH	Amylosaccharase	324 FFKSEAIW
TAKA-amylase	116 VDVVANH	TAKA-amylase	226 YCIGEVLD
CGTase	134 IDFAPNH	CGTase	253 FTFGEWFL
Motif II		Motif IV	
GTF180	1021 GIRVDAVDNVDVLLSI	GTF180	1131 FVRAHDSNAQDQI
DSR-S	547 GIRVDAVDNVDADLLQI	DSR-S	657 FVRAHDSVQTVI
CD1 (DSR-E)	523 GIRVDAVDNVDADLLQI	CD1 (DSR-E)	633 FIRAHDSVQTVI
CD2 (DSR-E)	2206 SIRIDAVDFIHNDTIQR	CD2 (DSR-E)	2317 IIAHADKGVQEKI
Amylosaccharase	282 ILRMDAVAFIWKQMGTS	Amylosaccharase	388 YVRSRDDDIGWTF
TAKA-amylase	202 GLRIDTVKHVQKDFWPG	TAKA-amylase	292 FVENHDDNPRFAS
CGTase	225 GIRVDAVKHMPGLGWQKS	CGTase	323 FIDNHDMDFKTS

Figure 8 : Motifs conservés entre les enzymes de la famille GH13 et GH70. En rouge : résidus strictement conservés correspondant à la diade catalytique et au stabilisateur de l'intermédiaire covalent. En noir et gras : autres résidus conservés entre les deux familles d'enzymes. En bleu : résidus formant le sous-site -1 de la GTF180- Δ N (Vujicic-Zagar *et al.*, 2010).

1.3.3.B Sous-sites de fixation au niveau du site actif

La résolution de la structure de la GTF180- Δ N en complexe avec le saccharose et le maltose a permis de cartographier les sous-sites de fixation -1, +1, +2 et +3 (Figure 9) (Vujicic-Zagar *et al.*, 2010). Ainsi, six des sept résidus établissant une interaction avec l'unité glucosyle positionnée au niveau du sous-site -1 sont strictement conservés entre les glucane-saccharases des familles GH13 et GH70 (figure 8 et 9).

Parmi ces résidus on retrouve le nucléophile Asp1025 et l'acide/base Glu1063 orienté vers le carbone anomérique de l'unité glucosyle du sous-site -1 et l'oxygène de la liaison glycosidique, respectivement. On retrouve également les résidus Asp1136 (stabilisateur de l'état de transition), His1135, Arg1023 et Asp1504. Le septième résidu est la Gln1509, strictement conservé au sein de la famille GH70 et remplacé par une histidine chez les enzymes de la famille GH13. En présence de saccharose, 5 autres résidus établissent des liaisons hydrogène avec l'unité fructosyle au niveau du sous-site +1. Il s'agit du Glu1063, Trp1065, Asn1029, Asp1136 et Gln1140.

Quatre sites de fixation du maltose ont été identifiés et nommés M1, M2, M3 et M4. Les sites M3 et M4 se trouvent à la surface des domaines A et B, respectivement, et sont donc éloignés du site actif. De plus, ces deux sites ne sont pas conservés chez les GH70. Les sites M1 et M2 se situent dans la gorge catalytique. La molécule de maltose en M2 est en mesure d'établir des liaisons hydrogène avec les résidus Asp1028, Trp1065, Trp1067 et Ala978 par l'intermédiaire de molécules d'eau. L'unité glucosyle en +1 du maltose en M1 est insérée entre les résidus Trp1065, Leu938 et Leu981. Une liaison hydrogène est établie entre le résidu Asn1029 et le groupe hydroxyle du C4 en +1. L'unité glucosyle en +2 établit quant à elle des interactions de van der Waals avec le Trp1065. Notons que le groupement hydroxyle du C6 en +1 est orienté vers l'acide/base Glu1063. Il est donc favorablement positionné pour attaquer le complexe covalent β -D-glucosyl-enzyme et établir une liaison de type $\alpha(1\rightarrow6)$.

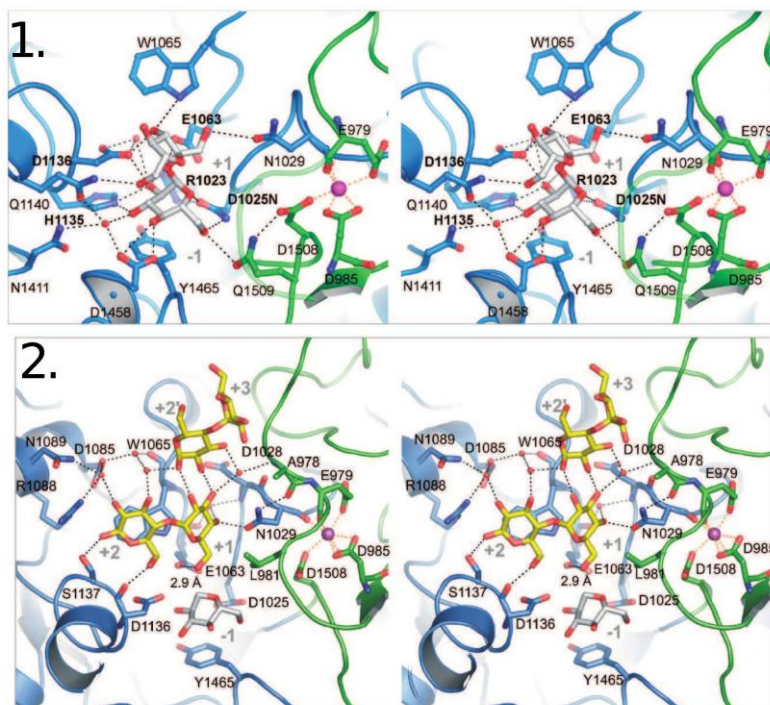


Figure 9 : Vues stéréo de la GTF180- Δ N en complexe avec du saccharose et du maltose. 1. Saccharose lié aux sous-sites -1 et +1 du mutant inactif D1025N de la GTF180- Δ N. Les résidus en gras correspondent aux résidus strictement conservés entre les glucane-saccharases des familles GH13 et GH70. 2. Maltose M1 lié aux sous-sites +1 et +2 et Maltose M2 lié aux sous-sites +2 et +3. Les résidus du domaine A sont représentés en bleu. Les résidus du domaine B sont représentés en vert. La sphère magenta représente l'ion calcium. Les sphères rouges représentent les molécules d'eau. Les pointillés noirs représentent les liaisons hydrogène. (Codes PDB : 3HZ3 et 3KLL).

La résolution de la structure de la GTF-SI en complexe avec de l'acarbose et du maltose a également permis de cartographier les sous-sites de fixation du domaine catalytique de l'enzyme avec ces ligands (Ito *et al.*, 2011). Si les 10 résidus qui composent le sous-site -1 de la GTF-SI sont relativement bien conservés entre les familles GH70 et GH13 (Figure 10), en revanche, ceux du sous-site +1 ne le sont pas. Sur les 6 résidus qui composent le sous-site +1 identifié en présence d'acarbose, 4 sont identiques au sous-site +1 de la GTF180 (Glu515, Asp588, Asn481 et Trp517), les deux autres correspondant aux résidus Leu433 et Asp480.

Comparé à la GTF180, un seul site de fixation du maltose a été identifié chez la GTF-SI et se trouve au niveau des sous-sites +1, +2 et une partie en +3 (Figure 10).

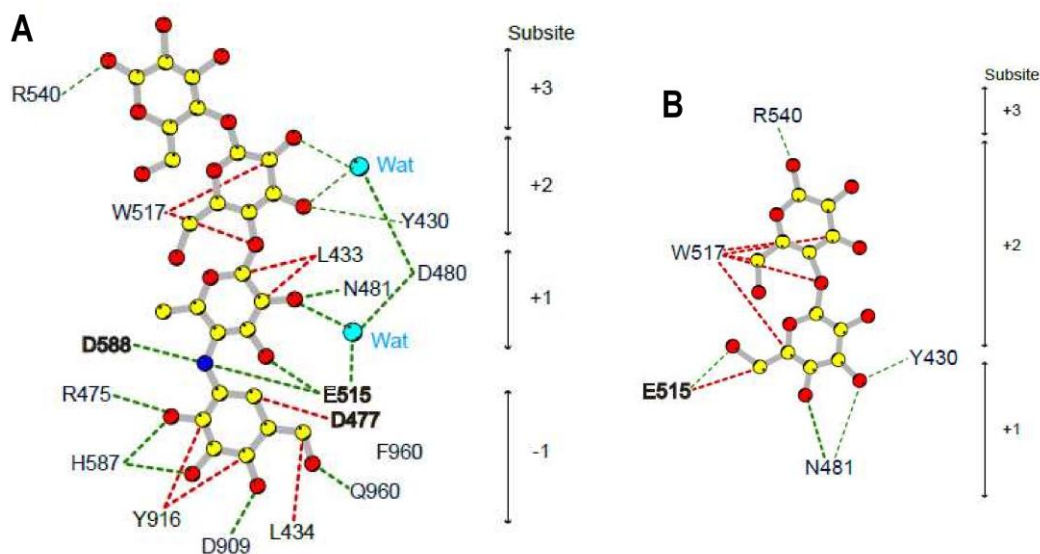


Figure 10 : Représentation schématique des interactions (A) acarbose-GTF-SI et (B) maltose-GTF-SI. Les résidus de la diade catalytique et le stabilisateur de l'intermédiaire covalent sont représentés en gras. Pointillés verts et gras : liaisons hydrogènes ; pointillés rouges et gras : interactions hydrophobes ; pointillés verts : interactions hydrophiles faibles. Tiré de (Ito *et al.*, 2011).

La superposition réalisée par Brison *et al.* de la structure de ΔN_{123} -GBD-CD2 avec la GTF180- ΔN en complexe avec une molécule de saccharose montre une forte similarité de structure au niveau des sous-sites -1 et +1 (Brison, 2010). L'ensemble des résidus du sous-site -1 défini pour la GTF180- ΔN sont strictement conservés chez ΔN_{123} -GBD-CD2. Au niveau du sous-site +1, 5 des 7 résidus interagissant avec l'unité fructosyle chez la GTF180- ΔN sont conservés chez ΔN_{123} -GBD-CD2. Il s'agit des résidus Leu2166, Leu2167, Glu2248, Gln2326 et Asp2322. Les résidus Asn1029 et Trp1065 de GTF180- ΔN , très bien conservés chez les glucane-saccharases GH70, sont substitués par Phe2214 et

Ala2249 respectivement. Ces deux résidus n'établissent pas de liaisons hydrogène avec l'unité fructosyle du saccharose. Ceci laisse suggérer que l'unité fructosyle est faiblement liée au sous-site +1 de ΔN_{123} -GBD-CD2, ce qui est en accord avec la valeur du K_m apparent pour le saccharose qui est 32 fois plus important chez ΔN_{123} -GBD-CD2 que celle déterminée pour GTF180- ΔN (Brison, 2010).

1.3.4 Mécanisme catalytique des glucane-saccharases de la famille GH70

La forte similarité structurale existant entre les enzymes des familles GH13 et GH70 suggère qu'elles partagent un même mécanisme catalytique. En effet, avant même la résolution de la structure de la GTF-180- ΔN , des alignements de séquences des enzymes de la famille GH13 et GH70 avaient déjà mis en évidence six des sept résidus les mieux conservés au sein de la famille GH13 (figure 9), parmi lesquels on retrouve les résidus de la diade catalytique (Asp551 et Glu589 chez DSR-S) et l'acide aspartique stabilisateur de l'intermédiaire covalent (Asp662 chez DSR-S), résidus essentiels à la catalyse (Devulapalle *et al.*, 1997; Kato *et al.*, 1992; MacGregor *et al.*, 1996; Monchois *et al.*, 1997). De plus, l'intermédiaire glucosyl-enzyme a été isolé pour la glucane-saccharase de *Streptococcus sobrinus*, démontrant ainsi que la catalyse se produit via cet intermédiaire (Mooser and Iwaoka, 1989).

1.3.5 Mécanisme de polymérisation des glucane-saccharases de la famille GH70

Les travaux de Moulis *et al.* ont permis de proposer un schéma réactionnel détaillé de la formation de polymère à partir du saccharose (Moulis *et al.*, 2006b). La synthèse de polymère démarre par l'hydrolyse du saccharose. Le saccharose, puis le glucose, jouent le rôle de molécules acceptrices et seraient donc glucosylés pour former des oligosaccharides. L'élongation se poursuit via le transfert de l'unité glucosyle à l'extrémité non-réductrice des accepteurs précédemment formés. Avec la diminution de la concentration en saccharose, le glucose est préférentiellement glucosylé pour former des oligosaccharides jusqu'à l'obtention d'un polymère de haute masse molaire. En fin de réaction, la concentration en fructose devenant majoritaire, ce dernier jouerait également le rôle d'accepteur pour former des isomères de saccharose tels que le leucrose (Moulis *et al.*, 2006b).

Moulis *et al.* proposent ainsi un mécanisme de polymérisation par l'extrémité non réductrice selon un mode semi-processif faisant intervenir un seul domaine catalytique. Ces conclusions sont confirmées par la résolution de la structure cristallographique de la GTF-180- ΔN qui démontre la présence d'un seul site catalytique comprenant un résidu nucléophile.

I.3.6 Mécanisme de réaction d'accepteur des glucane-saccharases de la famille GH70

L'introduction d'une molécule acceptrice exogène présentant un ou plusieurs groupements hydroxyle libres, comme par exemple les sucres, alcools, composés aromatiques ou dérivés d'acides aminés, entraîne la formation d'oligosaccharides ou de glucoconjugués (Bertrand *et al.*, 2006; Côté and Dunlap, 2003; Côté *et al.*, 2009; Moon *et al.*, 2007; Remaud-Simeon *et al.*, 2000; Seibel *et al.*, 2006).

Les accepteurs peuvent être classés en deux catégories :

- les accepteurs forts, tels le maltose et l'isomaltose, permettant la formation de fortes quantités d'oligosaccharides au détriment de la synthèse de glucane.

- les accepteurs faibles, moins bien reconnus qui sont glucosylés avec des rendements plus faibles (Robyt and Eklund, 1983).

De même que pour la réaction de polymérisation, l'unité glucosyle du saccharose est transférée du côté non-réducteur de la molécule acceptrice (Côté and Robyt, 1984; Germaine and Schachtele, 1976).

I.3.7 Déterminants de la spécificité des glucane-saccharases de la famille GH70

En l'absence de structure tridimensionnelle jusqu'en 2010, l'identification des facteurs structuraux influençant la spécificité des glucane-saccharases de la famille GH70 a été guidée par des comparaisons de structures primaires d'enzymes de spécificités variées. C'est ainsi qu'au début des années 1990, Shimamura *et al.* ont remarqué que certains résidus très conservés chez les dextrane-saccharases synthétisant un polymère soluble (formé majoritairement de liaisons de type $\alpha(1\rightarrow6)$) ne sont pas conservés chez les enzymes synthétisant un polymère insoluble (formé majoritairement de liaisons de type $\alpha(1\rightarrow3)$) (Shimamura *et al.*, 1994). En particulier, les enzymes synthétisant un polymère soluble possèdent une thréonine (Thr667 chez DSR-S) à proximité de l'acide aspartique stabilisateur de l'intermédiaire covalent, alors que les enzymes synthétisant un polymère insoluble possèdent à cette même position un acide aspartique. Des travaux de mutagenèse ont confirmé l'importance de ces résidus dans la spécificité des enzymes de la famille GH70. En effet, le remplacement de cette thréonine (Thr589) chez la GTF-D de *Streptococcus mutans* GS 5 (synthétisant un polymère formé à 70 % de liaisons de type $\alpha(1\rightarrow6)$) par un acide aspartique ou un acide glutamique aboutit à la synthèse d'un polymère présentant 85 et 98 % de liaisons de type $\alpha(1\rightarrow3)$, respectivement. Inversement, le remplacement de l'acide aspartique (Asp567) chez la GTF-B de *Streptococcus mutans* GS 5 (synthétisant un polymère formé à 88 % de liaisons de type $\alpha(1\rightarrow3)$) par une thréonine aboutit à la synthèse d'un polymère présentant 24 % de liaison de type $\alpha(1\rightarrow6)$. Cette même mutation générée

chez la GTF-I de *Streptococcus downei* Mfe 28 (synthétisant un polymère formé à 88 % de liaisons $\alpha(1\rightarrow3)$) engendre la production d'une fraction de polymère soluble (Monchois *et al.*, 2000).

Les résidus précédemment cités sont situés au niveau de la séquence consensus dénommée « motif IV ». D'autres résidus présents sur ce motif ont également fait l'objet de travaux de mutagenèse dirigée. En effet, on peut remarquer que la séquence consensus SEV directement en aval de l'acide aspartique stabilisateur de l'intermédiaire covalent ($^{663}\text{SEV}^{665}$ chez DSR-S) est divergente chez les glucane-saccharases synthétisant des polymères de structures différentes. Chez la reuterane-saccharase GTF-A de *Lb. reuteri* 121, cette séquence est remplacée par le tripeptide $^{1134}\text{NNS}^{1136}$. De même, chez l'alternane-saccharase ASR de *Ln. citreum* B-1355 (alternant les liaisons $\alpha(1\rightarrow3)/\alpha(1\rightarrow6)$) et le second domaine catalytique de la dextrane-saccharase DSR-E de *Ln. citreum* B-1299 (spécifique des liaisons $\alpha(1\rightarrow2)$), on retrouve les tripeptides $^{768}\text{YDA}^{770}$ et $^{2323}\text{KGV}^{2325}$, respectivement.

Le remplacement du tripeptide $^{1134}\text{NNS}^{1136}$ par la séquence SEV chez l'enzyme GTF-A de *Lb. reuteri* 121 (spécifique des liaisons $\alpha(1\rightarrow4)$) provoque une augmentation importante du taux de liaisons $\alpha(1\rightarrow6)$ dans les produits synthétisés (84 % du glucane synthétisé). Cependant, ce changement de spécificité est accompagné d'une chute de l'affinité de l'enzyme pour le saccharose, révélant l'importance du motif IV dans la liaison au substrat (Kralj *et al.*, 2005b). Les auteurs ont également muté les positions P1026V et I1029V (motif II, V553 et V556 chez DSR-S) en supplément des mutations $\text{NNS}^{1134}\text{SEV}$. Le mutant obtenu synthétise alors un polymère constitué de 92 % de liaisons de type $\alpha(1\rightarrow6)$, laissant supposer que la spécificité de cette enzyme peut être contrôlée en jouant sur la nature des acides aminés présents sur les motifs II et IV. Les mêmes auteurs ont étudié le rôle de chaque résidu du tripeptide $^{1134}\text{NNS}^{1136}$ de l'enzyme GTF-A. Des mutations ont donc été générées de façon indépendante ou en combinaison, pour obtenir ainsi les mutants $^{1134}\text{SNS}^{1136}$, $^{1134}\text{NES}^{1136}$, $^{1134}\text{NNV}^{1136}$, $^{1134}\text{SNV}^{1136}$, $^{1134}\text{SES}^{1136}$ et $^{1134}\text{NEV}^{1136}$. Les résultats ont souligné le rôle important joué par l'asparagine 1134 dont la mutation affecte systématiquement la spécificité de l'enzyme en faveur de la synthèse de liaisons de type $\alpha(1\rightarrow6)$ (Kralj *et al.*, 2006).

Van Leeuwen *et al.* ont caractérisé le glucane synthétisé par le mutant V1027P:S1137N:A1139S, affectant des résidus présents sur les motifs II et IV de la dextrane-saccharase GTF-180 de *Lb. reuteri* 180. Le glucane synthétisé par ce mutant est formé à 60% de liaisons $\alpha(1\rightarrow6)$, 28 % de liaisons $\alpha(1\rightarrow3)$ et 12 % de liaisons $\alpha(1\rightarrow4)$. Ainsi, ce mutant présente une nouvelle spécificité par rapport à l'enzyme sauvage, qui ne synthétise pas de liaisons de type $\alpha(1\rightarrow4)$ (Van Leeuwen *et al.*, 2008a).

Suivant une approche similaire, van Leeuwen *et al.* ont caractérisé les polymères synthétisés par une série de mutants de la dextrane-saccharase GTF-180 de *Lb. reuteri* 180. Les mutations ciblent encore une fois le tripeptide $^{1137}\text{SNA}^{1139}$ en aval de l'acide aspartique stabilisateur du complexe covalent,

additionné du résidu glutamine Q1140 (Q666 chez DSR-S). Ainsi, par de simples mutations, les auteurs ont obtenu une large gamme de polysaccharides modifiés. En effet, ces derniers présentent des taux de branchement de 8 à 20 % (pour les variants ¹¹³⁷SNAH¹¹⁴⁰ et ¹¹³⁷YDAQ¹¹⁴⁰ respectivement), des taux de résidus liés en $\alpha(1\rightarrow6)$ variant de 52 à 86 % (¹¹³⁷YDAQ¹¹⁴⁰ et ¹¹³⁷SNAH¹¹⁴⁰ respectivement), des taux de résidus liés en $\alpha(1\rightarrow3)$ variant de 14 à 39 % (¹¹³⁷SNAH¹¹⁴⁰ et ¹¹³⁷YDAQ¹¹⁴⁰ respectivement) et des taux de résidus liés $\alpha(1\rightarrow4)$ pouvant atteindre 8 % (¹¹³⁷YDAQ¹¹⁴⁰) (Van Leeuwen *et al.*, 2009).

Le mutant YDA768SEV du variant tronqué de l'alternane-saccharase ASR C-APY del construit par Moulis *et al.* (Joucla *et al.*, 2006) a montré une très forte altération dans la capacité à synthétiser des polymères de haute masse molaire. De même, en présence de saccharose et de maltose, ce mutant perd sa capacité à alterner les liaisons osidiques. Ce dernier semble donc incapable de reconnaître les unités glucosyle liées en $\alpha(1\rightarrow3)$ afin de les rallonger. Par conséquence, le mutant synthétise des oligodextranes composés exclusivement de liaisons $\alpha(1\rightarrow6)$ (Moulis *et al.*, 2006b). Les mêmes auteurs ont entrepris la construction des mutants S663Y, SEV663YDA, SEV663NNS et SEVQTVI663KGVQEKV de la dextrane-saccharase DSR-S afin de mimer la séquence des enzymes ASR, GTF-A et du second domaine catalytique de DSR-E. Les résultats montrent que tous les mutants ont perdu la capacité de produire du polymère. Ces mutations affectent également la capacité de transglucosylation vers l'accepteur maltose.

En utilisant une approche d'ingénierie semi-rationnelle, les modifications des résidus R624 et V630 de la glucansaccharase GTF-R de *Streptococcus oralis* ATCC10557 (motif III, R659 et V665 chez DSR-S respectivement), proches du stabilisateur du complexe covalent, contribuent à un changement radical de la spécificité de l'enzyme, passant ainsi de la synthèse d'un dextrane à la synthèse d'un mutane, avec toutefois des perturbations dans l'activité et l'efficacité catalytique de l'enzyme. Les modifications du résidu S628, quant à elles, favorisent la synthèse d'oligosaccharides et isomères de saccharose au détriment de la synthèse de polymère (Hellmuth *et al.*, 2008).

Funane *et al.* ont également remarqué que les résidus lysine du domaine catalytique pouvaient avoir un rôle dans la fixation du substrat et du dextrane en cours d'élongation. Ainsi, la mutation de la thréonine T350 de l'enzyme DSR-S en lysine a pour effet de favoriser la synthèse de liaisons $\alpha(1\rightarrow6)$. Il en est de même lorsque l'on mute la serine S455 en lysine. Quant à lui, le double mutant T350L:S455L synthétise un glucane composé à 86 % de liaisons $\alpha(1\rightarrow6)$, 8 % de liaisons $\alpha(1\rightarrow3)$ et 4 % de liaisons $\alpha(1\rightarrow2)$. C'est à ce jour le seul exemple d'ingénierie rationnelle aboutissant à la génération d'un mutant capable de synthétiser des liaisons de type $\alpha(1\rightarrow2)$.

La construction rationnelle de dextrane-saccharases chimères issues de DSR-S et DSR-T5 de *Ln. mesenteroides* NRRL B-512F, par l'échange de séquences peptidiques entre ces deux enzymes, a

permis d'obtenir des variants synthétisant des glucanes composés de liaisons glucosidiques $\alpha(1\rightarrow6)$, $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow4)$ à des teneurs variables (Funane *et al.*, 2004). Les auteurs proposent que la séquence peptidique comprise entre les résidus Y340 et N510 de la DSR-S, située en amont du tonneau $(\beta/\alpha)_8$, soit impliquée dans la formation de liaisons $\alpha(1\rightarrow4)$. De même, la région comprise entre les résidus N917 et K1131, positionnée dans la zone C-terminale du domaine catalytique, serait responsable du contrôle de la liaison synthétisée.

Enfin, Kralj *et al.* ont également construit des chimères à partir de variants tronqués des reuterane-saccharases GTF-A et GTF-O, toutes deux synthétisant des polymères formés de liaisons $\alpha(1\rightarrow4)$ et $\alpha(1\rightarrow6)$ mais avec des rapports différents (Kralj *et al.*, 2008). Les auteurs ont d'abord constaté que la partie C-terminale du domaine catalytique de ces enzymes, comprenant les éléments $\beta1$, $\alpha1$, $\beta2$, $\alpha2$, $\beta3$ du tonneau $(\beta/\alpha)_8$ et le GBD, n'a pas ou peu d'influence sur la spécificité des enzymes ciblées, ce qui diverge des résultats obtenus par Funane *et al.* précédemment cités. En revanche, la spécificité de liaison serait contrôlée par la région N-terminale du domaine catalytique (en amont du tonneau $(\beta/\alpha)_8$) et par la région comprenant les éléments $\alpha3$, $\beta4$, $\alpha4$, $\beta5$, $\alpha5$, $\beta6$, $\alpha6$ et $\beta7$.

Ces travaux sont résumés dans le tableau ci-après.

Tableau 2 : Ingénierie de la spécificité des glucane-saccharases de la famille GH70 : mutations et effets.

Enzyme	Mutant	Région	Effet de la mutation	Référence
GTF-B (<i>S. mutans</i> GS 5)	D567T	IV	Glucane soluble/Insoluble: 24/76%	(Shimamura <i>et al.</i> , 1994)
	T589D	IV	Glucane soluble/Insoluble: 15/85%	
GTF-D (<i>S. mutans</i> GS 5)	T589E	IV	Glucane soluble/Insoluble: 2/98%	(Monchois <i>et al.</i> , 2000b)
	D569T	IV	Augmentation de la production de glucane soluble	
GTF-A (<i>Lb. reuteri</i> 121)	Sauvage		46% $\alpha(1\rightarrow6)$, 54% $\alpha(1\rightarrow4)$	(Kralj <i>et al.</i> , 2005b; Kralj <i>et al.</i> , 2006)
	P1026V	II	augmentation des liaisons $\alpha(1\rightarrow6)$	
	P1026V/I1029V/A1066N/ N1134S/N1135E/S1136V	II et IV	93% $\alpha(1\rightarrow6)$, 7% $\alpha(1\rightarrow4)$	
	P1026V/I1029V/N1134S/ N1135E/S1136V	II et IV	92% $\alpha(1\rightarrow6)$, 8% $\alpha(1\rightarrow4)$	
	N1134S/N1135E/S1136V	IV	85% $\alpha(1\rightarrow6)$, 15% $\alpha(1\rightarrow4)$	
	N1134S	IV	82% $\alpha(1\rightarrow6)$, 18% $\alpha(1\rightarrow4)$	
	N1135E	IV	47% $\alpha(1\rightarrow6)$, 53% $\alpha(1\rightarrow4)$	
	S1136V	IV	43% $\alpha(1\rightarrow6)$, 57% $\alpha(1\rightarrow4)$	
	N1134S/S1136V	IV	70% $\alpha(1\rightarrow6)$, 30% $\alpha(1\rightarrow4)$	
	N1134S/N1135E	IV	75% $\alpha(1\rightarrow6)$, 25% $\alpha(1\rightarrow4)$	
	N1135E/S1136V	IV	42% $\alpha(1\rightarrow6)$, 58% $\alpha(1\rightarrow4)$	
	N1134D	IV	59% $\alpha(1\rightarrow6)$, 41% $\alpha(1\rightarrow4)$	
	N1134Q	IV	45% $\alpha(1\rightarrow6)$, 55% $\alpha(1\rightarrow4)$	
	N1134E	IV	48% $\alpha(1\rightarrow6)$, 52% $\alpha(1\rightarrow4)$	
	N1134G	IV	53% $\alpha(1\rightarrow6)$, 47% $\alpha(1\rightarrow4)$	
	N1134A	IV	64% $\alpha(1\rightarrow6)$, 36% $\alpha(1\rightarrow4)$	
	N1134Y	IV	45% $\alpha(1\rightarrow6)$, 55% $\alpha(1\rightarrow4)$	
	N1134H	IV	47% $\alpha(1\rightarrow6)$, 53% $\alpha(1\rightarrow4)$	
	GTF180 (<i>Lb. reuteri</i> 180)	Sauvage		
V1027P/S1137N/A1139S		IV	60% $\alpha(1\rightarrow6)$, 28% $\alpha(1\rightarrow3)$, 12% $\alpha(1\rightarrow4)$	
A1139S		IV	66% $\alpha(1\rightarrow6)$, 34% $\alpha(1\rightarrow3)$	
A1139L		IV	62% $\alpha(1\rightarrow6)$, 38% $\alpha(1\rightarrow3)$	
N1138D		IV	65% $\alpha(1\rightarrow6)$, 35% $\alpha(1\rightarrow3)$, <1% $\alpha(1\rightarrow4)$	
Q1140E		IV	68% $\alpha(1\rightarrow6)$, 29% $\alpha(1\rightarrow3)$, 3% $\alpha(1\rightarrow4)$	
Q1140A		IV	60% $\alpha(1\rightarrow6)$, 28% $\alpha(1\rightarrow3)$, 12% $\alpha(1\rightarrow4)$	
Q1140H		IV	86% $\alpha(1\rightarrow6)$, 14% $\alpha(1\rightarrow3)$	
S1137N/A1139V		IV	58% $\alpha(1\rightarrow6)$, 40% $\alpha(1\rightarrow3)$, 2% $\alpha(1\rightarrow4)$	
N1138E/A1139V		IV	65% $\alpha(1\rightarrow6)$, 34% $\alpha(1\rightarrow3)$, 1% $\alpha(1\rightarrow4)$	
S1137N		IV	61% $\alpha(1\rightarrow6)$, 35% $\alpha(1\rightarrow3)$, 4% $\alpha(1\rightarrow4)$	
S1137Y	IV	58% $\alpha(1\rightarrow6)$, 36% $\alpha(1\rightarrow3)$, 6% $\alpha(1\rightarrow4)$		
S1137N/A1139S	IV	59% $\alpha(1\rightarrow6)$, 37% $\alpha(1\rightarrow3)$, 4% $\alpha(1\rightarrow4)$		

	S1137Y/N1138D	IV	52% $\alpha(1\rightarrow6)$, 40% $\alpha(1\rightarrow3)$, 8% $\alpha(1\rightarrow4)$	
ASR (<i>Ln. mesenteroides</i> NRRL B1299)	Y768S/D769E/A770V	IV	Altération dans la capacité à alterner les liaisons	
	S663Y	IV		
DSR-S (<i>Ln. mesenteroides</i> NRRL B512F)	SEV663YDA	IV	Transglucosylation sur maltose affectée et perte de la capacité à produire un polymère	(Moulis <i>et al.</i> , 2006b)
	SEV663NNS	IV		
	SEVQTVI663KGVQEKV	IV		
	R624G:V630I:D717A	IV	synthèse de mutane	
	R624G	IV	augmentation des liaisons $\alpha(1\rightarrow3)$	
GTF-R (<i>S. oralis</i>)	V630I	IV	augmentation des liaisons $\alpha(1\rightarrow3)$	(Hellmuth <i>et al.</i> , 2008)
	D717A		Pas d'effet sur la spécificité	
	R624G/V630I	IV	synthèse de mutane	
	DSR-S STS1		glucane soluble/glucane insoluble : 95/5	
	DSR-S STS2		glucane soluble/glucane insoluble : 70/30	
DSR-S/DSR-T5 (<i>Ln. mesenteroides</i> NRRL B512-F)	DSR-S STS3		glucane soluble/glucane insoluble : 10/90	(Funane <i>et al.</i> , 2004)
	DSR-T5 TST1		glucane soluble/glucane insoluble : 30/70	
	DSR-T5 TST2		glucane soluble/glucane insoluble : 30/70	
	DSR-T5 TST3		glucane soluble/glucane insoluble : 95/5	
	T350K		94 % $\alpha(1\rightarrow6)$, 6 % $\alpha(1\rightarrow3)$	
DSR-S (<i>Ln. mesenteroides</i> NRRL B512-F)	S455K		93 % $\alpha(1\rightarrow6)$, 7 % $\alpha(1\rightarrow3)$	(Funane <i>et al.</i> , 2005)
	T350K/S455K		86 % $\alpha(1\rightarrow6)$, 8 % $\alpha(1\rightarrow3)$, 4 % $\alpha(1\rightarrow2)$	
	GTFA-O1-dN-RS (chimère)		38% $\alpha(1\rightarrow6)$, 62% $\alpha(1\rightarrow4)$	
	GTFA-O2-dN-RS (chimère)	II,III and IV	41% $\alpha(1\rightarrow6)$, 59% $\alpha(1\rightarrow4)$	
	GTFA-O3-dN-RS (chimère)		45% $\alpha(1\rightarrow6)$, 55% $\alpha(1\rightarrow4)$	(Kralj <i>et al.</i> , 2008)
GTF-A/GTF-O (<i>Lb. reuteri</i> 121 and <i>Lb. reuteri</i> ATCC 55730, respectivement)	GTFO-A1-dN-RS (chimère)		46% $\alpha(1\rightarrow6)$, 54% $\alpha(1\rightarrow4)$	
	GTFO-A2-dN-RS (chimère)	II,III and IV	37% $\alpha(1\rightarrow6)$, 63% $\alpha(1\rightarrow4)$	
	GTFO-A3-dN-RS (chimère)		26% $\alpha(1\rightarrow6)$, 74% $\alpha(1\rightarrow4)$	

Depuis la création de la classification CAZy en 1991, regroupant les enzymes agissant sur les sucres, 110 glucane-saccharases de la famille GH70 ont été recensées. La diversité de spécificités de ces biocatalyseurs en fait de formidables outils pour la valorisation du saccharose.

Des avancées majeures ont été réalisées ces dernières années pour mieux comprendre le mode d'action de ces enzymes et les facteurs régissant leur spécificité. Ainsi, la conservation entre les enzymes de la famille GH70 et GH13, des acides aminés « clés » impliqués dans la liaison au substrat et dans la catalyse, au niveau du sous-site -1 du domaine catalytique est en accord avec le fait que ces deux familles partagent un même mécanisme de rétention de la configuration anomérique du donneur. De même, la caractérisation biochimique de la dextrane-saccharase DSR-S a permis de mettre en évidence un mécanisme de polymérisation semi-processif avec élongation du polymère par l'extrémité non-réductrice, nécessitant un seul site catalytique (Moulis *et al.*, 2006b). Ce mécanisme a récemment été entériné par la résolution de la première structure d'une glucane-saccharase de la famille GH70, la GTF-180- Δ N (Vujicic-Zagar *et al.*, 2010).

De nombreux travaux d'ingénierie rationnelle et semi-rationnelle réalisés sur les glucane-saccharases de la famille GH70 ont porté sur l'identification de déterminants de la spécificité de ces enzymes. En l'absence de données cristallographiques, les recherches ont ciblé les motifs conservés entre les glucane-saccharases de la famille GH70 et GH13. Les mutations opérées au niveau du motif IV, en particulier, indiquent que ces régions seraient les plus impliquées dans la spécificité de ces enzymes. Cependant, d'autres régions moins conservées semblent jouer un rôle important dans la spécificité (Funane *et al.*, 2005; Kralj *et al.*, 2008). La résolution récente de la structure cristallographique de trois enzymes de la famille GH70, de spécificités différentes, devrait permettre d'approfondir les connaissances des relations structure/fonction des glucane-saccharases ainsi que des facteurs à l'origine de leur spécificité de liaison.

Partie 2 : Ingénierie d'enzymes

II. INGENIERIE D'ENZYMES

L'ingénierie d'enzymes est devenue, au cours de la dernière décennie, un outil puissant pour le modelage des propriétés des protéines à des fins industrielles ou pour la compréhension des relations structure/fonction d'un biocatalyseur. Deux approches ont été développées : l'ingénierie rationnelle et l'ingénierie combinatoire. L'approche rationnelle requiert une profonde connaissance des caractéristiques structurales du biocatalyseur étudié et de leurs contributions à la fonction de l'enzyme. Cette approche est souvent limitée par le manque de structure tridimensionnelle, et l'effet de certaines mutations proches ou éloignées du site actif pouvant être bénéfiques est difficile voire impossible à prédire. C'est pour palier ce problème que l'approche aléatoire (ou évolution dirigée) a été développée. Elle consiste à générer, dans un premier temps, de la diversité moléculaire par le biais de mutations au sein du gène codant pour la protéine d'intérêt et, dans un second temps, à trier les protéines mutées de manière à isoler le biocatalyseur possédant la(les) propriété(s) désirée(s). Parce que le nombre de mutants à trier peut-être important, la principale limitation de l'approche combinatoire est la nécessité de disposer de méthodes de sélection et/ou criblage à haut-débit et sensibles.

Ce chapitre est dédié à l'ingénierie des protéines par évolution dirigée. Il décrit de manière non exhaustive, à l'aide d'exemples, les méthodes employées pour générer et trier la diversité.

II.1 Ingénierie par évolution dirigée : vue générale

II.1.1 Principe

L'approche d'évolution dirigée vise à modifier les propriétés des enzymes en mimant les principes de l'évolution naturelle, basée sur l'itération de mutations génétique et de sélection phénotypique, mais de façon accélérée et *in vitro*.

Au laboratoire l'évolution dirigée consiste à générer un éventail de mutants enzymatiques, par mutagenèse et/ou recombinaisons aléatoires et isoler le ou les mutants ayant des propriétés améliorées par rapport à (aux) l'enzyme(s) parentale(aux) (figure 11).

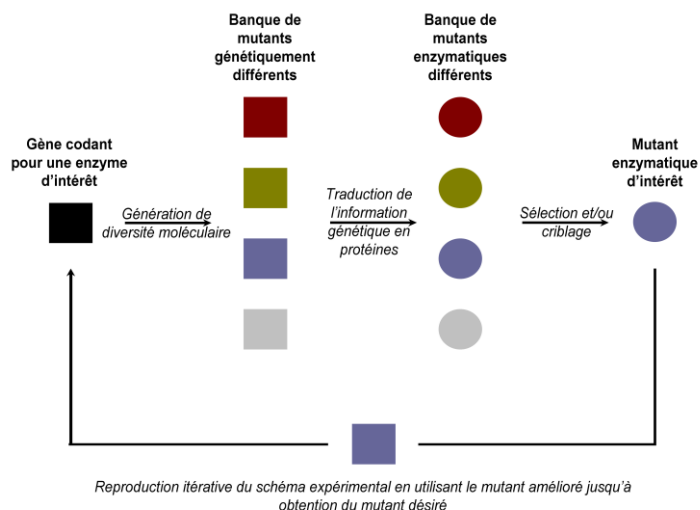


Figure 11 : Principe de l'évolution dirigée.

II.1.2 Générer la diversité par mutagenèse aléatoire

Comme son nom l'indique, la mutagenèse aléatoire consiste à générer une ou des mutations au sein d'un gène codant pour l'enzyme d'intérêt de façon aléatoire. La technique la plus courante est la PCR à erreur ou « **error-prone PCR** » (epPCR). L'epPCR consiste à utiliser la technique classique de PCR dans des conditions réduisant la fidélité de l'ADN polymérase. L'enzyme la plus utilisée est la DNA polymérase I de *Thermus aquaticus* (*Taq* polymérase) dont le taux d'erreurs varie de 10^{-4} (1 erreur pour 10 000 bases) à 10^{-6} (1 erreur pour 1 million de bases) en fonction des conditions réactionnelles (Cline *et al.*, 1996; Eckert and Kunkel, 1990; Tindall and Kunkel, 1988). Le taux de mutations peut être augmenté en jouant sur différents paramètres comme la concentration en *Taq* polymérase, l'augmentation du temps de polymérisation (Cadwell and Joyce, 1992), le remplacement de Mg^{2+} par du Mn^{2+} (Lin-Goerke *et al.*, 1997), la variation des ratios en nucléotides (Cadwell and Joyce, 1992; Fromant *et al.*, 1995) ou l'addition d'analogues de nucléotides (Spee *et al.*, 1993; Zaccolo *et al.*, 1996).

Malgré sa simplicité, l'epPCR souffre de plusieurs inconvénients qui limitent la qualité des banques générées. Il s'agit des biais mutationnels engendrés par la polymérase employée (favorisant les substitutions AT→GC et AT→TA) et l'impossibilité de muter deux nucléotides successifs dans un codon (Wong *et al.*, 2006).

La société française MilleGen™ a breveté une technique appelée MutaGen™ reposant sur l'utilisation des polymérases infidèles humaines Pol β et Pol η (Bouayadi *et al.*, 2002). Cette technique a récemment été utilisée pour l'obtention de variants thermostables de l'amylosaccharase de *Neisseria polysaccharea* (Emond *et al.*, 2008a; Emond *et al.*, 2008c). Les auteurs ont ainsi isolé un double mutant

10 fois plus thermostable que l'enzyme sauvage. Ainsi, ces polymérases génèrent des taux de mutation et des biais mutationnels différents de ceux de la *Taq* polymérase, ce qui permet de faire varier le spectre mutationnel lors de la création des banques de variants et ainsi de générer plus de diversité. Cependant, cette technique ne permet pas non plus d'introduire des mutations adjacentes. Récemment, les méthodes de mutagenèse aléatoire TriNEx (TriNucleotide Exchange) (Baldwin *et al.*, 2008), TIM (Transposon Integration mediated Mutagenesis) (Hoeller *et al.*, 2008) et SeSaM-Tv+ (transversion-enriched Sequence Saturation Mutagenesis) (Wong *et al.*, 2008) ont été développées afin de palier ce problème, parfois au détriment de la simplicité et de la robustesse.

II.1.3 Générer la diversité par recombinaison aléatoire

La recombinaison aléatoire repose sur une redistribution de l'ADN issu de gènes parentaux afin de générer une banque de variants hybrides. La recombinaison est dite « homologue » quand les gènes à recombiner présentent un taux d'identité fort, généralement supérieur à 60 %. C'est le cas lorsque les gènes parentaux sont issus de gènes codant pour des protéines homologues ou d'une banque de variants obtenus par mutagenèse aléatoire. On parle alors de « **family shuffling** » et « **DNA shuffling** » respectivement (Cramer *et al.*, 1998; Stemmer, 1994a; Stemmer, 1994b). Sur le plan technique, ces deux approches sont fondées sur le même principe. Les gènes à recombiner sont dans un premier temps fractionnés de façon aléatoire par la DNaseI. Les fragments obtenus sont alors réassemblés lors d'une PCR sans amorces engendrant la formation d'hétéroduplexes, et amplifiés jusqu'à obtention de fragments de la même taille que les gènes parentaux (figure 12).

Tout comme l'epPCR, la technique de DNA shuffling est assez simple à mettre en œuvre. Cependant elle présente certaines limites. D'une part, l'emploi de la PCR rend difficile l'utilisation de ces techniques sur des gènes longs (> 3kb). D'autre part, le nombre de recombinaisons reste faible (1 à 4 par séquence) du fait de la génération d'hétéroduplexes nécessitant la présence de régions de fortes identités, constituant les points de recombinaisons (Joern *et al.*, 2002).

Comparé au DNA shuffling, le family shuffling présente l'intérêt de tirer profit de la diversité structurale naturelle de protéines de même fonction. La technique a été appliquée au brassage de quatre gènes codant pour des cephalosporinases de classe C. partageant 60 à 82 % d'homologie. Ceci a abouti à l'identification d'un clone présentant une résistance au moxalactame 270 à 540 fois supérieure à celle des protéines parentales (Cramer *et al.*, 1998).

Depuis la description de ces techniques au début des années 1990, de multiples méthodes alternatives ont été proposées. Ce sont soit des améliorations apportées aux méthodes classiques comme la substitution de la DNaseI par des enzymes de restriction (Kikuchi *et al.*, 1999) ou l'utilisation d'ADN

simple brin (Kikuchi *et al.*, 2000; Zha *et al.*, 2003b) pour faciliter la formation d'hétéroduplexes, soit de nouvelles techniques comme les méthodes RPR « Random Priming Recombinaison » (Shao *et al.*, 1998), StEP (Staggered Extension Process) (Zhao *et al.*, 1998; Zhao and Zha, 2006), RACHITT (Random CHimeragenesis on Transient Template) (Coco *et al.*, 2001) ou L-Shuffling breveté par la société Proteus (Dupret *et al.*, 2002) qui permettent d'augmenter le taux de recombinaisons.



Figure 12 : Techniques de DNA et family shuffling (Cramer *et al.*, 1998; Stemmer, 1994a).

Lorsque le taux d'identité entre les séquences est faible (< 50%), la génération d'évènements de recombinaison est difficile. Pour palier cet obstacle, il est possible d'avoir recours à des méthodes de recombinaison dites « non homologues ». C'est le cas des techniques ITCHY (Incremental Truncation for the Creation of Hybrid enZYmes) (Ostermeier *et al.*, 1999) et SHIPREC (Sequence-Homology Independent RECombinaison) (Sieber *et al.*, 2001) (figure 13). Avec une même approche mais des stratégies différentes, les produits obtenus par ces deux méthodes sont les mêmes, des banques de mutants hybrides arborant un seul point de croisement. Même si le potentiel de ces techniques a été mis en évidence par l'identification de variants hybrides (points de croisement situés au niveau de régions de faible identité) fonctionnels, la génération d'un unique point de croisement limite l'exploration de l'espace de séquence théorique. En combinant l'approche ITCHY et le DNA shuffling, Lutz *et al.* ont

défini la méthode SCRATCHY permettant d'accroître le taux de croisement d'un facteur 2 à 3 pour le brassage de deux gènes présentant moins de 50 % d'identité (Lutz *et al.*, 2001).

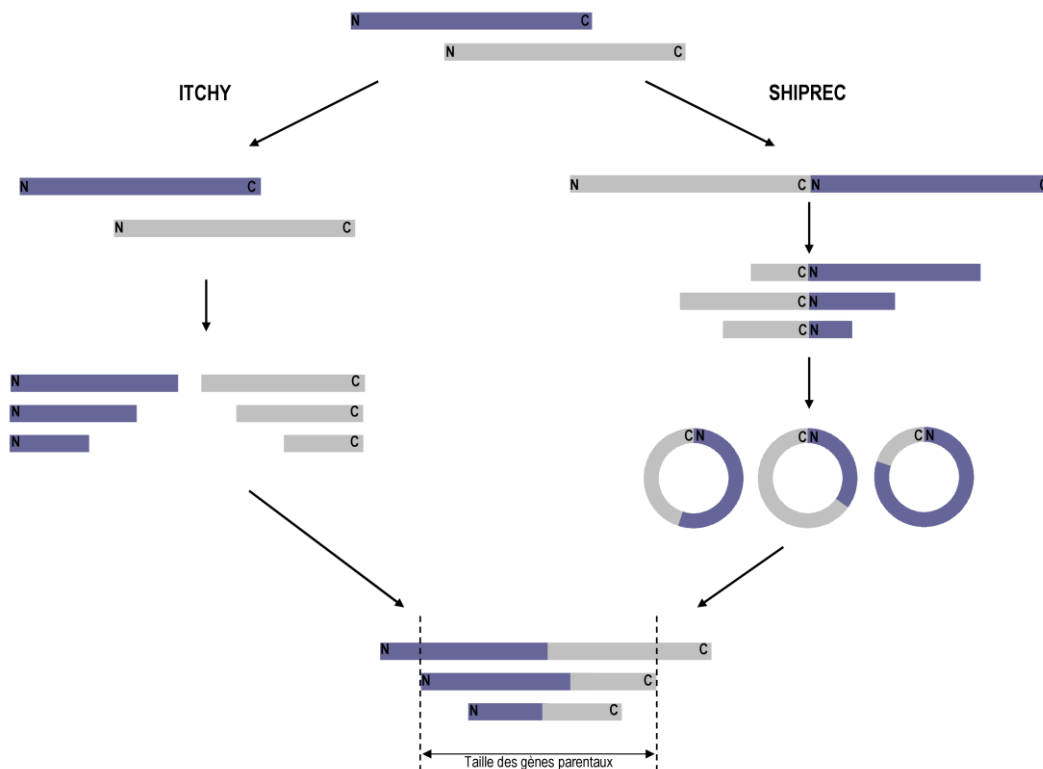


Figure 13 : Techniques de recombinaison non homologue ITCHY et SHIPREC (Ostermeier *et al.*, 1999; Sieber *et al.*, 2001). La première étape de la méthode ITCHY consiste à tronquer indépendamment les deux gènes parentaux en N- et C-terminal respectivement. Les fragments sont ensuite ligués pour aboutir à un mélange de séquences hybrides. La méthode SHIPREC commence par la fusion des deux gènes parentaux suivie d'une digestion contrôlée par la DNaseI. Les fragments sont ensuite circularisés et digérés au niveau du point de fusion originel aboutissant à une inversion de séquence générant la banque de mutants hybrides. Dans les deux méthodes, une dernière étape consiste à purifier les fragments dont la taille est identique à celle des gènes parents.

II.1.4 Rationaliser l'aléatoire : mutagenèse et recombinaison semi-aléatoire

La conception semi-aléatoire (ou semi-rationnelle) a pour but de combiner les connaissances des relations structure/fonction d'un biocatalyseur avec les techniques d'évolution dirigée (création des banques, sélection, criblage). L'approche semi-aléatoire se focalise donc sur le brassage d'un ou de plusieurs résidus ciblés, susceptibles d'avoir un impact sur la propriété recherchée.

II.1.4.A Mutagenèse de saturation

Parmi les différentes approches semi-aléatoires, la plus répandue est la mutagenèse de saturation. Elle consiste à introduire des mutations à l'aide d'oligonucléotides synthétiques, contenant un ou plusieurs

codons dégénérés ciblant la ou les positions choisies au préalable. L'incorporation est réalisée au cours d'une réaction de PCR en utilisant soit le gène d'intérêt, soit le plasmide portant le gène d'intérêt, comme matrice. Cette approche peut être utilisée pour sonder chaque résidu, de manière individuelle, et identifier celui ou ceux impliqués dans la propriété à modifier (DeSantis *et al.*, 2003b) ou bien pour muter plusieurs résidus simultanément (Sylvestre, 2010).

L'emploi d'oligonucléotides dégénérés permet d'introduire toutes les mutations possibles à une position donnée. Ainsi, par l'utilisation d'un codon dégénéré de type NNN (N étant le code de dégénérescence désignant les quatre nucléotides A, C, G ou T), il est possible d'explorer toute la diversité apportée par l'ensemble des vingt acides aminés (Oliphant *et al.*, 1986). Toutefois, le mélange équimolaire des quatre nucléotides entraîne un biais en faveur des acides aminés qui ont le plus d'entrées dans le code génétique (acides aminés les plus représentés tel la leucine, l'arginine ou la sérine) et génère également beaucoup de codons stop. Des codons dégénérés alternatifs tel NNS ou NNK (S = G ou C ; K = G ou T), dont la troisième base est partiellement dégénérée, permettent d'équilibrer la représentation des 20 acides aminés (32 codons possibles) tout en limitant la présence de codons stop (1 codon). Le choix du code de dégénérescence s'avère crucial lorsque l'on souhaite muter plusieurs positions simultanément. En effet, couvrir 95 % de l'espace de séquences généré par la mutation d'une position avec la dégénérescence NNK requiert le criblage de 94 variants (Reetz *et al.*, 2008b), ce qui ne nécessite pas de crible haut-débit. En revanche lorsque le nombre de positions à muter passe à 2, 3 ou 4, le nombre de variants à cribler pour couvrir 95 % de l'espace de séquence augmente de façon vertigineuse et passe à 3 066, 98 163 et 314 1251 respectivement. Cela demande donc une capacité de criblage suffisante pour explorer toute la diversité générée, ce qui n'est pas toujours possible. Reetz *et al.* proposent alors de contourner ce problème en utilisant un code de dégénérescence qui couvre l'ensemble des structures et propriétés des acides aminés tel le codon dégénéré NDT (D = A, G ou T ; 12 acides aminés, pas de codon stop). Ceci permet de réduire considérablement la taille de la banque à cribler (34, 403, 5 175 et 62 118 variants pour 1, 2, 3 et 4 positions respectivement). En appliquant cette stratégie à l'ingénierie de l'énantiosélectivité de l'époxide hydrolase d'*Aspergillus niger* vis-à-vis d'un substrat non reconnu par l'enzyme sauvage (3 résidus ciblés), Reetz *et al.* ont obtenu 511 mutants positifs après criblage de 5000 variants issus d'une banque NDT contre 38 issus d'une banque NNK, soit 13 fois plus (Reetz *et al.*, 2008b).

II.1.4.B Mutagenèse de saturation couplée à la recombinaison

La mutagenèse de saturation peut être couplée à des techniques de recombinaisons. C'est le cas des méthodes DOGS (Degenerate Oligonucleotide Gene Shuffling) (Gibbs *et al.*, 2001), ADO (Assembly of

Designed Oligonucléotides) (Zha *et al.*, 2003a) et de shuffling synthétique (Ness *et al.*, 2002). Ces méthodes font intervenir l'emploi massif de longs oligonucléotides dégénérés (60 à 140 pb) dont le niveau de pureté influe directement sur les banques produites. Sur ce constat, Herman *et al.* ont développé une méthode appelée ISOR (Incorporating Synthetic Oligonucleotides via Gene Reassembly) (figure 14) permettant l'incorporation, à n'importe quelle position, d'oligonucléotides dégénérés courts (20 à 30 pb) codant pour des substitutions, insertions ou délétions (Herman and Tawfik, 2007). Cette méthode est à la fois puissante et flexible car elle permet de cibler des résidus aux positions adjacentes ou répartis sur l'ensemble du gène d'intérêt et de contrôler la fréquence des mutations en jouant sur les concentrations des oligonucléotides. Sa mise en œuvre pour la diversification de 45 positions d'un gène codant pour une ADN méthyltransférase a abouti à la création d'une série de banques de gènes présentant en moyenne une à six mutations, en fonction de la concentration en oligonucléotides dégénérés, couvrant l'ensemble des positions ciblées (Herman and Tawfik, 2007).

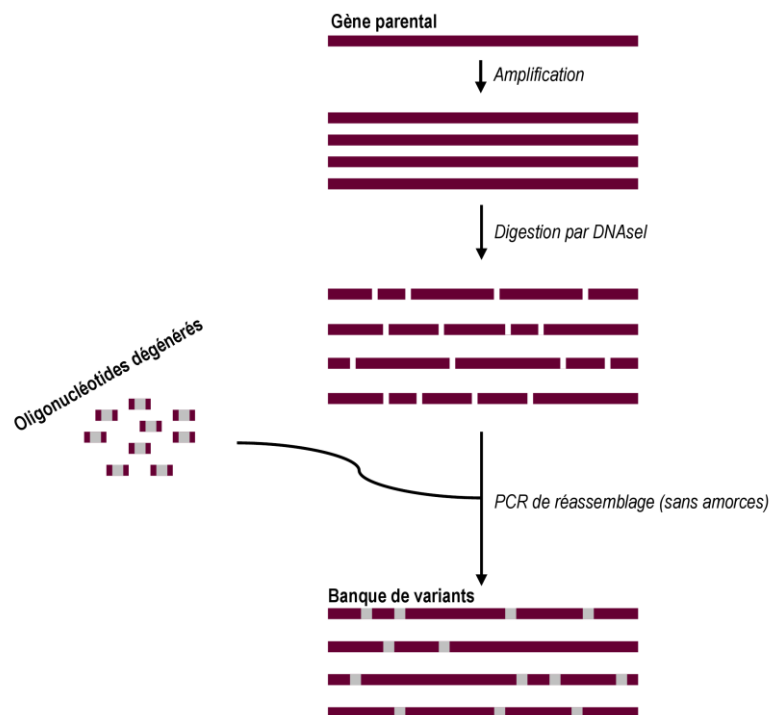


Figure 14 : Technique d'incorporation d'oligonucléotides dégénérés par réassemblage de gène ISOR (Herman and Tawfik, 2007). Dans un premier temps le gène codant pour la protéine d'intérêt est amplifié afin d'augmenter le nombre de copies. Cette étape est suivie d'une digestion par la DNaseI générant divers fragments couvrant l'ensemble du gène initial. Les fragments sont ensuite mélangés aux oligonucléotides dégénérés (la dégénérescence est portée au centre de l'oligonucléotide, flanquée de part et d'autre de séquences correspondant au gène parent). Pour finir, les oligonucléotides dégénérés sont incorporés lors du réassemblage des fragments grâce à une dernière étape de PCR sans amorces.

II.1.5 Complémentarité des approches rationnelles et aléatoires

L'approche rationnelle requiert une parfaite connaissance des relations structure/fonction de la protéine étudiée. Cette démarche est bien adaptée à l'ingénierie de fonctions en liens avec le site actif (modification de l'énantiosélectivité, la spécificité de substrat, l'efficacité). En revanche cette approche est peu appropriée lorsque l'on souhaite adapter une enzyme à des températures, pH ou solvants inhabituels. Cela est dû à la complexité des arrangements polypeptidiques assurant le maintien de la structure moléculaire. Dans ce dernier cas, la mutagenèse et/ou la recombinaison aléatoire restent des approches de choix. De plus, si les approches d'ingénierie rationnelle sont très efficaces, il arrive néanmoins que certaines mutations induisent le contraire des effets recherchés. Ceci est dû à une modification de la structure induite par la ou les mutations qui ne peuvent être prédites par modélisation moléculaire à partir de la structure de l'enzyme sauvage. Il en est de même lorsque l'on brasse par recombinaison aléatoire des fragments de peptides provenant de protéines parentales dont l'identité entre les séquences est faible. L'algorithme SCHEMA (Meyer *et al.*, 2003; Voigt *et al.*, 2002) permet de réduire la probabilité de déstabilisation des protéines chimériques par prédiction des points de recombinaisons favorables au bon repliement des biocatalyseurs hybrides. SCHEMA utilise les données structurales pour définir des blocs de séquences (régions) pouvant être échangés entre protéines parentales tout en limitant les perturbations susceptibles d'engendrer une rupture des contacts entre les chaînes latérales des acides aminés. Cet algorithme a été utilisé dernièrement pour la génération de cellobiohydrolases chimères thermorésistantes par le brassage de trois cellobiohydrolases fongiques (Heinzelman *et al.*, 2009). Des 6 500 combinaisons possibles, 48 chimères ont été générées, 23 se sont révélées actives parmi lesquelles cinq présentant une meilleure thermostabilité que celles des enzymes sauvages. Un modèle mathématique développé pour évaluer les effets individuels de chaque bloc sur la thermostabilité a entraîné la construction de 25 nouvelles chimères. Dix de ces dernières ont été produites sous forme active et ont montré, là encore, une thermostabilité supérieure à celles des enzymes sauvages.

II.2. Trier la diversité : Sélection et criblage

L'évolution dirigée, ou ingénierie combinatoire, est basée sur deux étapes distinctes et indissociables : (1) la génération de la diversité moléculaire et (2) le tri de la diversité. Cette deuxième étape nécessite de disposer d'une méthode capable d'isoler rapidement parmi l'ensemble des variants générés celui ou ceux exhibant la propriété recherchée. Si les méthodes permettant de créer la diversité sont génériques, en revanche les approches de sélection ou de criblage nécessitent d'être développées pour

chaque enzyme et chaque propriété à optimiser (Aharoni *et al.*, 2005b). Le point clé dans la conception d'enzymes optimisées consiste donc à développer des méthodes efficaces de sélection et/ou de criblage à haut-débit, sensibles (afin d'isoler des clones présentant une faible activité), reproductibles (pour discerner de faibles améliorations) et robustes (pour recueillir des informations dépendantes du clone lui-même et non de variables environnementales). Le développement de toute méthode de criblage ou de sélection est basé sur le lien existant entre le gène, l'enzyme pour lequel il code et son activité.

La présentation faite ci-après recense les méthodes haut-débit appliquées à la sélection et/ou criblage de larges banques de variants.

II.2.1 Techniques de sélection *in vivo*

La sélection *in vivo* relie la survie cellulaire à l'activité enzymatique recherchée. Le principe est basé sur le fait que l'activité de l'enzyme d'intérêt confère à la cellule hôte un phénotype lui permettant de croître dans des conditions de cultures données. Cette stratégie est simple et permet de trier un grand nombre de variants en favorisant la croissance de clones exprimant les gènes d'intérêt tout en éliminant ceux qui ne sont pas désirés. Deux approches sont possibles, la complémentation génétique et la complémentation chimique (figure 15). La première étape commune à ces deux approches consiste à transformer les cellules hôtes par la banque plasmidique. La plupart des cas recensés reposent sur l'utilisation d'un microorganisme bien caractérisé, comme *Escherichia coli*.

La complémentation génétique est d'une grande efficacité dans de nombreux cas portant sur l'évolution dirigée d'enzymes, en vue d'améliorer l'activité catalytique (Parikh *et al.*, 2006; Van Der Veen *et al.*, 2004), modifier la spécificité (DeSantis *et al.*, 2003a; Griffiths *et al.*, 2004), modifier l'énantioselectivité (Boersma *et al.*, 2008; Reetz *et al.*, 2008a), améliorer l'expression (Ju *et al.*, 2005). Cependant, la complémentation génétique est limitée à des stratégies permettant de combler une auxotrophie, de résister à un agent toxique ou de convertir un substrat non métabolisable en nutriment pour la cellule hôte. La complémentation chimique constitue une alternative. Dans cette approche la complémentation est liée non pas à la réaction catalysée en elle-même mais au produit formé dont la synthèse aura pour effet d'activer ou reconstituer un facteur de transcription favorisant l'expression d'un gène essentiel à la survie et la croissance de la cellule. Une méthode de complémentation chimique a récemment été développée chez *E. coli* (Van Sint Fiet *et al.*, 2006).

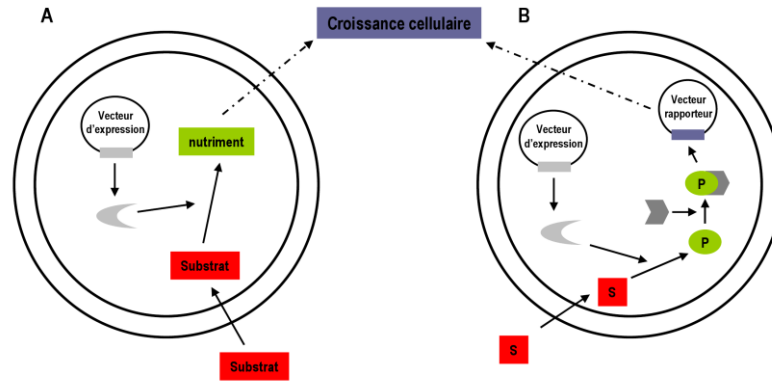


Figure 15 : Stratégie générale de sélection par complémentation. (A) Sélection par complémentation génétique sur substrat non métabolisable. Dans la bactérie, l'enzyme d'intérêt convertit le substrat non métabolisable en substrat métabolisable lui fournissant ainsi l'énergie nécessaire pour croître. (B) Sélection par complémentation chimique. L'enzyme d'intérêt convertit le substrat en produit qui forme un complexe actif avec un facteur de transcription favorisant l'expression d'un gène essentiel pour la croissance bactérienne (gène de résistance par exemple) porté par un vecteur rapporteur. Adapté de (van Sin Fiet *et al.*, 2006).

II.2.2 Techniques de sélection *in vitro*

II.2.2.A La technologie du « display »

La technologie du « display » est un outil puissant permettant l'exploration des interactions entre protéines et ligands. Elle repose sur le lien physique entre génotype et phénotype grâce à l'exposition des variants protéiques à la surface d'un phage (phage-display), d'une cellule (cell-surface display) ou par formation d'un complexe ARNm-ribosome-protéine (ribosome display).

Les premières méthodes développées mettant en jeu du phage-display pour la sélection de variants enzymatiques actifs sont basées sur l'affinité de l'enzyme pour un analogue de substrat, substrat suicide ou analogue de l'état de transition (Bai and Feng, 2004; Fujii *et al.*, 1998; Soumillion *et al.*, 1994; Vanwetswinkel *et al.*, 2000) (figure 16A). Ces méthodes, qualifiées d'indirectes, sont adaptées pour la recherche de variants stables ou de spécificités modifiées. D'autres stratégies permettent la sélection de variants enzymatiques sur la base de l'activité catalytique. Ceci nécessite de créer un lien physique entre le phage et le substrat afin d'éviter la dispersion du produit de la réaction. Le variant enzymatique actif transforme ainsi le substrat lié au phage en un produit qui reste attaché au phage. Le produit est alors utilisé pour isoler le variant d'intérêt en provoquant soit l'immobilisation du phage sur un support solide (figure 16B) soit son relargage (figure 16C). Le phage display est un outil rapide pour la sélection de banques de variants lors de campagnes d'évolution dirigée.

La technologie du « cell-surface display » a également gagné en rapidité notamment grâce au développement du tri-cellulaire par cytométrie en flux (FACS). Le FACS est une méthode à haut-débit

qui permet d'identifier, par marquage fluorescent du substrat, un grand nombre de cellules en des temps très courts. Cette technique a été utilisée dans divers travaux d'évolution dirigée (Aharoni *et al.*, 2005a; Aharoni *et al.*, 2006; Bernath *et al.*, 2004; Griffiths and Tawfik, 2003; Santoro *et al.*, 2002) et peut être combinée avec la technologie de compartimentation *in vitro*. Faire un tri sur l'activité enzymatique par cell-surface display nécessite là encore de capturer le produit de la réaction à la surface de la cellule. L'utilisation de bactéries Gram négative dont la surface est naturellement chargée négativement permet l'ancrage d'un substrat fluorescent conjugué à une queue polycationique comme démontrée pour la sélection de variants de la protéase OmpA de *E. coli* (Olsen *et al.*, 2000). La transformation du substrat en produit entraîne un changement de la fluorescence ce qui permet de distinguer les variants actifs des autres. Une autre approche consiste à immobiliser la cellule dans des gouttes de polyacrylamide fixées par la suite sur support solide. Ceci est réalisé de sorte qu'une goutte contienne en moyenne une seule cellule. Les microcolonies sont ensuite équilibrées avec un substrat chromogène ou fluorogène et sélectionnées à l'aide d'un microscope. La méthode a permis l'isolement de souches d'*E. coli* exprimant une β -lactamase en surface à partir d'un mélange de bactéries n'exprimant pas l'enzyme en excès (Freeman *et al.*, 2004). La levure est également un système exploité dans le cell-surface display (Colby *et al.*, 2004; Lipovsek *et al.*, 2007).

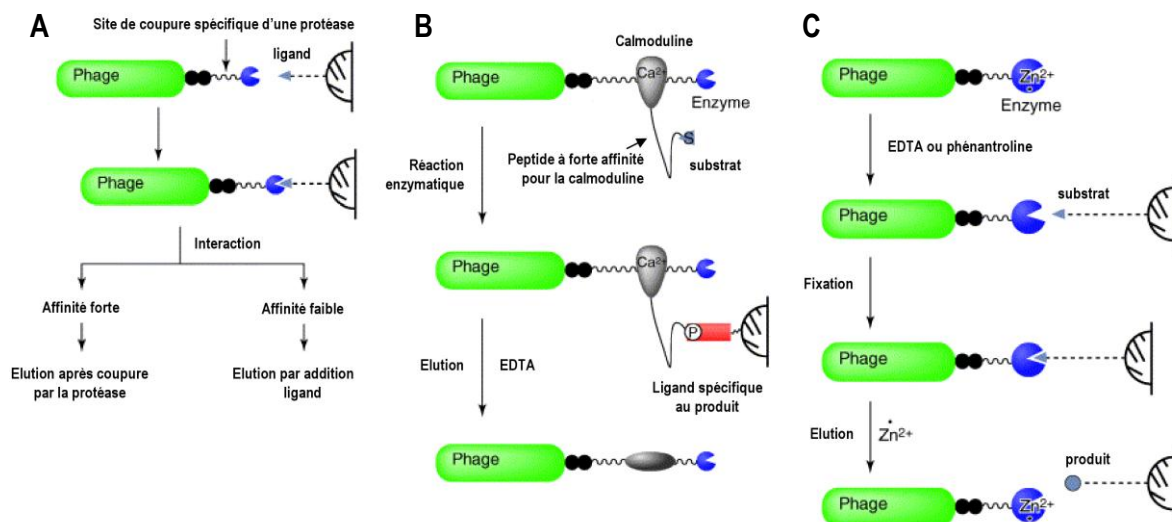


Figure 16 : Stratégie de sélection par phage display adapté de (Fernandez-Gacio *et al.*, 2003). (A) Sélection indirecte. Le variant enzymatique présenté à la surface du phage est sélectionné par affinité pour un ligand spécifique attaché sur un support solide. (B) Sélection directe. Le substrat est lié au complexe enzyme-calmoduline-phage via un peptide de liaison à la calmoduline. Après catalyse, le produit se fixe par affinité à un ligand attaché sur support solide, d'après (Demartis *et al.*, 1999). (C) Sélection directe de métalloenzyme par phage-display. L'enzyme est dans un premier temps inactivée par extraction du co-facteur métallique Zn^{2+} et se lie avec le substrat attaché sur support solide. L'ajout du co-facteur entraîne la conversion du substrat en produit et conduit à l'éluion du phage, d'après (Ponsard *et al.*, 2001).

La méthode de « ribosome display » est quant à elle réalisée *in vitro*. La formation du complexe ARNm-ribosome-protéine est réalisée lors de la traduction à l'aide d'extraits cellulaires procaryotes ou eucaryotes. La génération d'un tel complexe est possible grâce à l'absence de codon stop sur l'ARNm, bloquant ainsi la séparation entre la protéine et le ribosome. L'emploi d'un peptide espaceur en fusion C-terminal avec la protéine d'intérêt permet à cette dernière d'adopter un bon repliement fonctionnel. Cette technique a été utilisée pour la sélection de fragments d'anticorps ou peptides à forte affinité pour un ligand (Boersma *et al.*, 2007) (figure 17). Il existe également des exemples d'utilisation du ribosome display pour isoler des variants sur la base de leur efficacité catalytique (Amstutz *et al.*, 2002; Takahashi *et al.*, 2002) mais ils restent assez rares car la technique requiert de fortes affinités pour le ligand ce qui n'est pas toujours le cas.

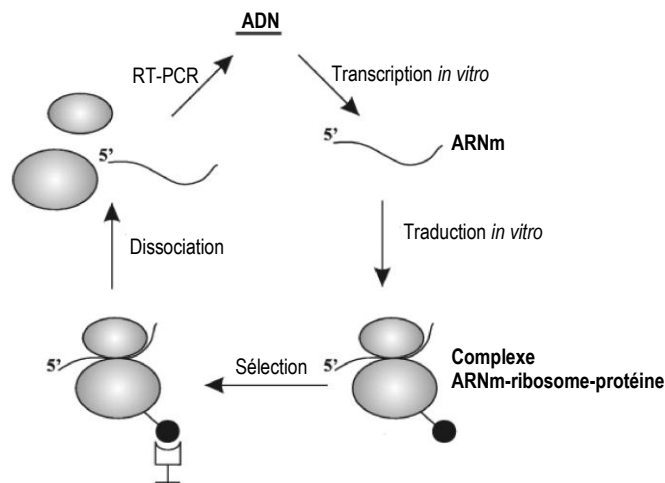


Figure 17 : Schéma général de la stratégie de sélection par ribosome display adapté de (Boersma *et al.*, 2007). Dans un système *in vitro*, l'ADN double brin codant pour la protéine d'intérêt est transcrit en ARNm. L'ARNm étant dépourvu de codon stop, le relargage du variant enzymatique lors de la traduction est alors empêché ce qui aboutit à la formation du complexe ARNm-ribosome-protéine. Une sélection est ensuite réalisée par capture du complexe par affinité de la protéine pour un ligand. Après dissociation du complexe, l'ARNm peut être amplifié par RT-PCR isolant ainsi la séquence codante pour le variant sélectionné.

II.2.2.B La compartimentation *in vitro*

Dans l'approche de compartimentation *in vitro* (IVC), les réactions sont confinées dans les gouttes d'une émulsion d'eau dans l'huile (Figure 18) (Squires and Quake, 2005). Des millions de compartiments d'environ 1 à 5 femtolitres sont produits en parallèle dans chaque émulsion ($>10^{10}$ dans 1 ml d'émulsion) (Aharoni *et al.*, 2005b; Griffiths and Tawfik, 2006). Chacun se comporte comme une cellule artificielle comprenant un seul gène codant pour la protéine d'intérêt et toute la machinerie permettant la synthèse de cette protéine. Ceci permet de sélectionner un très grand nombre de variants

dans un espace miniaturisé et dans un laps de temps court (5.10^7 réactions en 7 h) (Agresti *et al.*, 2010). Il est également possible, avec ce système, de moduler facilement les conditions de sélection (température, pH, etc.).

Une première approche utilisant l'IVC consiste à lier physiquement les gènes codants pour les variants enzymatiques à leur substrat. Dans le cas où la protéine variante est active, le substrat est transformé en un produit toujours attaché au gène qui peut être directement isolé en utilisant une méthode de détection du produit de la réaction (Cohen *et al.*, 2004; Tawfik and Griffiths, 1998). Une autre approche consiste à utiliser des substrats fluorophores et de coupler l'IVC à la cytométrie en flux. La synthèse du variant enzymatique et la réaction sont réalisées dans une première émulsion. Cette émulsion est ensuite convertie en une double émulsion d'eau dans huile et injectée dans un cytomètre afin d'isoler les émulsions fluorescentes (Mastrobattista *et al.*, 2005). Une méthode de double émulsion a également été décrite pour l'encapsulation de bactéries exprimant une banque de variants enzymatiques (Aharoni *et al.*, 2005a).

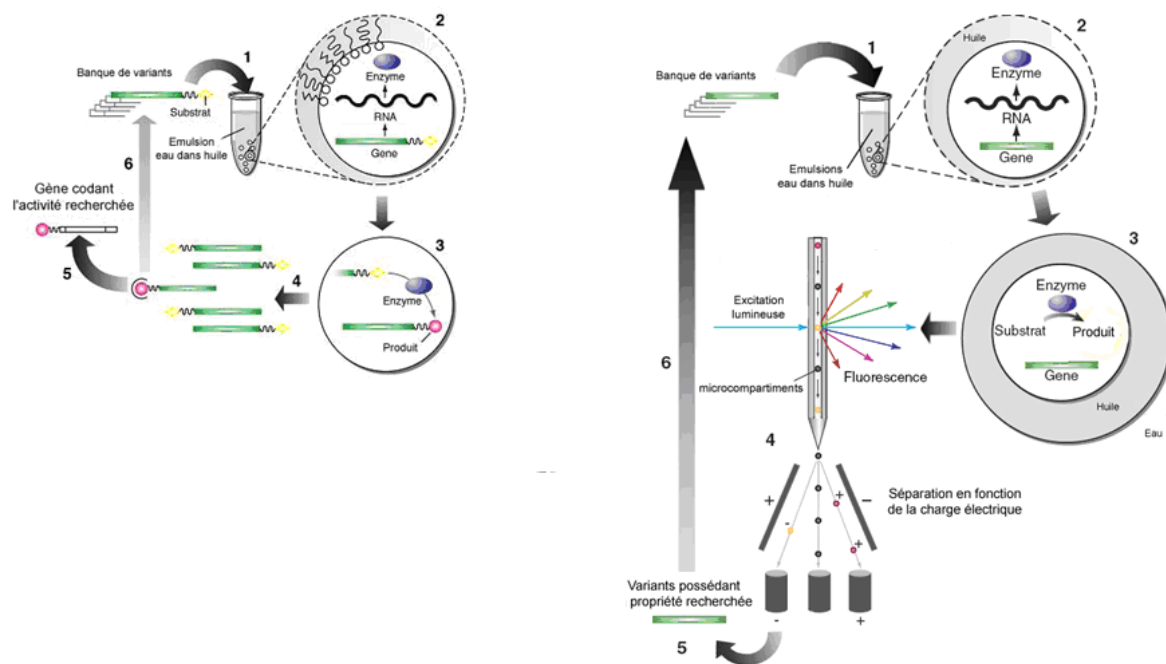


Figure 18 : Stratégies de sélection par compartimentation *in vitro*. (A) Couplage direct entre le gène et le produit de la réaction enzymatique utilisé pour la sélection de méthyltransférases (Cohen *et al.*, 2004). 1 : la banque de gènes et le mélange réactionnel permettant la transcription et traduction sont dispersés dans une émulsion eau dans huile (un gène par émulsion) ; 2 : chaque gène est transcrit et traduit ; 3 : l'enzyme active convertit le substrat en produit qui reste attaché au gène ; 4 : les émulsions sont cassées et les compartiments d'eau sont combinés ; 5 : les gènes codant pour des mutants enzymatiques actifs sont sélectionnés pour être caractérisés ou ; 6 : liés à un nouveau substrat pour d'autres tours de sélection (enrichissement de la banque de variants positifs). (B) Analyse par cytométrie de double émulsions employée pour la sélection de nouvelles β -galactosidases (Mastrobattista *et al.*, 2005). 1 : 1 : dispersion de la banque de gènes et le mélange réactionnel dans une émulsion eau dans huile ; 2 : transcription et traduction des gènes à l'intérieur de la microgouttelette ; 3 : transformation du substrat non fluorescent en produit fluorescent par les variants actifs et conversion de l'émulsion eau dans huile en double émulsion ; 4 : séparation des microgouttelettes fluorescentes des non-fluorescentes par FACS ; 5 : récupération des gènes codant pour une enzyme active et amplification par PCR ; 6 : enrichissement de la banque de variants positifs.

II.2.3 Criblage à haut-débit

Le criblage à haut-débit (high-throughput screening) repose sur la caractérisation individuelle et rapide de chaque variant enzymatique d'une banque. Ceci nécessite donc de miniaturiser et automatiser les procédures employées.

Lors d'un criblage, la production des variants enzymatiques est généralement réalisée de manière recombinante dans une cellule hôte. Le système le plus utilisé est *E. coli* chez les procaryotes et *S. cerevisiae* chez les eucaryotes. L'isolement des clones peut être ensuite réalisé en milieu gélosé lors de la croissance des colonies ou en milieu liquide, dans des mini-cultures au format microplaque.

Une grande majorité des protocoles développés repose sur l'utilisation de substrats ou capteurs chromophores ou fluorophores (Reymond and Babiak, 2007). Les clones exprimant un variant enzymatique d'intérêt sont alors repérés par apparition d'une coloration détectable, après transformation du produit, par simple analyse visuelle (coloration de la colonie, halo entourant une colonie) ou via l'utilisation d'un appareillage spécifique (lecteur de microplaque à fluorescence ou UV/VIS, etc.).

Le criblage en milieu gélosé peut être appliqué directement après transformation du système hôte. Il permet d'analyser, avec un protocole relativement simple et rapide, un grand nombre de variants à la fois (~400 000 clones/semaine). Cependant cette approche n'est pas quantitative. L'emploi de l'imagerie digitale permet de palier partiellement ce problème (Delagrave *et al.*, 2001; Joo *et al.*, 1999; Koné *et al.*, 2009). Le criblage en milieu gélosé peut également être utilisé comme première étape servant à éliminer de manière rapide les variants inactivés par les mutations générées. On parle alors de criblage primaire. C'est le cas de la méthode développée récemment par Champion *et al.* pour l'isolement de variants d'amylosaccharases actifs sur saccharose (substrat spécifique de glucane-saccharases) (Champion *et al.*, 2010). Cette méthode emploie un milieu de culture contenant du saccharose et du bleu de bromothymol, indicateur coloré qui devient jaune à pH acide. L'activité d'une amylosaccharase ou d'un variant actif d'amylosaccharase sur saccharose entraîne la libération de fructose et son utilisation par les cellules, aboutissant à la formation de composés acides diminuant ainsi le pH du milieu. Cette méthode a permis de réduire de 95 % une banque de mutants (8 000 clones pré-criblés) et ainsi facilité l'étape de criblage secondaire (Champion *et al.*, 2010).

Le criblage en milieu liquide est quant à lui plus complexe car il comporte plus d'étapes et nécessite un appareillage plus lourd pour la miniaturisation des essais, l'addition des réactifs et la détection des hits. Après transformation des banques plasmidiques, les clones sont dans un premier temps inoculés sur milieu gélosé. Les colonies sont ensuite repiquées et mises en culture en format microplaque (96 ou 384 puits). Après croissance des microorganismes, et expression génétique des variants enzymatiques,

les réactions sont elles-mêmes réalisées en microplaques dans des volumes allant de 5 à 2 000 μL . La manipulation de faibles volumes s'avère parfois problématique, à cause de l'évaporation rapide des solutions et des difficultés rencontrées pour le mélange du milieu réactionnel. La détection du produit réactionnel est ensuite réalisée par colorimétrie, fluorométrie ou chémoluminescence (Reymond and Babiak, 2007). Les récents développements rendent possible l'adaptation d'instruments analytiques lourds (chromatographie, spectrométrie de masse, RMN, etc.) aux protocoles de criblage à haut-débit. Ceci a pour effet de gagner en sensibilité et reproductibilité dans les mesures. Les travaux menés par le groupe du professeur Manfred T. Reetz sur l'énantioselectivité de catalyseurs et biocatalyseurs sont de bons exemples de la diversité des méthodes de criblage. Ces dernières font appel à la spectrométrie UV/VIS (Reetz *et al.*, 1997), la thermographie infrarouge (Reetz *et al.*, 2001a), la spectrométrie de masse (Reetz *et al.*, 1999), l'électrophorèse capillaire (Reetz *et al.*, 2000a), le dichroïsme circulaire (Reetz *et al.*, 2000b), la chromatographie en phase gazeuse (Reetz *et al.*, 2001b) et permettent la détermination de 700 à 20 000 excès énantiomériques par jours. Ce groupe a également fait la preuve de concept de l'utilisation de la RMN à haut-débit pour la résolution de mélanges racémiques ou la désymétrisation d'un substrat prochiral portant des groupes énantiotopiques, par l'emploi de *pseudo*-énantiomères ou composés *pseudo*-prochiraux marqués au C^{13} (Reetz *et al.*, 2002). Cette méthode n'a cependant jamais été utilisée pour le criblage de banques de biocatalyseurs. La RMN a également été utilisée pour le criblage de la spécificité de transglutaminases (TGases) envers leur substrat. Le principe repose sur le marquage isotopique au N^{15} de l'ovalbumine catalysé par une TGase (remplacement du groupe amine de la chaîne latérale d'une glutamine par un groupe amine marqué au N^{15}). Une analyse de l'ovalbumine marquée par RMN 2D ($\text{H}^1\text{-N}^{15}$) permet d'identifier des signaux correspondant aux résidus glutamine marqués de l'albumine. Plus la TGase réagit, plus il y a de signaux, ce qui traduit une enzyme moins spécifique (Shimba *et al.*, 2002).

Le tableau 3, ci-après, recense des exemples de méthodes de criblage pour l'optimisation de propriétés enzymatiques telles que l'activité, la thermostabilité, la spécificité de substrat, etc.

Tableau 3 : Exemples de méthodes de criblage utilisées et taille des banques analysées pour l'évolution dirigée d'enzymes.

Enzyme	Objectif du crible	Nombre de variants analysés	Référence
Criblage sur milieu solide			
Galactose oxydase	amélioration de l'efficacité catalytique	100 000	(Delagrave <i>et al.</i> , 2001)
P450 oxygénase	amélioration de l'activité enzymatique	200 000	(Joo <i>et al.</i> , 1999)
β -glycosidase	amélioration de la transglycosylation	16 000	(Koné <i>et al.</i> , 2009)
β -galactosidase	modification de la spécificité	10 000	(Parikh and Matsumura, 2005)
Amylosaccharase	Recherche de variants actifs (crible primaire)	8 000	(Champion <i>et al.</i> , 2010)
Criblage en milieu liquide			
Amylosaccharase	amélioration de la thermostabilité	2 700	(Emond <i>et al.</i> , 2008a; Emond <i>et al.</i> , 2007)
P450 peroxydase	modification de la spécificité	16 000	(Bloom <i>et al.</i> , 2006)
Oléandromycine glucosyltransferase	élargir la promiscuité	1 000	(Williams <i>et al.</i> , 2007)
Cyclodextrine glucanotransferase	résistance à un inhibiteur	12 000 (banque epPCR) 8 000 (banque shuffling)	(Kelly <i>et al.</i> , 2008)
Acetyltransferase	amélioration de la thermostabilité	1 000 (1er tour) 700 (2nd tour)	(Leemhuis <i>et al.</i> , 2008)

II.3 Ingénierie de la spécificité de transglycosidases vis-à-vis de leurs substrats et produits

La production d'hydrates de carbone complexes, par synthèse chimique ou extraction à partir de ressources biologique, est souvent fastidieuse et faiblement rentable (Muthana *et al.*, 2009). Dans ce contexte, la voie enzymatique et l'usage de CAZymes est une alternative qui offre beaucoup d'avantages en termes de production, de régiospécificité et de conditions de synthèse. Beaucoup d'efforts ont été développés ces dernières années pour concevoir par ingénierie des protéines, des enzymes de spécificité contrôlée et/ou capables de glycosyler un large panel de molécules.

Les **glycosynthases** sont des mutants de glycoside-hydrolases dont le résidu nucléophile catalytique a été substitué par un résidu non-nucléophile (alanine, glycine ou sérine). Ces mutants sont donc incapables d'hydrolyser une liaison glycosidique et de former un complexe covalent glycosyl-enzyme, nécessaire à la réaction de transglycosylation. Cependant, en présence d'un substrat donneur de glycosyle activé par un groupement de type fluorure, de configuration anomérique opposée à celle du substrat naturel (ce qui mime l'intermédiaire covalent glycosyl-enzyme), ces mutants sont capables de transférer l'unité glycosyle vers une molécule acceptrice (Shaikh and Withers, 2008). Ce sont donc des outils très intéressants pour la glycodiversification. Des travaux d'évolution dirigée ont permis d'améliorer les performances de ces enzymes mais également d'agrandir le répertoire de donneurs et accepteurs utilisables par les glycosynthases. Pour ce faire, Mayer *et al.* proposèrent une méthode de criblage sur milieu solide par couplage enzymatique (Mayer *et al.*, 2001). Cette méthode est basée sur

une co-expression de deux gènes, l'un codant pour la protéine d'intérêt et l'autre codant pour l'endo-cellulase (Cel5A) de *Cellulomonas fimi* qui catalyse le clivage des 4-méthylumbelliféryl cellobioside et cellobiotrioside pour former la 4-méthylumbelliférone, un produit fluorescent. Les auteurs ont alors criblé l'activité d'une banque de variants de la glycosynthase d'*Agrobacterium sp.* (Abg-glycosynthase), mutée par mutagenèse de saturation au niveau du nucléophile E358 pour la synthèse de méthylumbelliféryl β -cellooligosaccharides à partir du fluorure d' α -D-glucopyranosyle (donneur) et du 4-méthylumbelliféryl β -D-glucopyranoside (accepteur). Le mutant E358G a ainsi été identifié et isolé. L'efficacité de ce mutant a par la suite été améliorée par deux cycles de mutagenèse à erreurs en appliquant le même protocole de criblage comportant quelques améliorations (un seul plasmide portant le gène d'intérêt et le gène codant pour la Cel5A, et addition d'une étape de perméabilisation de la paroi cellulaire pour diminuer les quantités de donneurs et accepteurs) (Kim *et al.*, 2004). Ainsi, les auteurs ont obtenu le variant A19T:E358G:Q248R:M407V présentant une amélioration de la spécificité vis à vis du substrat d'un facteur 27 comparé au mono-mutant E358G.

Une seconde approche de criblage de l'activité des glycosynthases a été développée par Lin *et al.* qui ont employé la complémentation chimique par un système triple hybride chez la levure *S. cerevisiae* (Lin *et al.*, 2004). Ici, l'activité de la glycosynthase est reliée à la transcription du gène rapporteur *leu2* essentiel pour la survie de la cellule auxotrophe pour la leucine. L'expression de ce gène est contrôlée par le facteur de transcription DBD-AD dont les deux domaines, DBD et AD, ont été volontairement séparés. Leur liaison n'est possible que par l'intermédiaire d'une glycosynthase active. Les auteurs ont alors appliqué cette approche sur une banque de monomutants de la glycosynthase Cel7B d'*Humicola insolens*, isolant ainsi le variant E197S avec une efficacité catalytique améliorée d'un facteur 5.

Récemment, Ben-David *et al.* ont proposé une méthode universelle pour le criblage de glycosynthases basé sur la chute de pH (révélée au rouge de méthyle) engendrée par la libération de l'acide fluorhydrique lors du transfert de l'unité glycosyle vers l'accepteur (Ben-David *et al.*, 2008). Appliquée au criblage d'une banque de variants de β -xylosidase XynB2(E335G) glycosynthase de *Geobacillus stearothermophilus* obtenus par epPCR, la méthode a permis d'isoler un variant avec un k_{cat} amélioré d'un facteur 35 accompagné d'une augmentation du K_m envers le donneur et l'accepteur d'un facteur 17 et 930 respectivement.

Les enzymes de la famille des **glycosyltransférases** emploient des sucres activés pour la synthèse de liaisons osidiques. Les spécificités variées de ces enzymes en font aussi des outils de glycosylation très attractifs. L'ingénierie des glycosyltransférases pour le contrôle de la spécificité est également très prometteuse. Aharoni *et al.* ont développé une méthode utilisant la cytométrie en flux pour la sélection à haut-débit d'une banque de plus de 10^6 mutants de la sialyltransférases CstII de *Campylobacter jejuni*

obtenu par epPCR (Aharoni *et al.*, 2006). Les sialyltransférases catalysent le transfert de l'acide sialique (Neu5Ac) à partir de l'acide cytidine monophosphate N-acétylneuraminique (CMP-Neu5Ac) vers le groupement carbohydrate de divers glycolipides et glycoprotéines. La méthode proposée est basée sur la co-expression de deux gènes portés par deux plasmides, l'un codant pour la protéine d'intérêt et l'autre codant pour une CMP-Neu5Ac synthétase. Une fois transformées, les cellules ont été incubées avec du Neu5Ac et un sucre marqué par un fluorophore (lactose ou galactose). La CMP-Neu5Ac synthétase catalyse la transformation *in situ* du Neu5Ac en CMP-Neu5Ac fournissant ainsi le substrat donneur. En présence d'un variant actif, l'acide sialique est transféré sur le sucre marqué qui reste piégé dans la cellule. Après lavage (élimination des substrats non utilisés) les clones ont été triés par FACS. Les auteurs ont observé après séquençage que la mutation F91Y est fréquemment présente chez les actifs issus de la sélection. Le mutant CstIIIF91Y a donc été construit et caractérisé. Ce dernier présente une efficacité catalytique 153 et 367 fois supérieure envers le lactose et le galactose marqués respectivement, comparé au lactose et galactose non marqués. L'analyse de la structure tridimensionnelle du variant CstIIIF91Y a montré que le remplacement de la phénylalanine en position 91 par une tyrosine crée une poche hydrophobe complémentaire en taille du fluorophore présent sur les substrats, favorisant la liaison du substrat marqué à l'enzyme. Ceci est un bon exemple de biais que peuvent engendrer l'utilisation de substrats marqués ou d'analogues de substrats sur l'activité réelle du variant vis-à-vis d'un substrat naturel.

Williams *et al.* ont rapporté l'évolution dirigée de l'oléandomycine glycosyltransferase (OleD) de *Streptomyces antibioticus* qui catalyse la synthèse d'oléandomycine glucosylée à partir d'UDP-glucose (donneur) et d'oléandomycine (accepteur) (Williams *et al.*, 2007). Le criblage d'une banque de variants a été réalisé en microplaques en présence d'UDP-glucose et de 4-méthylumbelliférone (fluorescent), accepteur dont la vitesse de conversion en 4-méthylumbelliféryl β -D-glucopyronoside (non-fluorescent) est 300 fois moins importante qu'avec l'accepteur naturel. Cette campagne a abouti à la caractérisation d'un triple mutant dont l'efficacité catalytique envers l'accepteur non naturel est 30 fois plus élevée. En conséquence, une banque d'une vingtaine de donneurs de glycosyle non naturels et 6 différents accepteurs ont été testés. Les résultats obtenus ont confirmé l'élargissement de la promiscuité des variants envers les substrats (donneurs et accepteurs) par rapport à la protéine sauvage.

L'emploi de substrats activés et non naturels peut être très coûteux et donc peu intéressant pour les applications à faible valeur ajoutée. Des travaux d'ingénierie ont donc été menés sur des outils de transglycosylation utilisant des substrats naturels. Pour ce faire Koné *et al.* ont développé un crible utilisant l'imagerie digitale pour l'évolution dirigée de β -glucosidases et l'amélioration du taux de transglycosylation T/H (T/H correspond à l'activité transglycosidase/activité hydrolase) (Koné *et al.*,

2009). Après transformation par une banque plasmidique, les cellules ont été inoculées sur une membrane de nitrocellulose, elle-même déposée sur un milieu de culture solide. La membrane a ensuite été transférée, après croissance des cellules, sur un milieu minimum contenant le substrat donneur (X-Gal) et un inducteur activant l'expression des variants. Au bout d'un certain temps la membrane a été une nouvelle fois transférée sur un milieu minimum contenant le X-Gal et du cellobiose (accepteur). L'activité est suivie par libération du chromophore du X-Gal. Ainsi, une image a été prise après incubation des colonies avec uniquement du X-Gal où l'apparition de cellules colorées est rattachée à une réaction d'hydrolyse. Une deuxième image a été prise après incubation avec le donneur et l'accepteur où l'intensité de couleur est cette fois rattachée à une réaction de transglycosylation. Le T/H peut donc être calculé par comparaison des intensités de coloration des colonies avec et sans la présence d'accepteur. L'application de cette stratégie à une banque de variants de la β -glucosidase de *Thermus aquaticus* a permis l'identification de mutants dont le T/H présentait une augmentation globale d'un facteur 70 (Koné *et al.*, 2009).

De récents travaux ont porté sur l'ingénierie semi-rationnelle d'une α -transglucosidase pour la synthèse d'oligosaccharides mimant le motif antigénique des sérotypes 1b et 3a de *Shigella flexneri* (Champion *et al.*, 2009). L'objectif était de glucosyler deux accepteurs, le *N*-acetyl-D-glucosamine et le L-rhamnopyranose à partir du saccharose (donneur) pour former les motifs α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acetyl- β -D-glucosamine et α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranose. L'analyse détaillée de la structure de l'amylosaccharase de *Neisseria polysaccharea* (ASNp) en complexe avec les accepteurs de l'étude a permis d'identifier 7 cibles (résidus du sous-site +1), qui ont été systématiquement remplacées par les 19 acides aminés possibles. Les 133 monomutants obtenus ont été criblés par réaction d'accepteur réalisée en microplaques suivie d'une analyse des produits synthétisés par HPLC. Ainsi deux mutants, ASNpI228Y et ASNpF290K ont révélé des améliorations nettes de leur efficacité catalytique envers le L-rhamnopyranose et le *N*-acetyl-D-glucosamine respectivement, accepteurs peu ou pas reconnus par la protéine sauvage.

En ce qui concerne les glucane-saccharases de la famille GH70, les travaux d'ingénierie ont principalement porté sur la modification de la spécificité de liaison de ces enzymes et sur l'identification des facteurs structuraux impliqués dans cette spécificité (cf. paragraphe I.3.7, partie I). De telles approches nécessitent la caractérisation de la structure des produits synthétisés par les variants enzymatiques et requièrent l'utilisation de diverses méthodes incluant la spectroscopie RMN, la spectroscopie de masse, les dégradations enzymatiques et chimiques, l'analyse par méthylation et acétylation. Ces méthodes sont lourdes, longues et de ce fait peu adaptées à des débits élevés de criblage. De ce fait, les travaux d'ingénierie de la spécificité des glucane-saccharases de la famille

GH70 ont été principalement focalisés sur des approches rationnelles et semi-rationnelles, sur un espace de séquence limité (motif IV en particulier), afin de générer des banques de tailles raisonnées (de l'ordre de la dizaine de variants en général). Seul Hellmuth *et al.* décrivent le criblage de 2 000 mutants, issus de l'ingénierie semi-rationnelle de la glucane-saccharase GTF-R, par une analyse visuelle du changement d'apparence et de viscosité des milieux de synthèse (Hellmuth *et al.*, 2008).

La présentation des différentes techniques d'ingénierie d'enzymes exposée dans ce chapitre n'est pas exhaustive. Il est cependant facile de s'imaginer le large éventail de possibilités qu'il est possible d'utiliser pour modifier les propriétés d'un biocatalyseur. Chaque approche présente avantages et limites. Le choix final de la méthode de création de la diversité sera donc fonction du panel de connaissances disponibles sur les relations structure/fonction de l'enzyme d'intérêt. Toutefois, la sélection et/ou le criblage des mutants les plus performants est l'autre paramètre clé à prendre en considération, quelle que soit l'approche envisagée. Le nombre de variants générés lors de la construction d'une banque peut être considérable, de 10^4 à 10^{12} suivant la méthode et les étapes expérimentales (clonage, transformation de microorganismes, etc.). Il faut donc être capable d'explorer cette diversité avec précision et à haut-débit, afin de maximiser l'exploitation des informations. Enfin, les méthodes de sélection et criblage ne sont pas universelles et sont bien souvent spécifiques des biocatalyseurs pour lesquels elles sont développées.

Concernant l'ingénierie de la spécificité des glucane-saccharases de la famille GH70, il n'existait au début de cette thèse aucune méthode de criblage haut-débit permettant de discriminer des variants de CAZymes de spécificités variées, ce qui limite considérablement l'espace de séquence exploré. Un des défis que nous avons relevé a donc été de développer une méthodologie de sélection et de criblage de la spécificité des glucane-saccharases suffisamment puissante pour analyser de larges banques de variants.

Partie 3 : α -glucanes produits à partir de saccharose

III. α -glucanes produits à partir de saccharose

Ce chapitre vise à faire la synthèse des données structurales, des propriétés et des applications des α -glucanes synthétisés par les glucane-saccharases. Il n'est volontairement pas étendu aux autres α -glucanes constituant les polysaccharides de réserve des plantes, des champignons et des animaux, en l'occurrence l'amidon, le pullulane et le glycogène qui ont été beaucoup plus étudiés et dont la biosynthèse met en jeu une machinerie complexe de glycosyltransférases, d'enzymes de branchement et de débranchement.

III.1 Caractéristiques structurales

Comme présenté précédemment, la plupart des α -glucanes produits à partir du saccharose sont synthétisés par les bactéries lactiques des genres *Lactobacillus*, *Leuconostoc*, *Streptococcus*, *Exiguobacterium* et *Weissella*, ainsi que par les bactéries des genres *Neisseria*, *Deinococcus* et *Alteromonas*. Ces polymères sont définis comme des homopolysaccharides constitués d'unités glucosyle associées par des liaisons osidiques de type $\alpha(1\rightarrow6)$, $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$ ou $\alpha(1\rightarrow2)$, présentes dans la chaîne principale et/ou au niveau de ramifications. Les fonctions physiologiques de ces polymères, notamment des dextrans, ne sont pas pleinement connues. En effet, les bactéries productrices de dextrans ne dégradent généralement pas le polymère qu'elles produisent, ce qui indique que le dextrane n'interviendrait pas dans le stockage d'énergie (Leathers, 2005). En revanche, le mutane produit par les *Streptocoques* de la flore buccale est connu pour participer à la formation de la plaque dentaire, fournissant ainsi un environnement pour la prolifération de ces bactéries. De manière générale, il est donc possible que ces polymères facilitent la colonisation de certains habitats. Ils pourraient aussi jouer un rôle protecteur contre la dessiccation ou l'attaque par d'autres organismes.

III.1.1 Structure des α -glucanes

La caractérisation structurale des α -glucanes est très complexe en raison de la taille de ces polymères ainsi que de la variété et des multiples arrangements possibles des liaisons osidiques les composant. A l'heure actuelle, la caractérisation des α -glucanes produits par les glucane-saccharases concerne presque exclusivement les dextrans. La diversité ces polymères a été mise en évidence dès les années 50 par analyse des produits générés par oxydation périodique (Jeanes *et al.*, 1954). Par la suite, d'autres techniques d'analyse structurale incluant la méthylation, l'acétolyse partielle, la

dégradation de Smith, l'hydrolyse enzymatique, la RMN du carbone 13 et du proton ont permis d'avoir des informations plus précises sur la distribution ou la longueur des ramifications (Larm *et al.*, 1971; Van Leeuwen *et al.*, 2008d).

La spectroscopie de résonance magnétique nucléaire constitue depuis les années 70 une méthode de choix pour l'identification et la quantification des liaisons osidiques constituant les α -glucanes. Différentes études ont permis l'attribution des signaux ^1H et ^{13}C de différents α -glucanes (Cheetham *et al.*, 1991; Seymour *et al.*, 1979a; Seymour *et al.*, 1979b; Shimamura, 1989). Ainsi, les dextrans sont caractérisés par un taux d'unités glucosyle liées à plus de 50 % par des liaisons osidiques de type $\alpha(1\rightarrow6)$ dans la chaîne principale et des taux variables de branchements en $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$ et $\alpha(1\rightarrow2)$. L'alternane produit par la souche *Ln. citreum* B-1355 est composé de liaisons osidiques $\alpha(1\rightarrow6)$ et $\alpha(1\rightarrow3)$ alternées et présente quelques ramifications en $\alpha(1\rightarrow3)$. Les mutanes sont quant à eux des polymères composés à plus de 50 % de liaisons de type $\alpha(1\rightarrow3)$. Les reuteranes, dont la découverte est très récente (Kralj *et al.*, 2005a; Kralj *et al.*, 2002), sont constitués à plus de 50 % de liaisons $\alpha(1\rightarrow4)$. Enfin, les $\alpha(1\rightarrow4)$ -glucanes produits par les amylosaccharases sont très similaires à l'amylose, le constituant linéaire de l'amidon des plantes, excepté en ce qui concerne leur degré de polymérisation, beaucoup plus faible (Potocki-Veronese *et al.*, 2005; Potocki de Montalk *et al.*, 2000). Plus récemment, van Leeuwen *et al.* ont proposé une méthode globale comprenant des données de RMN 1D du proton (Tableau 4), 2D ^{13}C - ^1H , méthylation, hydrolyse acide partielle, hydrolyse enzymatique et dégradation de Smith pour caractériser en détail la structure des α -glucanes produits par les glucane-saccharases GTF180 et GTFA (Figure 19) (Van Leeuwen *et al.*, 2008d).

Tableau 4 : Attribution des déplacements chimiques par RMN ^1H pour différents glucanes solubilisés dans du D_2O à 27°C (Van Leeuwen *et al.*, 2008d).

Glucane	H1	H2	H3	H4	H5	H6
[α -D-Glcp-(1 \rightarrow 2)-]	5,084	3,55	3,78	3,45	3,932	3,84/3,77
[α -D-Glcp-(1 \rightarrow 3)-]	5,359	3,562	3,76	3,431	4,01	3,87/3,80
[α -D-Glcp-(1 \rightarrow 4)-]	5,399	3,59	3,67	3,418	3,72	3,84/3,74
[α -D-Glcp-(1 \rightarrow 6)-]	4,961	3,55	3,72	3,422	3,72	3,84/3,78
[α -D-Glcp-(1 \rightarrow 3)-]	5,386/5,371	3,68	3,9	3,66	4,03	3,85/3,74
[α -D-Glcp-(1 \rightarrow 4)-]	5,399	3,63	3,96	3,65	3,85	3,85/3,82
[α -D-Glcp-(1 \rightarrow 6)-]	4,961/4,969	3,57	3,72	3,49	3,93	3,78/3,983

Les dextrans commerciaux produits par la souche *Ln. mesenteroides* NRRL B-512F ont aussi été caractérisés en détail. Ils comprennent plus de 95 % de liaisons $\alpha(1\rightarrow6)$ et des branchements liés en $\alpha(1\rightarrow3)$. Larm *et al.* ont estimé qu'environ 40 % de ces ramifications sont composés d'une seule unité glucosyle et 45 % de deux unités. Les autres branchements (environ 15%) seraient formés de plusieurs

dizaines d'unités glucosyle (Larm *et al.*, 1971). Par des analyses de diffusion de la lumière, Ioan *et al.* estiment que la longueur des ramifications peut atteindre 179 unités glucosyle (Ioan *et al.*, 2001). Cette structure serait très proche de celle du dextrane synthétisé par la souche *Weissella confusa* (Figure 20), récemment analysée par une méthodologie similaire à l'approche employée par van Leeuwen *et al.* (Maina *et al.*, 2011). En effet, ce dextrane est composé à ~97 % de liaisons $\alpha(1\rightarrow6)$ avec des ramifications en $\alpha(1\rightarrow3)$. Bien que les résultats n'aient pas permis aux auteurs de conclure sur la longueur des ramifications, ces derniers suggèrent la présence de branchements de longueur supérieure à deux unités glucosyle.

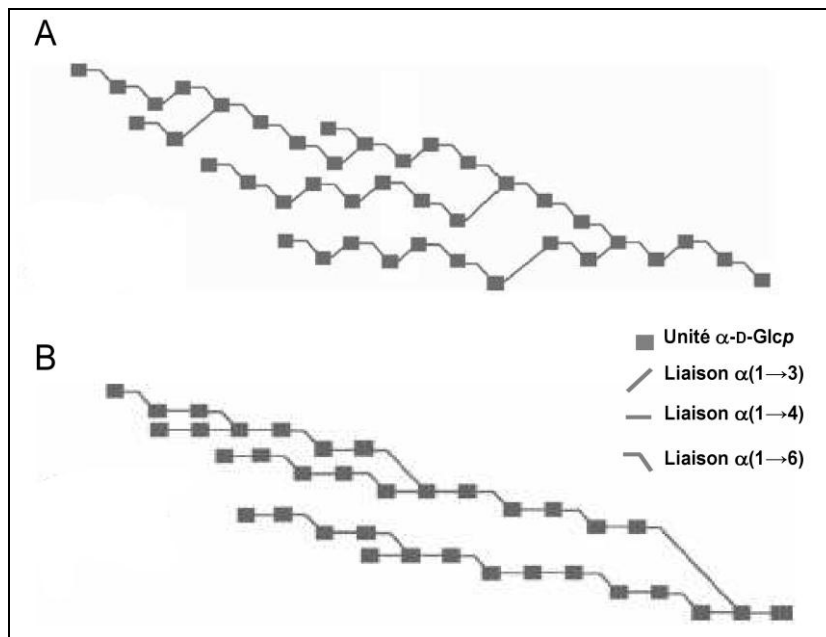


Figure 19 : Représentation schématique du (A) dextrane et (B) reuterane synthétisés par les glucane-saccharases GTF180 et GTF-A, respectivement. Adapté de (van Leeuwen *et al.*, 2008b; Van Leeuwen *et al.*, 2008c).

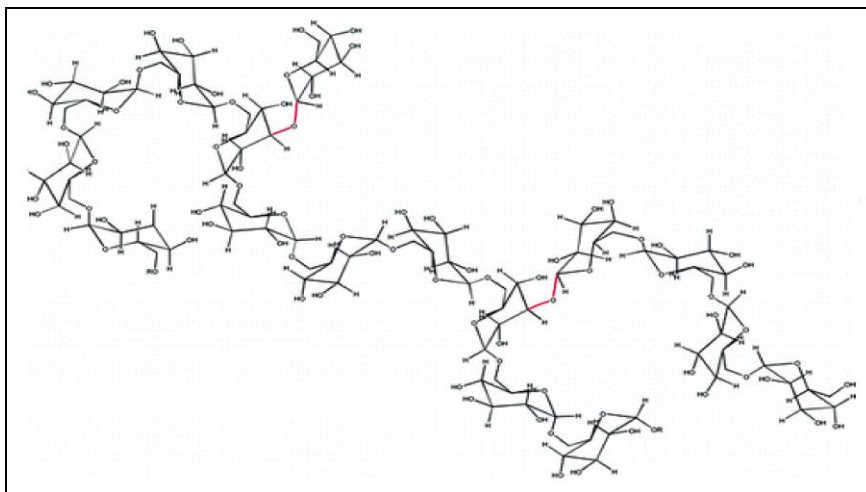


Figure 20 : Représentation schématique du dextrane synthétisé par la souche *Weissella confusa*. Les liaisons en rouge correspondent aux liaisons $\alpha(1\rightarrow3)$. Adapté de (Maina *et al.*, 2011).

Enfin, en raison de la complexité et la diversité structurale des α -glucanes, très peu de travaux ont été consacrés à la résolution de la structure atomique de ces polymères. Citons les travaux de Popov *et al.* et de Montesanti *et al.*, qui ont permis d'obtenir des cristaux de longueur supérieure à 5 μm à partir de fractions purifiées de malto-oligosaccharides et de courtes chaînes d'amylose de degrés de polymérisation variables (2 à 30), synthétisés par l'amylosaccharase de *N. polysacchara* (Montesanti *et al.*, 2010; Popov *et al.*, 2006). Grâce à ces cristaux, la première structure de polymère a pu être obtenue par microdiffraction des rayons X, avec une résolution exceptionnelle (1,3 Å) (Popov *et al.*, 2009). En ce qui concerne les dextrans, seuls ont été reportés les travaux de Guizard *et al.* (Guizard, 1981), qui ont porté sur la caractérisation de la structure tridimensionnelle de dextrans de faibles masses moléculaires (DP = 60) synthétisés par voie chimique. Les clichés de diffraction électronique des monocristaux obtenus, de quelques dizaines d'Angstroms d'épaisseur, ont permis de construire des modèles structuraux de deux polymorphes obtenus par cristallisation à des températures variées (Figure 21). Néanmoins, il n'existe à l'heure actuelle aucun travail de caractérisation structurale à l'échelle atomique d' α -glucanes de hautes masses molaires synthétisés par voie enzymatique.

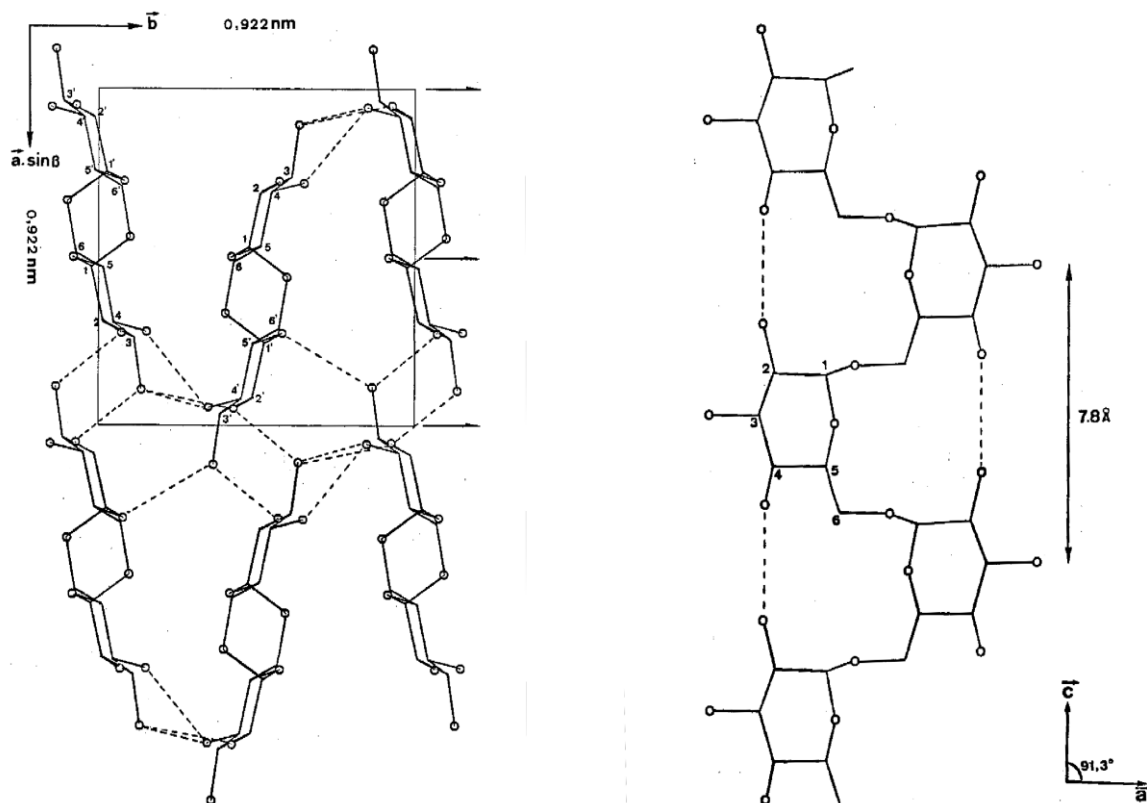


Figure 21 : Modèle structural du dextrane (polymorphe haute température) suivant 2 plans de vue différents. Tiré de (Guizard, 1981)

III.1.2 Tailles des α -glucanes

L'estimation de la taille et de la conformation en solution des α -glucanes est essentielle pour leur caractérisation. La masse moléculaire est généralement déterminée par chromatographie d'exclusion de taille (SEC) couplée ou non à un détecteur de diffusion de la lumière. Les études réalisées indiquent toutes que les dextrans natifs présentent une masse moléculaire très élevée, variant de 10^5 à 10^9 Da suivant les enzymes responsables de la synthèse. Il est important de souligner que les différentes études sont souvent difficiles à comparer entre elles car les équipements employés, notamment les colonnes d'exclusion, ne présentent pas toutes les mêmes caractéristiques. De plus, les valeurs de masse moléculaire supérieure à 10^6 Da se situent dans la limite supérieure des masses détectables par les techniques d'exclusion, les polymères de très haute masse molaire étant parfois difficiles à distinguer des agrégats. En outre, l'agrégation de ces polymères, tout comme les problèmes de solubilisation, diminue les rendements d'élution de SEC, la fraction éluée pouvant donc ne pas être représentative de l'ensemble de l'échantillon. Enfin, les conditions réactionnelles (pH, température, concentration en saccharose) ont aussi un effet sur la taille et le degré de branchement des dextrans (Braswell *et al.*, 1962; Braswell and Stern, 1959; Falconer *et al.*, 2011; Kim *et al.*, 2003; Sabatie *et al.*, 1988) ce qui, là encore, ne facilite pas la comparaison des résultats. Le tableau 5 recense des masses molaires de quelques α -glucanes.

Tableau 5 : Masse molaire de certains α -glucanes trouvée dans la littérature.

Glucane	Enzyme	souche	Masse molaire (g.mol ⁻¹)	Référence
Dextrane	DSR-S (native)	<i>Ln. mesenteroides</i> NRRL B512-F		
	DSR-S (recombinante)		$> 10^7$	(Moulis <i>et al.</i> , 2006a)
	DSR-S vardel Δ 4N (mutant tronqué)			
	GTF180 (native)	<i>Lb. reuteri</i> 180	$3,6 \cdot 10^7$	(Kralj <i>et al.</i> , 2004a)
	GTF180 (mutant tronqué)		$3,2 \cdot 10^7$	(Van Leeuwen <i>et al.</i> , 2009)
	GTF-kg15 (native)	<i>Lb. sakei</i> Kg15	$2,7 \cdot 10^7$	
	GTF-kg15 (native)	<i>Lb. fermentum</i> Kg3	$2,4 \cdot 10^7$	(Kralj <i>et al.</i> , 2004a)
	GTF-33 (native)	<i>Lb. parabuchnerii</i> 33	$2 \cdot 10^5$	
Reuterane	GTF-A (native)	<i>Lb. reuteri</i> 35-5	$4 \cdot 10^7$	
	GTF-A (recombinante)		$8 \cdot 10^7$	
	GTF-A (native)	<i>Lb. reuteri</i> 121	$4 \cdot 10^7$	(Kralj <i>et al.</i> , 2004b)
	GTF-A (recombinante)		$4,5 \cdot 10^7$	
	GTF-A (mutant tronqué)		$5 \cdot 10^7$	
	GTF-O (native)	<i>Lb. reuteri</i> BioGaia	$2,8 \cdot 10^7$	
	GTF-O (recombinante)		$4,5 \cdot 10^7$	(Kralj <i>et al.</i> , 2005a)
Alternane	ASR (native)	<i>Ln. citreum</i> B-1355	$1 \cdot 10^7$	(Côté, 1992)
	ASR (native)	<i>Ln. citreum</i> B-1355	$3,5-4,4 \cdot 10^7$	(Striegel <i>et al.</i> , 2009)

Une alternative à l'analyse par chromatographie d'exclusion couplée à la diffusion de la lumière est l'utilisation d'un système de fractionnement par flux/force avec flux asymétrique (en anglais Asymmetrical Flow Field-Flow Fractionation, ou AFFFF). Le système de fractionnement AFFFF est constitué d'une cellule de séparation, asymétrique formée d'une paroi solide et d'une paroi d'accumulation perméable (membrane d'ultrafiltration) séparées par un canal d'élution (Figure 22). Le seuil de coupure de la membrane d'ultrafiltration est typiquement de l'ordre de 10 kDa, qui est donc la limite inférieure pour la détermination de la masse molaire. La séparation a lieu dans un canal de faible épaisseur dans lequel s'écoule l'éluant, suivant une distribution parabolique. Ce flux transporte l'échantillon de manière axiale le long du canal. La séparation est rendue possible par l'application d'un champ de force externe, perpendiculaire au canal (flux croisé), qui entraîne les molécules vers la paroi d'accumulation. La distance entre la paroi d'accumulation et le centre de gravité de chaque macromolécule dépend du coefficient de diffusion de ces dernières. Ainsi, les molécules les plus éloignées de la paroi sont plus rapidement éluées, grâce au profil parabolique du flux axial. Par opposition à la chromatographie par exclusion de taille, ce système permet d'éluer les petites molécules en premier. L'absence de phase stationnaire limite toute dégradation de l'échantillon.

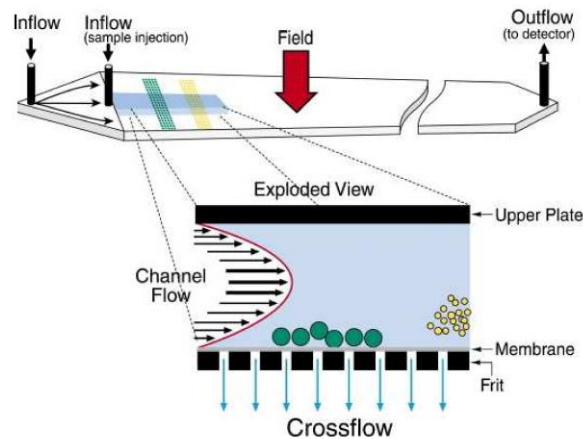


Figure 22 : Principe de la séparation par AFFFF. Représentation schématique du système de séparation AFFFF. A l'intérieur de la cellule les échantillons sont sous l'influence de deux vecteurs de flux qui les transportent vers le détecteur (channel flow) et vers la paroi d'accumulation (crossflow). La séparation dans le temps est dépendante de la position de l'échantillon dans la cellule, elle-même déterminée par le coefficient de diffusion de chaque échantillon. Tiré de (Yohannes *et al.*, 2011).

En AFFFF, le temps de rétention t_R s'écrit en fonction du coefficient de diffusion :

$$t_R = \frac{t^0 \cdot V_c \cdot w^2}{6 \cdot V^0 \cdot D} \quad (1)$$

avec t^0 le temps mort (correspondant à la traversée du volume de la cellule), V^0 le volume de la cellule, w l'épaisseur de la cellule, V_c le débit de flux croisé et D le coefficient de diffusion. Ce coefficient de diffusion peut s'écrire en fonction du rayon hydrodynamique de la molécule suivant l'équation approximée de Stokes-Einstein :

$$D = \frac{k.T}{6.\pi.\eta.r_H} \quad (2)$$

où k est la constante de Boltzmann, T la température, η la viscosité de l'éluant et r_H le rayon hydrodynamique de l'échantillon. Le rayon hydrodynamique de l'échantillon peut alors être exprimé en fonction du temps de rétention :

$$r_H = \frac{V^0.k.T}{V_c.w^2.t^0.\pi.\eta} t_R \quad (3)$$

Ainsi, l'AFFFF permet de fractionner des macromolécules qui diffèrent en taille et conformation. Cette caractéristique est particulièrement intéressante lorsque l'AFFFF est couplé à une détection de la diffusion multi-angulaire de la lumière laser (en anglais Multiangle Laser Light Scattering, ou MALLS). Ceci permet de déterminer les masses moléculaires et le rayon de giration des macromolécules grâce à la théorie de la diffusion de lumière. En effet, soumise à un rayonnement électromagnétique, la molécule se polarise et diffuse cette lumière. L'intensité de la lumière diffusée dépend, entre autre, de l'angle d'observation, de la masse molaire et du rayon de giration. L'équation de Zimm décrit la relation entre ces différents paramètres (Zimm, 1948):

$$\frac{K.c}{R_\theta} = \frac{1}{M_w.P(\theta)} + 2A_2.c \quad (4)$$

où K est une constante optique, c la concentration de l'échantillon, R_θ le rapport d'excès de Rayleigh du soluté, M_w la masse molaire, A_2 le second coefficient viriel et $P(\theta)$ le facteur de forme qui dépend de la taille et de la conformation de la molécule et décrit la dépendance angulaire de l'intensité de la lumière diffusée :

$$P(\theta) = \frac{1}{1 + \frac{16\pi^2}{3\lambda^2} \langle R_G \rangle^2 \sin^2 \left(\frac{\theta}{2} \right)} \quad (5)$$

avec λ la longueur d'onde de la radiation incidente, θ l'angle d'observation et R_G le rayon de giration.

En pratique, la masse molaire (M_i) et le rayon de giration (R_{Gi}) sont déterminés pour chaque tranche (i) de pic d'élugramme (Figure 23) en utilisant l'extrapolation de Zimm :

$$\left(\frac{K.c}{R_\theta}\right)_i = \frac{1}{M_i} \left(1 + \frac{16\pi^2 n^2}{3\lambda^2} R_{Gi}^2 \sin^2\left(\frac{\theta}{2}\right)\right) \quad (6)$$

où n correspond à l'indice de réfraction du solvant. La concentration c_i est déterminée à partir de l'indice de réfraction différentiel. La masse molaire moyenne et la moyenne quadratique du rayon de giration sont calculées par la sommation de l'ensemble des tranches d'un pic :

$$\bar{M}_w = \frac{\sum_i c_i M_i}{\sum_i c_i} \quad \text{et} \quad \bar{R}_G = \frac{\sum_i c_i M_i R_{Gi}^2}{\sum_i c_i M_i} \quad (7) \text{ et } (8)$$

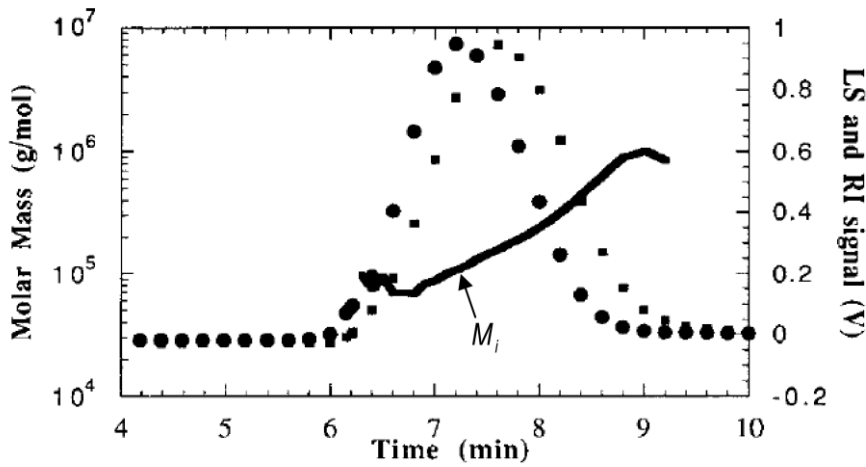


Figure 23 : Elugramme (réponse de diffusion de lumière (LS) à l'angle 90° et réponse de l'indice de réfraction (RI) et masses molaires en fonction du temps d'éluion pour une solution d'un dextrane commercial de 150 kDa solubilisé dans une solution de nitrate de sodium à 0,1 M. L'analyse a été réalisée à 60°C . Les cercles représentent la réponse de l'indice de réfraction, les carrés la réponse de diffusion de lumière. Tiré de (Viebke and Williams, 2000).

D'autres informations peuvent être tirées d'une analyse par AFFFF-MALLS. Sur la base de la théorie des polymères en solution, il existe différentes lois de puissance qui permettent de décrire la conformation moléculaire d'un polymère (Teraoka, 2002). Ainsi, la loi de puissance qui décrit la dépendance du rayon de giration à la masse molaire est la suivante :

$$R_{Gi} = K_G M_i^{\nu_G} \quad (9)$$

L'exposant ν_G dépend de la configuration du polymère, de la température et de l'interaction polymère-solvant (Rolland-Sabaté *et al.*, 2008). Pour un polymère sous forme (i) sphérique $\nu_G = 0.33$, (ii) de pelote statistique $\nu_G = 0.5-0.6$ et (iii) de bâtonnet $\nu_G = 1$. La valeur de ν_G correspond à la pente de la courbe du logarithme de R_{Gi} en fonction du logarithme de M_i (Figure 24). Les valeurs de ν_G de dextrans commerciaux rapportées dans la littérature sont répertoriées dans le tableau 6.

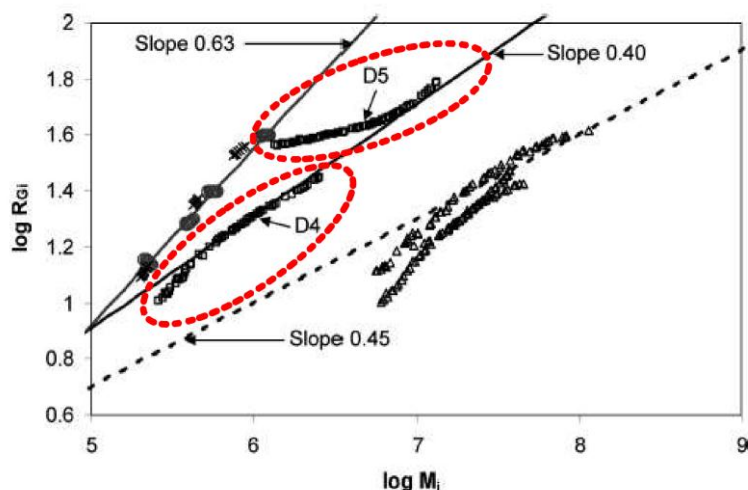


Figure 24 : Courbes de R_G en fonction de M_w de divers biopolymères solubilisés dans une solution d'azide de sodium concentrée à $0,2 \text{ g.L}^{-1}$. L'analyse a été réalisée à 40°C . (\square) Dextrans commerciaux D4 (483,1 kDa) et D5 (2 906 kDa), (\times) pullulane, (\circ) amylose et (Δ) glycogène de foie de lapin, huitre et phytoglycogène. Tiré de (Rolland-Sabaté *et al.*, 2008).

Tableau 6 : Relation entre la masse molaire et le rayon de giration pour des dextrans commerciaux.

Référence	Solvant	$M_w (\times 10^3 \text{ g.mol}^{-1})$	\bar{R}_G (nm)	V_G
(Roessner and Kulicke, 1994)	NaNO_3 0,1M	2730	47	0,50
(Wittgren and Wahlund, 1997)	NaCl 0,1 M	408,8 2540	16 34	0,36 0,27
(Viebke and Williams, 2000)	NaNO_3 0,1M	105 - 2100	-	0,49
(Rolland-Sabaté <i>et al.</i> , 2008)	NaN_3 0,2 g.L^{-1}	483,1 2906	19,2 47,7	0,37 ^a 0,36 ^b
(Ioan <i>et al.</i> , 2000b)	H_2O	132 - 2660	12 - 47	0,43
	NaOH 0,5 M	166 - 2700	14 - 50	0,40

^a valeur déterminée dans la région où $M_i > 5,81 \cdot 10^5 \text{ g.mol}^{-1}$

^b valeur déterminée dans la région où $M_i > 5,72 \cdot 10^6 \text{ g.mol}^{-1}$

A ce jour, sur l'ensemble des α -glucanes produits à partir de saccharose, seul les dextrans commerciaux de tailles inférieures à $3 \cdot 10^3$ kDa, ont été étudiés par AFFFF.

III.2 Propriétés et caractéristiques physico-chimiques

III.2.1 Propriétés rhéologiques

De par la proportion importante de liaisons $\alpha(1 \rightarrow 6)$, le dextrane synthétisé par *Leuconostoc mesenteroides* NRRL B-512F a une forte solubilité dans l'eau et dans d'autres solvants tels que l'urée, la glycine, le glycérol, le formamide, etc... (de Belder, 1990; Guizard, 1981; Jeanes, 1966). A faible concentration, les chaînes de dextrane en solution, que l'on peut représenter comme des pelotes

individuelles aléatoires (ou pelotes statistiques), ont un comportement Newtonien (de Belder, 1990). Au fur et à mesure que la concentration en polymère augmente, les interactions chaîne-chaîne sont favorisées aux dépens des interactions chaîne-solvant et les chaînes individuelles s'interpénètrent les unes avec les autres pour former un réseau de chaînes enchevêtrées aboutissant à la formation de pelotes compactes avec un comportement pseudoplastique (Ioan *et al.*, 2000a; McCurdy *et al.*, 1994). Ce phénomène est observé dans la figure 25 ci-dessous représentant la viscosité spécifique de dextrans en fonction de la concentration en solution.

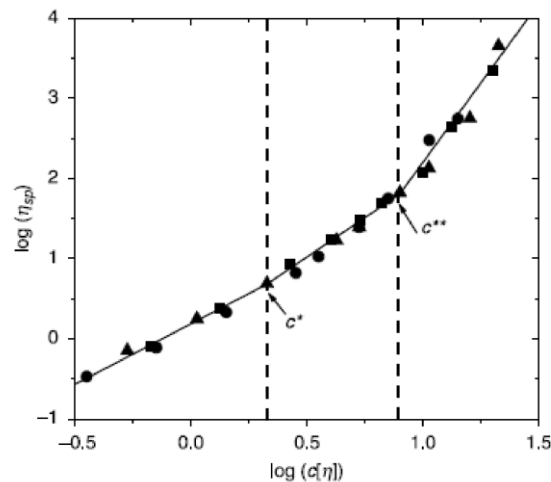


Figure 25 : Viscosité spécifique (η_{sp}) en fonction de la concentration réduite ($C[\eta]$) pour trois dextrans commerciaux de différentes masses molaires : $M_w = 2.10^6$ (■), (●) 2.10^5 et (▲) 5.10^5 g.mol⁻¹. Adapté de (Pinder *et al.*, 2006).

La courbe de la figure 25 définit 3 pentes correspondant à 3 domaines (dilué, semi-dilué et concentré) délimités par deux concentrations critiques C^* et C^{**} . En dessous de C^* , dans le domaine dilué, les chaînes de dextrans sont individualisées et la viscosité varie selon $\eta = C^{1.4}$. La région semi-diluée est délimitée par les deux concentrations critiques. Dans cette région, les pelotes s'enchevêtrent et se contractent avec l'augmentation de la concentration jusqu'à $C = C^{**}$, par laquelle la dimension des bobines se stabilise et reste constante. Dans cette région la viscosité varie selon $\eta = C^2$. Au-delà de C^{**} , la viscosité varie selon $\eta = C^{3.8}$. Plusieurs auteurs ont rapporté des valeurs très proches de dépendance entre concentration et viscosité des dextrans commerciaux (McCurdy *et al.*, 1994; Pinder *et al.*, 2006; Sabatié *et al.*, 1988). Cependant, Sabatié *et al.* n'observent pas de domaine semi-dilué pour des dextrans synthétisés à partir de la dextrane-saccharase de *Ln. mesenteroides* NRRL B-512F purifiée. Ces dextrans auraient une seule concentration critique (Sabatié *et al.*, 1988). De même, la dépendance de la viscosité en fonction de la concentration dans le domaine concentré peut largement excéder la valeur de puissance 3,8 obtenu pour les dextrans commerciaux.

Aux fortes concentrations, la viscosité d'une solution de dextrane est aussi fonction de la concentration, de la température et du poids moléculaire du polymère (Carrasco *et al.*, 1989; Sabatié *et al.*, 1988). L'évolution de la viscosité en fonction de la concentration est donnée pour différentes fractions de dextrane, dans la figure 26.

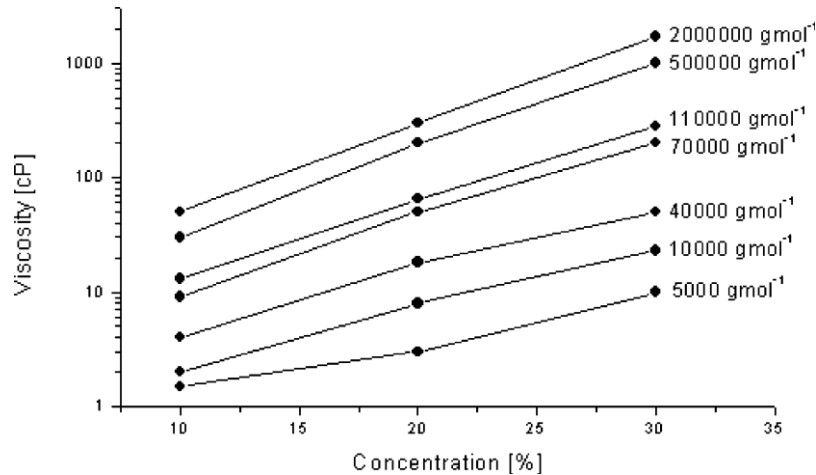


Figure 26 : Corrélation entre la viscosité et la concentration de différentes fractions de dextrane, à 25°C (Heinze *et al.*, 2006).

Les solutions très visqueuses peuvent se comporter simultanément comme des substances élastiques (comportement d'un gel) ou visqueuses (comportement d'un liquide). Ces deux états de réponse peuvent être différenciés par des mesures rhéologiques en mode dynamique. La solution est alors soumise à une contrainte dont la force varie de façon sinusoïdale en fonction de la fréquence imposée. La déformation de l'échantillon qui en résulte peut être modélisée par une fonction sinusoïdale de même fréquence. La réponse générale est caractérisée par un module de rigidité complexe G^* , donné par la relation :

$$G^* = \sqrt{G'^2 + G''^2} \quad (10)$$

G' étant le module élastique (ou module de stockage d'énergie) et G'' le module visqueux (ou module de perte d'énergie). G' est d'autant plus important que l'échantillon est à prédominance élastique ou fortement structuré, alors que G'' est d'autant plus élevé que l'échantillon est visqueux. A partir des valeurs de G' et G'' , il est possible de calculer la viscosité dynamique $|\eta^*|$ par la relation :

$$|\eta^*| = \sqrt{G'^2 + G''^2} / \omega \quad (11)$$

où ω représente la fréquence d'oscillation imposée (contrainte). Ces paramètres apportent des informations sur l'état et les arrangements (interactions ou non) entre les chaînes du polymère en solution (Figure 27).

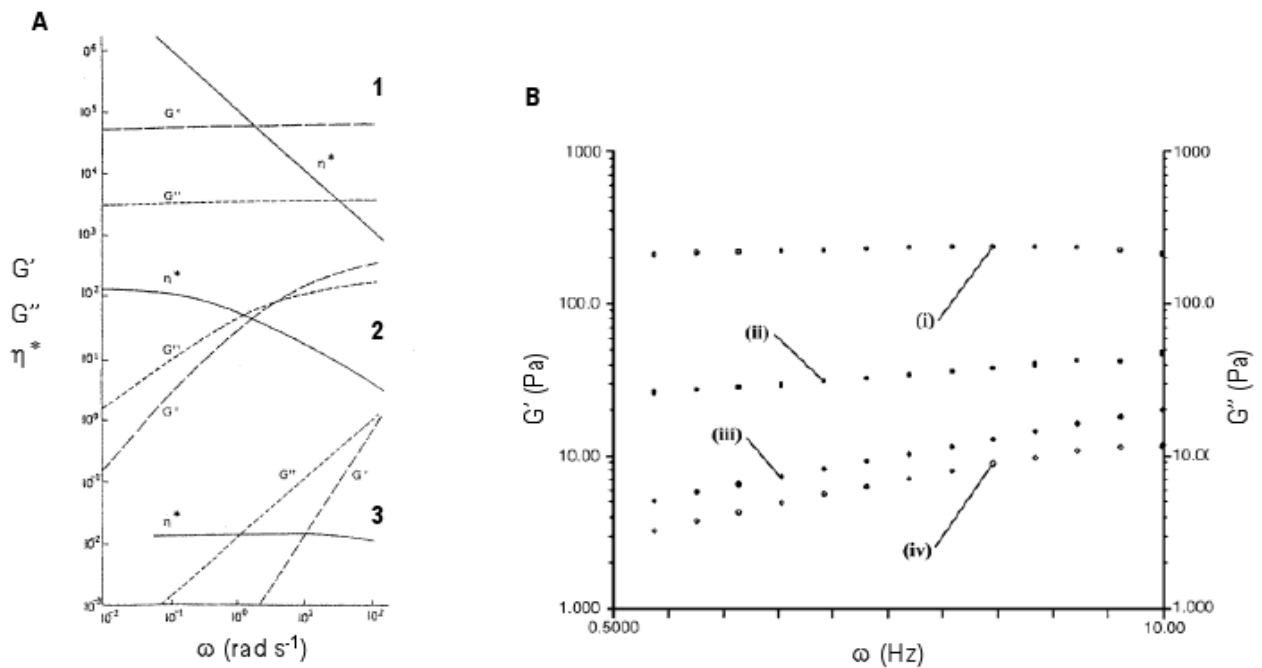


Figure 27 : (A) Evolution des modules élastique (G') et visqueux (G'') dans le cas (1) d'un gel, (2) d'une solution viscoélastique et (3) d'une solution diluée. (B) Mesures en mode dynamique dans le domaine linéaire (oscillation entre 0 à 10 Pa) pour des solutions de dextrans synthétisés par *Thio-DSR-S varden $\Delta 4N-His$* , variant tronqué de DSR-S (voir partie I, paragraphe I.3.5), à partir d'extraits cellulaires (\circ (i) et \bullet (ii); $G' < G''$ comportement typique d'une solution macromoléculaire; solution concentrée à 5 %) et à partir de l'enzyme pure (\square (iv) and \blacksquare (iii); $G' > G''$ comportement d'un gel; solution concentrée à 0,4 %) (Moulis *et al.*, 2006a).

Comparé au dextrane natif de *Ln. mesenteroides* NRRL B-512F, l'alternane natif de *Ln. citreum* B-1355 a une viscosité ~ 100 fois moins élevée pour des solutions concentrées à 5 % (p/v) (Côté, 1992). Cependant, des alternanes de poids moléculaire plus faibles ($< 10^6$), ont une forte solubilité dans l'eau, ce qui permet d'obtenir des solutions très concentrées (> 50 % (p/v)) dont les propriétés rhéologiques sont similaires à celles de la gomme arabique (Côté, 1992).

III.2.2 Propriété thermique : *Transition vitreuse*

Pour un polymère amorphe, ou pour des domaines amorphes d'un composé semi-cristallin, la transition vitreuse correspond au passage d'un état vitreux à un état caoutchoutique. Elle est induite par une élévation de température et s'étale sur un domaine de température plus ou moins large (Figure 28). La température de transition vitreuse T_g peut alors être déterminée en début ($T_{g\text{onset}}$), au milieu ($T_{g\text{midpoint}}$) ou en fin de transition ($T_{g\text{endset}}$). Elle est le plus souvent déterminée au point d'inflexion ($T_{g\text{midpoint}}$) de la courbe calorimétrique représentant l'évolution de la capacité calorifique en fonction de la température (Bizot *et al.*, 1997).

T_g définit donc deux domaines :

- × Au dessus de la T_g le polymère est plastique, c'est l'état « caoutchoutique ». Cet état est associé à l'existence de mouvements moléculaires induits par une relaxation globale (relaxation α) de la chaîne du polymère.

- × En dessous de la T_g le polymère est rigide et cassant, c'est l'état « vitreux ». Dans cet état les déplacements moléculaires sont restreints et se limitent à des mouvements de segments de chaîne (relaxation β) et de groupes hydroxyle (relaxation γ) (Montès and Cavaillé, 1999).

L'ajout de plastifiant abaisse la température de transition vitreuse. L'eau est un excellent plastifiant des α -glucanes dont la T_g dépend fortement de leur teneur en eau et s'étend sur des gammes de températures très étendues, allant de 400°C à l'état sec jusqu'à des valeurs négatives à fortes teneur en eau pour le dextrane T500 (Bizot *et al.*, 1997; Borde *et al.*, 2002). L'eau améliore la flexibilité du polymère par une augmentation du volume libre autour des chaînes et diminue ainsi la valeur de T_g (Figure 29).

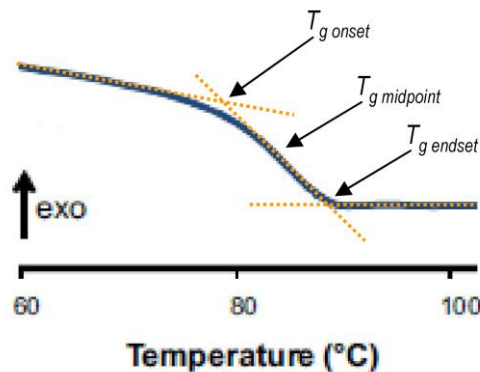


Figure 28 : Thermogramme obtenu par analyse calorifique différentielle illustrant la transition vitreuse pour un échantillon de dextrane de masse molaire de $6.10^3 \text{ g.mol}^{-1}$ dont la teneur en eau est de 8,3 % (g d'eau/100g de masse sèche). Adapté de (Gloria Hernández *et al.*, 2011).

La température de transition vitreuse est influencée par la structure moléculaire. Icoz *et al.* ont étudié l'effet de la masse molaire de dextrans commerciaux sur la T_g (Icoz *et al.*, 2005). Comme le montre la figure 29, la T_g augmente avec la masse molaire des dextrans jusqu'à un plateau. L'augmentation de la T_g aux faibles masses molaires peut s'expliquer par une diminution du volume libre avec l'augmentation du degré de polymérisation, limitant ainsi les mouvements des chaînes du polymère. Aux fortes masses molaires, la T_g devient constante, Slade et Levine ont attribué cette indépendance à la formation d'enchevêtrements entre les chaînes, au-delà d'une certaine longueur (Slade and Levine,

1991). Pour les dextrans, la T_g critique est atteinte dans une gamme de masses molaires entre $2,3 \cdot 10^4$ et $3 \cdot 10^4$ g.mol⁻¹ (Figure 29).

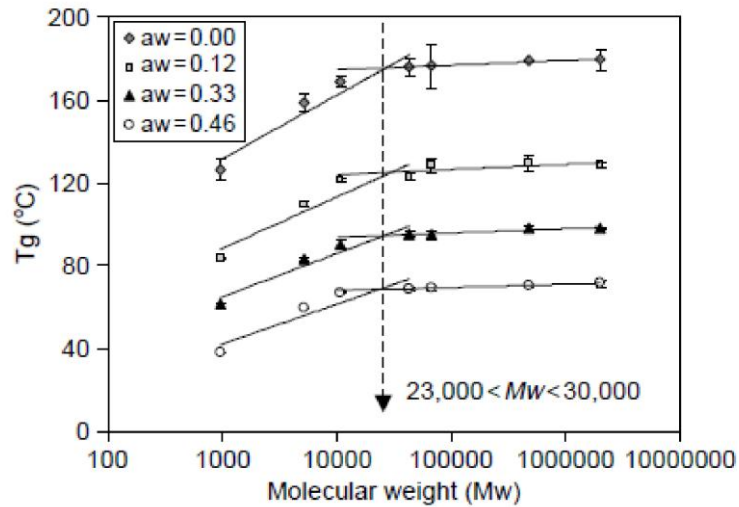


Figure 29 : Effet de la masse molaire et de la teneur en eau sur la T_g . Tiré de (Icoz *et al.*, 2005).

La nature des liaisons osidiques a aussi un rôle dans la valeur de T_g . Les liaisons de type $\alpha(1 \rightarrow 6)$ possèdent trois axes de rotation ce qui apporte plus de flexibilité aux chaînes et donc tend à diminuer la T_g (Bizot *et al.*, 1997). De même, des structures linéaires apportent moins de contraintes stériques facilitant ainsi les mouvements de chaînes (Borde *et al.*, 2002). Cependant, la présence de branchements abaisse également la T_g (Bizot *et al.*, 1997). En effet, l'amylopectine et le glycogène présentent des T_g inférieures à celle de l'amylose. Ceci est dû à l'effet plastifiant des liaisons $\alpha(1 \rightarrow 6)$ présentes au niveau des ramifications. De même, la longueur des branchements influence la T_g (Bizot *et al.*, 1997).

III.3 Applications des α -glucanes

III.3.1 Applications des dextrans

La production commerciale de dextrane a été initiée après la seconde guerre mondiale pour répondre à des besoins de substitution en plasma sanguin. La taille optimale des dextrans utilisés pour ce genre d'application, également appelé dextrans « cliniques », est de l'ordre de 40 à 100 kDa. En effet, les dextrans de taille trop faible sont rapidement éliminés de la circulation sanguine par les reins et sont donc inefficaces d'un point de vue thérapeutique. A l'inverse, les dextrans de très haute masse

moléculaire peuvent interférer avec le processus naturel de coagulation du sang et se révéler dangereux (Naessens *et al.*, 2005). Ainsi, l'intérêt des dextrans cliniques est dû à leur capacité de rétention d'eau (~ 20 ml d'eau/g de dextrane). De même, le dextrane de 40 kDa améliorerait le flux sanguin, probablement par réduction de la viscosité du sang et inhibition de l'agrégation des érythrocytes (de Belder, 1996). L'intérêt du dextrane produit par *Ln. mesenteroides* NRRL B-512F dans les applications médicales provient de son faible degré d'allergénicité et de sa forte solubilité. Cependant, une des conséquences de l'injection de dextrane dans le sang est le risque d'hypersensibilité immunitaire (~ 20 % de la population), par réaction anaphylactique, liée à la structure du polymère (Naessens *et al.*, 2005). Ce risque peut être réduit du facteur 35 en pré-injectant aux patients un dextrane de faible poids moléculaire (10 kDa) comme haptène monovalent (Messmer *et al.*, 1980).

Des dérivés de dextrans ont également été développés. Ils sont obtenus par fonctionnalisation des groupes hydroxyle libres. La plus importante utilisation de dérivés de dextrane concerne les supports chromatographiques commercialisés sous la marque Sephadex® (Flodin and Porath, 1961). Ce sont des gels hydrophiles insolubles possédant des pores de tailles variables et parfaitement contrôlées, utilisés pour la purification et la séparation de macromolécules comme les protéines, polysaccharides et acides nucléiques. Les gels sont obtenus par co-réticulation, à différents degrés, du dextrane avec de l'épichlorhydrine, balayant ainsi un large spectre de limites d'exclusion moléculaire. Des dérivés de la matrice Sephadex® initiale ont également été développés pour permettre d'autres techniques de séparation (interactions hydrophobes, échange d'ions, etc.) (Leathers, 2005).

D'autres dérivés comme les dextrane-sulfates ont par exemple une activité anticoagulante reconnue et sont utilisés comme substitut de l'héparine. Des études ont montré que les propriétés anticoagulantes sont plus fortes pour des dextrans sulfates de faibles masses molaires (Heinze *et al.*, 2006). Ces dérivés ont également été étudiés comme agents antiviraux, et particulièrement dans le traitement contre le virus de l'immunodéficience humaine (VIH) ou le virus de l'herpès (Piret *et al.*, 2000). Cette démonstration a fait l'objet d'un dépôt de brevet (Hershline, 2004). Par ailleurs, plusieurs travaux récents portent sur la formation d'hydrogels de dérivés de dextrane. Ces derniers trouvent des applications dans le génie tissulaire, dont le but est de maintenir, réparer ou reconstruire des tissus ou organes endommagés (Sun *et al.*, 2009). Les hydrogels sont également des transporteurs de protéines très efficaces (Van Tomme and Hennink, 2007). Un autre dérivé, le mercaptodextrane, a une forte affinité pour les ions métalliques lourds (argent, mercure, cuprique et aurique). Cette affinité, couplée à sa faible toxicité, suggère une utilisation possible du mercaptodextrane lors d'intoxications aiguës aux métaux lourds (Jellum *et al.*, 1973; Naessens *et al.*, 2005).

Les dextrans sont aussi connus pour leurs propriétés gélifiantes, texturantes, stabilisantes et épaississantes, exploitées de manière marginale dans le secteur de l'agroalimentaire en comparaison avec d'autres polysaccharides extraits de plantes (amidon, alginates, carraghénanes...) ou d'origine microbienne (xanthane par exemple). Cependant, les dextrans de faible degré de polymérisation (isomalto-oligosaccharides composés de 2 à 10 unités glucosyle) font l'objet d'un engouement croissant pour l'exploitation de leurs propriétés prébiotiques.

La notion de prébiotique a été formalisée en 1995 par Gibson and Roberfroid, et revisitée récemment par Roberfroid, donnant la définition suivante : « Un prébiotique est un ingrédient sélectivement fermenté qui permet des changements spécifiques, dans la composition et/ou l'activité de la microflore gastro-intestinale, conférant des avantages à la santé et au bien-être de l'hôte » (Gibson and Roberfroid, 1995; Roberfroid, 2007). Un composé prébiotique répond à trois critères : (1) il ne doit pas être métabolisé par l'hôte dans la partie supérieure du tractus digestif, de façon à atteindre le microbiome intestinal sous forme intacte ou partiellement dégradé; (2) il doit être fermenté par les bactéries intestinales, et (3) stimuler de manière sélective la croissance et/ou l'activité des bactéries intestinales considérées comme bénéfiques pour l'hôte (Roberfroid, 2007).

Les liaisons $\alpha(1\rightarrow6)$ étant plus difficilement dégradées par les enzymes digestives que les liaisons $\alpha(1\rightarrow4)$ (digérées par les α -amylases humaines), les isomalto-oligosaccharides (IMOS) arrivent partiellement intacts au niveau de l'intestin (Kohmoto *et al.*, 1992). Des études menées sur 7 patients âgés, atteints de constipations chroniques, traités à raison de 10 g d'IMOS ingérés par jour sur une période d'un mois, ont montré une amélioration nette du transit intestinal et des problèmes de constipation (Chen *et al.*, 2001). Les auteurs montrent ainsi que l'ingestion d'IMOS augmente l'activité bactérienne du colon et favorise le développement de *Bifidobactéries* (Kaneko, 1994; Kohmoto *et al.*, 1984; Kohmoto *et al.*, 1991). Des observations similaires ont été rapportées par Wang *et al.* sur des patients souffrant de constipation et d'hypercholestérolémie (Wang *et al.*, 2001). De plus, la consommation d'IMOS abaisserait de 18 % le taux de cholestérol total et de triglycérides dans le sang. Cependant, Chen *et al.* n'observent pas ces effets sur des patients présentant une cholestérolémie normale (Chen *et al.*, 2001). Plus récemment, l'addition d'IMOS dans le régime alimentaire du porc, à raison de 25 g/kg, a eu pour effet de réduire la concentration en glucose dans le sang (Li *et al.*, 2009). Cette propriété permettrait potentiellement de prévenir et/ou de traiter le diabète de type II (diabète non-insulino-dépendant).

Une autre catégorie d'oligosaccharides présente des effets prébiotiques similaires. Il s'agit des $\alpha(1\rightarrow6)$ -gluco-oligosaccharides (GOS) branchés par des liaisons osidiques de type $\alpha(1\rightarrow2)$, synthétisés à partir de saccharose et de maltose en utilisant l'activité enzymatique de la souche *Leuconostoc citreum* NRRL B-1299 (Dols *et al.*, 1997b; Remaud-Simeon *et al.*, 1994). Les liaisons

$\alpha(1\rightarrow2)$, rarement rencontrées dans la nature, confèrent aux GOS branchés une forte résistance à l'hydrolyse par les enzymes humaines (Flickinger *et al.*, 2000; Valette *et al.*, 1993). L'administration de GOS branchés à des souris obèses dans un état pré-diabétique a pour effet de diminuer l'intolérance au glucose. Pour des souris obèses et diabétiques, l'administration d' $\alpha(1\rightarrow6)/\alpha(1\rightarrow2)$ -GOS induit une diminution du taux de glucose sanguin après administration de glucose (Boucher *et al.*, 2003). Les auteurs émettent l'hypothèse que les GOS branchés auraient un effet sur la production d'incrétine, peptide insulinothropique produit par les cellules endocrines de la muqueuse intestinale, qui stimule la production d'insuline par les cellules beta du pancréas.

III.3.2 Applications de l'alternane

La structure très particulière de ce polymère lui confère une résistance aux glycoside-hydrolases humaines. De par sa faible viscosité et sa grande solubilité dans l'eau, Côté suggère dès 1992 d'utiliser l'alternane dans le domaine agroalimentaire, comme agent de charge non digestible ou faiblement calorique. D'autres auteurs proposent d'utiliser ce polymère dans diverses applications comme par exemple, agent dilatateur de vaisseaux sanguins, transporteur de composés actifs ou bien comme additif dans l'industrie du textile ou de la cosmétique (Leathers *et al.*, 1997; Leathers *et al.*, 1995; López-Munguía *et al.*, 1993).

Les oligoalternanes synthétisés à partir de l'alternane-saccharase, en présence de saccharose et d'accepteurs tel que le maltose, le gentibiose, le maltitol, le melibiose ou le raffinose présentent des propriétés prébiotiques en favorisant la croissance et l'activité de *Bifidobactéries* au détriment de bactéries pathogènes et coliformes (Holt *et al.*, 2005; Holt *et al.*, 2008). En 2002, Leathers *et al.* ont déposé un brevet pour la synthèse d'alternanes de poids moléculaires plus faibles, obtenus à partir de l'alternane produit par fermentation couplée de *Ln. citreum* B-1355 et de diverses souches de *Penicillium* sp. (Leathers *et al.*, 2002a; Leathers *et al.*, 2002b). Ces dérivés alternanes de faibles poids moléculaires trouveraient donc des applications comme agents texturants (Leathers *et al.*, 2003). Récemment le procédé d'obtention de ces alternanes a été optimisé par le remplacement des souches de *Penicillium* sp. par de la dextranase commerciale de *Chaetomium erraticum* (Leathers *et al.*, 2010).

III.3.3 Applications des mutanes

La forte proportion de liaisons $\alpha(1\rightarrow3)$ confère aux mutanes la propriété d'être insolubles dans l'eau. Connus pour leur rôle dans la formation de la plaque dentaire, les mutane-saccharases sont de ce fait, des cibles de choix pour la synthèse de vaccins anti-caries (Koga *et al.*, 2002).

III.3.4 Applications des amylopolysaccharides

L'amylosaccharase de *Neisseria polysaccharea* est non-seulement capable de synthétiser des $\alpha(1\rightarrow4)$ -glucanes linéaires de type amylose, mais aussi, dans certaines conditions réactionnelles, de modifier la structure d' $\alpha(1\rightarrow4)/\alpha(1\rightarrow6)$ -glucanes branchés exogènes, tels l'amylopectine (le constituant ramifié de l'amidon), les β -limites-dextrines (issues de l'hydrolyse de l'amylopectine par la β -amylase), et le glycogène (le polysaccharide de réserve énergétique des animaux). La morphologie, la structure et les propriétés (viscosité, cristallinité, teneur en amidon résistant) de ces molécules peuvent alors être contrôlées en variant les rapports entre le donneur et l'accepteur d'unités glucosyle utilisé, et la durée de la réaction (Potocki-Veronese *et al.*, 2005). Au même titre que l'amidon des plantes, ces polysaccharides pourraient donc être utilisés dans l'industrie agroalimentaire pour leurs propriétés texturantes, de complexation de principes actifs et en tant qu'amidons résistants présentant des propriétés prébiotiques.

Les glucane-saccharases sont des enzymes très versatiles, capables de synthétiser un très large panel de structures de polysaccharides et oligosaccharides. Le développement et la valorisation de ces produits présupposent de pouvoir établir les relations entre leur structure et leurs propriétés physico-chimiques et biologiques, qui ont été peu explorées. En effet, aujourd'hui, seuls les dextrans sont commercialisés, principalement pour des applications analytiques et thérapeutiques. L'autorisation de mise sur le marché européen de dextrans produits par la souche *Leuconostoc mesenteroides* NRRL B-512F pour des applications alimentaires est d'ailleurs relativement récente (janvier 2001), et est restreinte à leur utilisation en panification à hauteur de 5 % maximum. Les autres α -glucanes et gluco-oligosaccharides produits par les GS suscitent eux aussi un intérêt croissant pour des applications en tant qu'aliments fonctionnels de types prébiotiques et agents de charge peu caloriques. Néanmoins, la diversification des applications potentielles de ces polymères de structures variées, notamment en tant que matériaux bio-sourcés, est actuellement limitée par le manque de données relatives à leurs propriétés physico-chimiques.

Objectifs de la thèse

Malgré des avancées notoires ces dernières années dans la compréhension des relations structure-fonction des glucane-saccharases de la famille GH70, les facteurs impliqués dans la spécificité de ces enzymes sont encore très mal compris. Ceci est lié à la grande diversité de séquences des enzymes de cette famille, et à l'absence jusqu'en juillet 2010, de structure tri-dimensionnelle. De plus, les travaux d'ingénierie de la spécificité des glucane-saccharases ont été restreints à des approches rationnelles ou semi-rationnelles générant des espaces de séquences à explorer très limités, car il n'existait aucune méthode de criblage à haut-débit de la spécificité de liaison des CAZymes. Enfin, les relations structure/propriétés des α -glucanes synthétisés par les glucane-saccharases ont été également très peu explorées, et se limitent presque essentiellement aux travaux portant sur les dextrans commerciaux.

Les travaux effectués aux cours de cette thèse, réalisée dans le cadre du Réseau de Recherche INRA-CNRS « Intégrer une approche sociétale renouvelée de la chimie: les ressources renouvelables comme matières premières et sources de nouveaux produits et matériaux », avaient pour objectif de générer des outils de glycosylation originaux capables de synthétiser des α -glucanes et gluco-oligosaccharides de structures et de propriétés originales, afin d'envisager de nouvelles applications de ces biopolymères.

Les enjeux étaient multiples. Il s'agissait d'adapter les techniques d'ingénierie combinatoire à la création et au criblage de grandes banques de glucane-saccharases de spécificités variées. Sur le plan fondamental, l'objectif était d'identifier de nouveaux déterminants structuraux de la spécificité des glucane-saccharases GH70, et d'approfondir les connaissances structure-propriétés des α -glucanes.

Les travaux réalisés ont donc visé à :

- (i) Développer une stratégie multi-étapes de sélection et de criblage permettant d'isoler des variants de glucane-saccharases actifs et de les discriminer sur la base de leurs spécificités de liaison.
- (ii) Générer par ingénierie combinatoire une diversité importante sur des espaces de séquences encore non explorés et caractériser les propriétés des variants isolés, afin

d'identifier le rôle des résidus mutés dans le contrôle de la spécificité des glucane-saccharases.

- (iii) Caractériser la structure et les propriétés physico-chimiques des nouveaux biopolymères formés, pour envisager de nouvelles applications pour les α -glucanes.

Les trois chapitres suivants font la synthèse des principaux résultats obtenus.

Références bibliographiques

A

Agresti, J.J., Antipov, E., Abate, A.R., Ahn, K., Rowat, A.C., Baret, J.C., Marquez, M., Klibanov, A.M., Griffiths, A.D., Weitz, D.A., 2010. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *PNAS* **107**, 4004-4009.

Aharoni, A., Amitai, G., Bernath, K., Magdassi, S., Tawfik, D.S., 2005a. High-throughput screening of enzyme libraries: thiolactonases evolved by fluorescence-activated sorting of single cells in emulsion compartments. *Chemistry & Biology* **12**, 1281.

Aharoni, A., Griffiths, A.D., Tawfik, D.S., 2005b. High-throughput screens and selections of enzyme-encoding genes. *Curr. Opin. Chem. Biol.* **9**, 210.

Aharoni, A., Thieme, K., Chiu, C.P.C., Buchini, S., Lairson, L.L., Chen, H., Strynadka, N.C.J., Wakarchuk, W.W., Withers, S.G., 2006. High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* **3**, 609.

Albenne, C., Skov, L.K., Mirza, O., Gajhede, M., Feller, G., D'Amico, S., Andre, G., Potocki-Veronese, G., van der Veen, B.A., Monsan, P., Remaud-Simeon, M., 2004. Molecular basis of the amylose-like polymer formation catalyzed by *Neisseria polysaccharea* amylosucrase. *J Biol Chem* **279**, 726-734.

Amstutz, P., Pelletier, J.N., Guggisberg, A., Jerminus, L., Cesaro-Tadic, S., Zahnd, C., Plückthun, A., 2002. *In vitro* selection for catalytic activity with ribosome display. *Journal of the American Chemical Society* **124**, 9396.

André, I., Potocki Véronèse, G., Morel, S., Monsan, P., Remaud-Siméon, M., 2010. Sucrose-utilizing transglucosidases for biocatalysis. *Rauter AP, Queneau Y, Vogel P eds.; In Topics in Current Chemistry: Carbohydrates in sustainable development.*

Arguello-Morales, M., Sanchez-Gonzalez, M., Canedo, M., Quirasco, M., Farres, A., Lopez-Munguia, A., 2005. Proteolytic modification of *Leuconostoc mesenteroides* B-512F dextranase. *Antonie Van Leeuwenhoek* **87**, 131-141.

Argüello-Morales, M.A., Remaud-Simeon, M., Pizzut, S., Sarcabal, P., Willemot, R.M., Monsan, P., 2000. Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. *FEMS Microbiol. Lett.* **182**, 81-85.

B

Bai, Y., Feng, H., 2004. Selection of stably folded proteins by phage-display with proteolysis. *Eur. J. Biochem.* **271**, 1609.

Baldwin, A.J., Busse, K., Simm, A.M., Jones, D.D., 2008. Expanded molecular diversity generation during directed evolution by trinucleotide exchange (TriNEx). *Nucl. Acids Res.* **36**.

Banas, J.A., Russell, R.R., Ferretti, J.J., 1990. Sequence analysis of the gene for the glucan-binding protein of *Streptococcus mutans* Ingbritt. *Infect. Immun.* **58**, 667-673.

- Ben-David, A., Shoham, G., Shoham, Y.**, 2008. A universal screening assay for glycosynthases: directed evolution of glycosynthase XynB2 (E335G) suggests a general path to enhance activity. *Chemistry & Biology* **15**, 546.
- Bernath, K., Hai, M., Mastrobattista, E., Griffiths, A.D., Magdassi, S., Tawfik, D.S.**, 2004. *In vitro* compartmentalization by double emulsions: sorting and gene enrichment by fluorescence activated cell sorting. *Analytical Biochemistry* **325**, 151.
- Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., Remaud-Simeon, M.**, 2006. *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr. Res.* **341**, 855.
- Bizot, H., Le Bail, P., Leroux, B., Davy, J., Roger, P., Buleon, A.**, 1997. Calorimetric evaluation of the glass transition in hydrated, linear and branched polyanhydroglucose compounds. *Carbohydrate Polymers* **32**, 33-50.
- Bloom, J.D., Labthavikul, S.T., Otey, C.R., Arnold, F.H.**, 2006. Protein stability promotes evolvability. *Proceedings of the National Academy of Sciences* **103**, 5869-5874.
- Boersma, Y.L., Dröge, M.J., Quax, W.J.**, 2007. Selection strategies for improved biocatalysts. *FEBS Journal* **274**, 2181.
- Boersma, Y.L., Dröge, M.J., van der Sloot, A.M., Pijning, T., Cool, R.H., Dijkstra, B.W., Quax, W.J.**, 2008. A novel genetic selection system for improved enantioselectivity of *Bacillus subtilis* Lipase A. *ChemBioChem* **9**, 1110.
- Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J.**, 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769-781.
- Borde, B., Bizot, H., Vigier, G., Buleon, A.**, 2002. Calorimetric analysis of the structural relaxation in partially hydrated amorphous polysaccharides. I. Glass transition and fragility. *Carbohydr. Polym.* **48**, 83-96.
- Bouayadi, K., Kharrat, H., Louat, T., Servant, L., Cazaux, C., Hoffmann, J.S.**, 2002. Use of mutagenic polymerase for producing random mutations. *Patent WO/2002/0238756*.
- Boucher, J., Daviaud, D., Simeon-Remaud, M., Carpenne, C., Saulnier-Blache, J.S., Monsan, P., Valet, P.**, 2003. Effect of non-digestible gluco-oligosaccharides on glucose sensitivity in high fat diet fed mice. *J. Physiol. Biochem.* **59**, 169-173.
- Bozonnet, S., Dols-Laffargue, M., Fabre, E., Pizzut, S., Remaud-Simeon, M., Monsan, P., Willemot, R.M.**, 2002. Molecular characterization of DSR-E, an α -1,2 linkage-synthesizing dextranucrase with two catalytic domains. *J. Bacteriol.* **184**, 5753-5761.
- Braswell, E., Goodman, A., Stern, K.G.**, 1962. Studies on the enzymatic synthesis of dextran. Part II. *J. Pol. Sci.* **61**, 143.
- Braswell, E., Stern, K.G.**, 1959. Studies on the enzymatic synthesis of dextran. *J. Pol. Sci.* **41**, 467.

Brison, Y., 2010. Contribution à la compréhension du mécanisme de formation de dextrans ou glucooligosaccharides ramifiés en α -(1→2) par l'enzyme GBD-CD2: études cinétique et structurale. **Thèse de doctorat, INSA, Toulouse.**

C

Cadwell, R.C., Joyce, G.F., 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**, 28-33.

Campbell, J.A., Davies, G.J., Bulone, V., Henrissat, B., 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **326**, 929-942.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucl. Acids Res.* **37**, 233-238.

Carrasco, F., Chornet, E., Overend, R.P., Costa, J., 1989. A generalized correlation for the viscosity of dextrans in aqueous solutions as a function of temperature, concentration, and molecular weight at low shear rates. *J. Appl. Pol. Sci.* **37**, 2087.

Champion, E., 2008. Glucane-saccharases optimisées pour la glucosylation de molécules intervenant dans la composition de vaccins. **Thèse de doctorat, INSA, Toulouse.**

Champion, E., André, I., Moulis, C., Boutet, J., Descroix, K., Morel, S., Monsan, P.M., Mulard, L.A., Remaud-Siméon, M., 2009. Design of α -transglucosidases of controlled specificity for programmed chemoenzymatic synthesis of antigenic oligosaccharides. *J. Am. Chem. Soc.* **131**, 7379.

Champion, E., Moulis, C., Morel, S., Mulard, L., Monsan, P., Remaud-Siméon, M., André, I., 2010. A pH-based high-throughput screening of sucrose-utilizing transglucosidases for the development of enzymatic glucosylation tools. *ChemCatChem* **2**, 969.

Cheetham, N.W., Fiala-Beer, E., Walker, G.J., 1991. Dextran structural details from high-field proton NMR spectroscopy. *Carbohydr. Polym.* **14**, 149-158.

Chen, H.L., Lu, Y.H., Lin, J.J., Ko, L.Y., 2001. Effects of isomalto-oligosaccharides on bowel functions and indicators of nutritional status in constipated elderly men. *J. Am. Coll. Nutr.* **20**, 44-49.

Cline, J., Braman, J.C., Hogrefe, H.H., 1996. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* **24**, 3546-3551.

Coco, W.M., Levinson, W.E., Crist, M.J., Hektor, H.J., Darzins, A., Pienkos, P.T., Squires, C.H., Monticello, D.J., 2001. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat. Biotech.* **19**, 354.

Cohen, H.M., Tawfik, D.S., Griffiths, A.D., 2004. Altering the sequence specificity of HaeIII methyltransferase by directed evolution using *in vitro* compartmentalization. *Protein Eng Des Sel* **17**, 3-11.

Colby, D.W., Kellogg, B.A., Graff, C.P., Yeung, Y.A., Swers, J.S., Wittrup, K.D., 2004. Engineering antibody affinity by yeast surface display. *Methods in Enzymology*. Academic Press, p. 348.

Côté, G.L., 1992. Low-viscosity α -D-glucan fractions derived from sucrose which are resistant to enzymatic digestion. *Carbohydr. Polym.* **19**, 249.

Côté, G.L., Dunlap, C.A., 2003. Alternansucrase acceptor reactions with methyl hexopyranosides. *Carbohydr. Res.* **338**, 1961.

Côté, G.L., Dunlap, C.A., Vermillion, K.E., 2009. Glucosylation of raffinose via alternansucrase acceptor reactions. *Carbohydr. Res.* **344**, 1951.

Côté, G.L., Robyt, J.F., 1984. The formation of α -(1 \rightarrow 3) branch linkages by a glucansucrase from *Streptococcus mutans* 6715 producing a soluble glucan. *Carbohydr. Res.* **127**, 95.

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

Coutinho, P.M., Henrissat, B., 1999. Carbohydrate-active enzymes: an integrated database approach. In "Recent Advance in Carbohydrate Bioengineering". H.J. Gilbert, G. Davies, B. Henrissat and B. Svensson eds., The Royal Society of Chemistry, Cambridge, 3-12.

Cramer, A., Raillard, S.A., Bermudez, E., Stemmer, W.P., 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288.

D

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

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

de Belder, A.N., 1990. Dextran. *Uppsala Sweden: Pharmacia*.

de Belder, A.N., 1996. Medical applications of dextran and its derivatives. In S. Dumitriu ed., *Polysaccharides: in medicinal applications*. Marcel Dekker, Inc., New York, 275-296.

Delagrave, S., Murphy, D.J., Pruss, J.L.R., Maffia, A.M., Marrs, B.L., Bylina, E.J., Coleman, W.J., Grek, C.L., Dilworth, M.R., Yang, M.M., Youvan, D.C., 2001. Application of a very high-throughput digital imaging screen to evolve the enzyme galactose oxidase. *Protein Eng.* **14**, 261-267.

Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L., Neri, P., Winter, G., Neri, D., 1999. A strategy for the isolation of catalytic activities from repertoires of enzymes displayed on phage. *J. Mol. Biol.* **286**, 617.

DeSantis, G., Liu, J., Clark, D.P., Heine, A., Wilson, I.A., Wong, C.-H., 2003a. Structure-Based mutagenesis approaches toward expanding the substrate specificity of D-2-Deoxyribose-5-phosphate aldolase. *Bioorg. Med. Chem.* **11**, 43.

DeSantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z., Tomlinson, G., Huang, H., Tan, X., Bibbs, L., Chen, P., Kretz, K., Burk, M., 2003b. Creation of a productive, highly enantioselective nitrilase through Gene Site Saturation Mutagenesis (GSSM). *J. Am. Chem. Soc.* **125**, 11476.

Devulapalle, K.S., Goodman, S.D., Gao, Q., Hemsley, A., Mooser, G., 1997. Knowledge-based model of a glucosyltransferase from the oral bacterial group of *mutans streptococci*. *Protein Sci.* **6**, 2489.

Dols, M., Remaud-Simeon, M., Vignon, M., Monsan, P., 1997a. Characterisation of dextransucrases from *leuconostoc mesenteroides* NRRL B-1299. *Appl. Biochem. Biotechnol.* **62**, 47-59.

Dols, M., Simeon, M.R., Willemot, R.M., Vignon, M.R., Monsan, P.F., 1997b. Structural characterization of the maltose acceptor-products synthesized by *Leuconostoc mesenteroides* NRRL B-1299 dextransucrase. *Carbohydr. Res.* **305**, 549-559.

Dupret, D., Lefevre, F., Masson, J.M., 2002. Template-mediated, ligation-oriented method of nonrandomly shuffling polynucleotides. *Patent WO/2002/086121*.

E

Eckert, K.A., Kunkel, T.A., 1990. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* **18**, 3739-3744.

Emond, S., Andre, I., Jaziri, K., Potocki-Veronese, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., Remaud-Simeon, M., 2008a. Combinatorial engineering to enhance thermostability of amylosucrase. *Protein Sci.* **17**, 967-976.

Emond, S., Mondeil, S., Jaziri, K., André, I., Monsan, P., Remaud-Siméon, M., Potocki-Véronèse, G., 2008b. Cloning, purification and characterization of a thermostable amylosucrase from *Deinococcus geothermalis*. *FEMS Microbiol. Lett.* **285**, 25.

Emond, S., Mondon, P., Pizzut-Serin, S., Douchy, L., Crozet, F., Bouayadi, K., Kharrat, H., Potocki-Véronèse, G., Monsan, P., Remaud-Simeon, M., 2008c. A novel random mutagenesis approach using human mutagenic DNA polymerases to generate enzyme variant libraries. *Protein Eng. Des. Sel.* **21**, 267-274.

Emond, S., Potocki-Véronèse, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., Remaud-Simeon, M., 2007. Optimized and automated protocols for high-throughput screening of amylosucrase libraries. *Journal of Biomolecular Screening* **12**, 715-723.

F

Fabre, E., Bozonnet, S., Arcache, A., Willemot, R.M., Vignon, M., Monsan, P., Remaud-Simeon, M., 2005. Role of the two catalytic domains of DSR-E dextransucrase and their involvement in the formation of highly alpha-1,2 branched dextran. *J. Bacteriol.* **187**, 296-303.

Falconer, D.J., Mukerjea, R., Robyt, J.F., 2011. Biosynthesis of dextrans with different molecular weights by selecting the concentration of *Leuconostoc mesenteroides* B-512FMC dextransucrase, the sucrose concentration, and the temperature. *Carbohydr. Res.* **346**, 280.

Fernandez-Gacio, A., Uguen, M., Fastrez, J., 2003. Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol.* **21**, 408.

Ferretti, J.J., Gilpin, M.L., Russell, R.R., 1987. Nucleotide sequence of a glucosyltransferase gene from *Streptococcus sobrinus* Mfe28. *J. Bacteriol.* **169**, 4271-4278.

Flickinger, E.A., Wolf, B.W., Garleb, K.A., Chow, J., Leyer, G.J., Johns, P.W., Fahey, G.C., 2000. Glucose-based oligosaccharides exhibit different *in vitro* fermentation patterns and affect *in vivo* apparent nutrient digestibility and microbial populations in dogs. *J. Nutr.* **130**, 1267-1273.

Flodin, P., Porath, J., 1961. Process of separating materials having different molecular weight and dimensions. **U. S. patent 3002823**.

Freeman, A., Cohen-Hadar, N., Abramov, S., Modai-Hod, R., Dror, Y., Georgiou, G., 2004. Screening of large protein libraries by the "cell immobilized on adsorbed bead" approach. *Biotechnology and Bioengineering* **86**, 196.

Fromant, M., Blanquet, S., Plateau, P., 1995. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal. Biochem.* **224**, 347.

Fujii, I., Fukuyama, S., Iwabuchi, Y., Tanimura, R., 1998. Evolving catalytic antibodies in a phage-displayed combinatorial library. *Nat. Biotech.* **16**, 463.

Fujiwara, T., Hoshino, T., Ooshima, T., Sobue, S., Hamada, S., 2000. Purification, characterization, and molecular analysis of the gene encoding glucosyltransferase from *Streptococcus oralis*. *Infect. Immun.* **68**, 2475-2483.

Funane, K., Ishii, T., Ono, H., Kobayashi, M., 2005. Changes in linkage pattern of glucan products induced by substitution of Lys residues in the dextransucrase. *FEBS Lett.* **579**, 4739.

Funane, K., Ishii, T., Terasawa, K., Yamamoto, T., Kobayashi, M., 2004. Construction of chimeric glucansucrases for analyzing substrate-binding regions that affect the structure of glucan products. *Biosci. Biotechnol. Biochem.* **68**, 1912-1920.

Funane, K., Mizuno, K., Takahara, H., Kobayashi, M., 2000. Gene encoding a dextransucrase-like protein in *Leuconostoc mesenteroides* NRRL B-512F. *Biosci. Biotechnol. Biochem.* **64**, 29-38.

G

Germaine, G.R., Schachtele, C.F., 1976. *Streptococcus mutans* dextransucrase: mode of interaction with high-molecular-weight dextran and role in cellular aggregation. *Infect. Immun.* **13**, 365-372.

Gibbs, M.D., Nevalainen, K.M.H., Bergquist, P.L., 2001. Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling. *Gene* **271**, 13-20.

Gibson, G.R., Roberfroid, M.B., 1995. Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *J. Nutr.* **125**, 1401-1412.

Giffard, P.M., Jacques, N.A., 1994. Definition of a fundamental repeating unit in *Streptococcal* glucosyltransferase glucan-binding regions and related sequences. *J. Dent. Res.* **73**, 1133-1141.

Giffard, P.M., Simpson, C.L., Milward, C.P., Jacques, N.A., 1991. Molecular characterization of a cluster of at least two glucosyltransferase genes in *Streptococcus salivarius* ATCC 25975. *J. Gen. Microbiol.* **137**, 2577-2593.

Gilmore, K.S., Russell, R.R., Ferretti, J.J., 1990. Analysis of the *Streptococcus downei* gtfS gene, which specifies a glucosyltransferase that synthesizes soluble glucans. *Infect. Immun.* **58**, 2452-2458.

Gloria Hernández, H., Livings, S., Aguilera, J.M., Chiralt, A., 2011. Phase transitions of dairy proteins, dextrans and their mixtures as a function of water interactions. *Food Hydrocolloids* **25**, 1311-1318.

Griffiths, A.D., Tawfik, D.S., 2003. Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J* **22**, 24.

Griffiths, A.D., Tawfik, D.S., 2006. Miniaturising the laboratory in emulsion droplets. *Trends Biotechnol.* **24**, 395.

Griffiths, J.S., Cheriyan, M., Corbell, J.B., Pocivavsek, L., Fierke, C.A., Toone, E.J., 2004. A bacterial selection for the directed evolution of pyruvate aldolases. *Bioorg. Med. Chem.* **12**, 4067.

Guizard, C., 1981. Cristallisation et polymorphisme des glucanes liées (1-6). Etude de la structure tridimensionnelle des cristaux par microscopie électronique. **Thèse de doctorat, Université scientifique et médicale & Institut National Polytechnique de Grenoble, Grenoble.**

H

Ha, S.J., Seo, D.H., Jung, J.H., Cha, J., Kim, T.J., Kim, Y.W., Park, C.S., 2009. Molecular cloning and functional expression of a new amylosucrase from *Alteromonas macleodii*. *Biosci. Biotechnol. Biochem.* **73**, 1505-1512.

Hanada, N., Fukushima, K., Nomura, Y., Senpuku, H., Hayakawa, M., Mukasa, H., Shiroza, T., Abiko, Y., 2002. Cloning and nucleotide sequence analysis of the *Streptococcus sobrinus* gtfU gene that produces a highly branched water-soluble glucan. *Biochim. Biophys. Acta, Gen. Subj.* **1570**, 75.

Hanada, N., Isobe, Y., Aizawa, Y., Katayama, T., Sato, S., Inoue, M., 1993. Nucleotide sequence analysis of the gtfT gene from *Streptococcus sobrinus* OMZ176. *Infect. Immun.* **61**, 2096-2103.

Heinze, T., Liebert, T., Heublein, B., Hornig, S., 2006. Functional polymers based on dextran, Polysaccharides II, *Advances In Polymer Science*. Springer-Verlag Berlin, Berlin, pp. 199-291.

Heinzelman, P., Snow, C.D., Wu, I., Nguyen, C., Villalobos, A., Govindarajan, S., Minshull, J., Arnold, F.H., 2009. A family of thermostable fungal cellulases created by structure-guided recombination. *PNAS* **106**, 5610-5615.

Hellmuth, H., Wittrock, S., Kralj, S., Dijkhuizen, L., Hofer, B., Seibel, J., 2008. Engineering the glucansucrase GTFR enzyme reaction and glycosidic bond specificity: toward tailor-made polymer and oligosaccharide products. *Biochemistry* **47**, 6678-6684.

Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280 (Pt 2)**, 309-316.

Henrissat, B., Bairoch, A., 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293 (Pt 3)**, 781-788.

Henrissat, B., Bairoch, A., 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**, 695-696.

Henrissat, B., Davies, G., 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637.

Herman, A., Tawfik, D.S., 2007. Incorporating synthetic oligonucleotides via gene reassembly (ISOR): a versatile tool for generating targeted libraries. *Protein Eng. Des. Sel.* **20**, 219-226.

Hershline, R.K., 2004. Antiviral composition. *US Patent 6821958*.

Hoeller, B.M., Reiter, B., Abad, S., Graze, I., Glieder, A., 2008. Random tag insertions by Transposon Integration mediated Mutagenesis (TIM). *J. Microbiol. Methods* **75**, 251-257.

Holt, S.M., Miller-Fosmore, C.M., Côté, G.L., 2005. Growth of various intestinal bacteria on alternansucrase-derived oligosaccharides. *Lett. Appl. Microbiol.* **40**, 385.

Holt, S.M., Teresi, J.M., Côté, G.L., 2008. Influence of alternansucrase-derived oligosaccharides and other carbohydrates on α -galactosidase and α -glucosidase activity in *Bifidobacterium adolescentis*. *Lett. Appl. Microbiol.* **46**, 73.

I

Icoz, D.Z., Moraru, C.I., Kokini, J.L., 2005. Polymer-polymer interactions in dextran systems using thermal analysis. *Carbohydr. Pol.* **62**, 120-129.

Ioan, C.E., Aberle, T., Burchard, W., 2000a. Light scattering and viscosity behavior of dextran in semidilute solution. *Macromolecules* **34**, 326.

Ioan, C.E., Aberle, T., Burchard, W., 2000b. Structure properties of dextran. 2. Dilute solution. *Macromolecules* **33**, 5730-5739.

Ioan, C.E., Aberle, T., Burchard, W., 2001. Structure properties of dextran. 3. Shrinking factors of individual clusters. *Macromolecules* **34**, 3765-3771.

Ito, K., Ito, S., Shimamura, T., Kawarasaki, Y., Abe, K., Misaka, T., Kobayashi, T., Iwata, S., 2010. Crystallization and preliminary X-ray analysis of a glucansucrase from the dental caries pathogen *Streptococcus mutans*. *Acta Crystallogr., Sect. F: Struct.* **66**, 1086-1088.

Ito, K., Ito, S., Shimamura, T., Weyand, S., Kawarasaki, Y., Misaka, T., Abe, K., Kobayashi, T., Cameron, A.D., Iwata, S., 2011. Crystal structure of glucansucrase from the dental caries pathogen *Streptococcus mutans*. *J. Mol. Biol.* **In Press**.

J

Jeanes, A., 1966. Dextrans. *in: Encyclopedia of polymer science and ingeneering*, (Mark, H. F.; Bikales,

N. M.; Overberger, C. G.; Menges, G.; Kroschwitz, J. L., Eds.), New York: John Wiley & Sons **4**, 752-767.

Jeanes, A., Haynes, W., William, C., Rankin, J., Melvin, E., Austin, M., Clusray, J., Fisher, B., Tsuchiya, H., Rist, C., 1954. Characterisation and classification of dextrans from ninety six strains of bacteria. *J. Am. Chem. Soc.* **76**, 5041.

Jellum, E., Aaseth, J., Eldjarn, L., 1973. Mercaptodextran, a metal-chelating and disulphide-reducing polythiol of high molecular weight. *Biochem. Pharm.* **22**, 1179.

Jensen, M.H., Mirza, O., Albenne, C., Remaud-Simeon, M., Monsan, P., Gajhede, M., Skov, L.K., 2004. Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. *Biochemistry* **43**, 3104.

Joern, J.M., Meinhold, P., Arnold, F.H., 2002. Analysis of shuffled gene libraries. *J Mol Biol* **316**, 643-656.

Joo, H., Arisawa, A., Lin, Z., Arnold, F.H., 1999. A high-throughput digital imaging screen for the discovery and directed evolution of oxygenases. *Chem. Biol.* **6**, 699-706.

Joucla, G., Pizzut, S., Monsan, P., Remaud-Simeon, M., 2006. Construction of a fully active truncated alternansucrase partially deleted of its carboxy-terminal domain. *FEBS lett.* **580**, 763.

Ju, J., Misono, H., Ohnishi, K., 2005. Directed evolution of bacterial alanine racemases with higher expression level. *J. Biosci. Bioeng.* **100**, 246-254.

K

Kaneko, T., 1994. Effects of isomaltooligosaccharides with different degrees of polymerization on human fecal *Bifidobacteria*. *Biosci. Biotechnol. Biochem.* **58**, 2288-2290.

Kang, H.K., Kim, Y.M., Kim, D.M., 2008. Functional, genetic, and bioinformatic characterization of dextranase (DSRBCB4) gene in *Leuconostoc mesenteroides* B-1299CB4. *J. Microbiol. Biotechnol.* **18**, 1050-1058.

Kang, H.K., Oh, J.S., Kim, D., 2009. Molecular characterization and expression analysis of the glucanase DSRWC from *Weissella cibaria* synthesizing a $\alpha(1\rightarrow6)$ glucan. *FEMS Microbiol. Lett.* **292**, 33-41.

Kato, C., Nakano, Y., Lis, M., Kuramitsu, H.K., 1992. Molecular genetic analysis of the catalytic site of *streptococcus mutans* glucosyltransferases. *Biochem. Biophys. Res. Commun.* **189**, 1184-1188.

Kelly, R.M., Leemhuis, H., Gatjen, L., Dijkhuizen, L., 2008. Evolution toward small molecule inhibitor resistance affects native enzyme function and stability, generating acarbose-insensitive cyclodextrin glucanotransferase variants. *J Biol Chem* **283**, 10727-10734.

Kikuchi, M., Ohnishi, K., Harayama, S., 1999. Novel family shuffling methods for the *in vitro* evolution of enzymes. *Gene* **236**, 159.

- Kikuchi, M., Ohnishi, K., Harayama, S.**, 2000. An effective family shuffling method using single-stranded DNA. *Gene* **243**, 133.
- Kim, D., Robyt, J.F., Lee, S.Y., Lee, J.H., Kim, Y.M.**, 2003. Dextran molecular size and degree of branching as a function of sucrose concentration, pH, and temperature of reaction of *Leuconostoc mesenteroides* B-512FMCM dextransucrase. *Carbohydr. Res.* **338**, 1183.
- Kim, Y.W., Lee, S.S., Warren, R.A., Withers, S.G.**, 2004. Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire. *J Biol Chem* **279**, 42787-42793.
- Koga, T., Oho, T., Shimazaki, Y., Nakano, Y.**, 2002. Immunization against dental caries. *Vaccine* **20**, 2027.
- Kohmoto, T., Fukui, F., Takaku, H., Machida, Y., Auai, M., Mitsuoka, T.**, 1984. Effect of isomaltooligosaccharides on human fecal flora. *Bifidobacteria Microflora* **7**, 61-69.
- Kohmoto, T., Fukui, F., Takaku, H., Mitsuoka, T.**, 1991. Dose-response test of isomaltooligosaccharides for increasing fecal *Bifidobacteria*. *Agric. Biol. Chem.* **55**, 2157-2159.
- Kohmoto, T., Tsuji, K., Kaneko, T., Shiota, M., Fukui, F., Takaku, H., Nakagawa, Y., Ichikawa, T., Kobayashi, S.**, 1992. Metabolism of ¹³C-isomaltooligosaccharides in healthy men. *Biosci. Biotechnol. Biochem.* **56**, 937-940.
- Koné, F.M.T., Le Béhec, M., Sine, J.-P., Dion, M., Tellier, C.**, 2009. Digital screening methodology for the directed evolution of transglycosidases. *Protein Eng. Des. Sel.* **22**, 37-44.
- Konishi, N., Torii, Y., Yamamoto, T., Miyagi, A., Ohta, H., Fukui, K., Hanamoto, S., Matsuno, H., Komatsu, H., Kodama, T., Katayama, E.**, 1999. Structure and enzymatic properties of genetically truncated forms of the water-insoluble glucan-synthesizing glucosyltransferase from *Streptococcus sobrinus*. *J. Biochem.* **126**, 287-295.
- Kralj, S.**, 2004. Glucansucrases of *Lactobacilli*: characterization of genes, enzymes, and products synthesized. **PhD Thesis, Rijksuniversiteit Groningen, Groningen.**
- Kralj, S., Eeuwema, W., Eckhardt, T.H., Dijkhuizen, L.**, 2006. Role of asparagine 1134 in glucosidic bond and transglycosylation specificity of reuteransucrase from *Lactobacillus reuteri* 121. *Febs J.* **273**, 3735-3742.
- Kralj, S., Stripling, E., Sanders, P., van Geel-Schutten, G.H., Dijkhuizen, L.**, 2005a. Highly hydrolytic reuteransucrase from probiotic *Lactobacillus reuteri* strain ATCC 55730. *Appl. Environ. Microbiol.* **71**, 3942-3950.
- Kralj, S., van Geel-Schutten, G.H., Dondorff, M.M., Kirsanovs, S., van der Maarel, M.J., Dijkhuizen, L.**, 2004a. Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. *Microbiology* **150**, 3681-3690.
- Kralj, S., van Geel-Schutten, G.H., Rahaoui, H., Leer, R.J., Faber, E.J., van der Maarel, M.J., Dijkhuizen, L.**, 2002. Molecular characterization of a novel glucosyltransferase from *Lactobacillus*

reuteri strain 121 synthesizing a unique, highly branched glucan with α -(1→4) and α -(1→6) glucosidic bonds. *Appl. Environ. Microbiol.* **68**, 4283-4291.

Kralj, S., van Geel-Schutten, G.H., van der Maarel, M.J., Dijkhuizen, L., 2004b. Biochemical and molecular characterization of *Lactobacillus reuteri* 121 reuteransucrase. *Microbiology* **150**, 2099-2112.

Kralj, S., van Geel-Schutten, I.G., Faber, E.J., van der Maarel, M.J., Dijkhuizen, L., 2005b. Rational transformation of *Lactobacillus reuteri* 121 reuteransucrase into a dextransucrase. *Biochemistry* **44**, 9206-9216.

Kralj, S., Van Leeuwen, S.S., Valk, V., Eeuwema, W., Kamerling, J.P., Dijkhuizen, L., 2008. Hybrid reuteransucrase enzymes reveal regions important for glucosidic linkage specificity and the transglucosylation/hydrolysis ratio. *FEBS J.* **275**, 6002.

L

Larm, O., Lindberg, B., Svensson, S., 1971. Studies on the length of the side chains of the dextran elaborated by *Leuconostoc mesenteroides* NRRL B-512. *Carbohydr. Res.* **20**, 39.

Leathers, T.D., 2005. Dextran. *Polysaccharides and polyamides in the food industry, vol. 1.* Wiley-VCH, 233-255.

Leathers, T.D., Ahlgren, J.A., Cote, G.L., 1997. Alternansucrase mutants of *Leuconostoc mesenteroides* strain NRRL B-21138. *J. Ind. Microbiol. Biotechnol.* **18**, 278.

Leathers, T.D., Hayman, G.T., Cote, G.L., 1995. Rapid screening of *Leuconostoc mesenteroides* mutants for elevated proportions of alternan to dextran. *Curr. Microbiol.* **31**, 19.

Leathers, T.D., Nunnally, M.S., Ahlgren, J.A., Côté, G.L., 2003. Characterization of a novel modified alternan. *Carbohydr. Pol.* **54**, 107.

Leathers, T.D., Nunnally, M.S., Cote, G.L., 2002a. *Penicillium* isolates for modifying alternan, US patent 6479275.

Leathers, T.D., Nunnally, M.S., Côté, G.L., 2002b. Modification of alternan by novel *Penicillium* spp. *J. Ind. Microbiol. Biotechnol.* **29**, 177.

Leathers, T.D., Nunnally, M.S., Côté, G.L., 2010. Optimization of process conditions for enzymatic modification of alternan using dextransucrase from *Chaetomium erraticum*. *Carbohydr. Pol.* **81**, 732.

Leemhuis, H., Nightingale, K.P., Hollfelder, F., 2008. Directed evolution of a histone acetyltransferase - enhancing thermostability, whilst maintaining catalytic activity and substrate specificity. *FEBS Journal* **275**, 5635.

Li, Y.-J., Zhao, G.-Y., Du, W., Zhang, T.-J., 2009. Effect of dietary isomaltooligosaccharides on nutrient digestibility and concentration of glucose, insulin, cholesterol and triglycerides in serum of growing pigs. *Anim. Feed Sci. Technol.* **151**, 312.

Lin-Goerke, J.L., Robbins, D.J., Burczak, J.D., 1997. PCR-based random mutagenesis using manganese and reduced dNTP concentration. *Biotechniques* **23**, 409-412.

Lin, H., Tao, H., Cornish, V.W., 2004. Directed evolution of a glycosynthase via chemical complementation. *Journal of the American Chemical Society* **126**, 15051.

Lipovsek, D., Antipov, E., Armstrong, K.A., Olsen, M.J., Klibanov, A.M., Tidor, B., Wittrup, K.D., 2007. Selection of horseradish peroxidase variants with enhanced enantioselectivity by yeast surface display. *Chemistry & Biology* **14**, 1176.

López-Munguía, A., Pelenc, V., Remaud, M., Biton, J., Michel, J.M., Lang, C., Paul, F., Monsan, P., 1993. Production and purification of alternansucrase, a glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355, for the synthesis of oligoalternans. *Enzyme and Microbial Technology* **15**, 77.

Lutz, S., Ostermeier, M., Moore, G.L., Maranas, C.D., Benkovic, S.J., 2001. Creating multiple-crossover DNA libraries independent of sequence identity. *PNAS* **98**, 11248-11253.

M

MacGregor, E.A., Jespersen, H.M., Svensson, B., 1996. A circularly permuted alpha-amylase-type alpha/beta-barrel structure in glucan-synthesizing glucosyltransferases. *FEBS Lett.* **378**, 263-266.

Maina, N.H., Virkki, L., Pynnonen, H., Maaheimo, H., Tenkanen, M., 2011. Structural analysis of enzyme-resistant isomaltooligosaccharides reveals the elongation of α -(1 \rightarrow 3)-linked branches in *Weissella confusa* dextran. *Biomacromolecules* **12**, 409-418.

Mastrobattista, E., Taly, V., Chanudet, E., Treacy, P., Kelly, B.T., Griffiths, A.D., 2005. High-throughput screening of enzyme libraries: *in vitro* evolution of a β -galactosidase by fluorescence-activated sorting of double emulsions. *Chemistry & Biology* **12**, 1291.

Matsuura, Y., Kusunoki, M., Harada, W., Kakudo, M., 1984. Structure and possible catalytic residues of Taka-amylase A. *J. Biochem.* **95**, 697-702.

Mayer, C., Jakeman, D.L., Mah, M., Karjala, G., Gal, L., Warren, R.A.J., Withers, S.G., 2001. Directed evolution of new glycosynthases from *Agrobacterium* β -glucosidase: a general screen to detect enzymes for oligosaccharide synthesis. *Chemistry & Biology* **8**, 437.

McCurdy, R.D., Goff, H.D., Stanley, D.W., Stone, A.P., 1994. Rheological properties of dextran related to food applications. *Food Hydrocolloids* **8**, 609.

Messmer, K., Ljungström, K.G., Gruber, U.F., Richter, W., Hedin, H., 1980. Prevention of dextran-induced anaphylactoid reactions by hapten inhibition. *The Lancet* **315**, 975.

Meyer, M.M., Silberg, J.J., Voigt, C.A., Endelman, J.B., Mayo, S.L., Wang, Z.G., Arnold, F.H., 2003. Library analysis of SCHEMA-guided protein recombination. *Protein Sci.* **12**, 1686.

Mirza, O., Skov, L.K., Remaud-Simeon, M., Potocki de Montalk, G., Albenne, C., Monsan, P., Gajhede, M., 2001. Crystal structures of amylosucrase from *Neisseria polysaccharea* in complex with D-glucose and the active site mutant Glu328Gln in complex with the natural substrate sucrose. *Biochemistry* **40**, 9032-9039.

- Monchois, V., Remaud-Simeon, M., Russell, R.R., Monsan, P., Willemot, R.M., 1997.** Characterization of *Leuconostoc mesenteroides* NRRL B-512F dextransucrase (DSRS) and identification of amino-acid residues playing a key role in enzyme activity. *Appl. Microbiol. Biotechnol.* **48**, 465-472.
- Monchois, V., Reverte, A., Remaud-Simeon, M., Monsan, P., Willemot, R.-M., 1998.** Effect of *Leuconostoc mesenteroides* NRRL B-512F dextransucrase carboxy-terminal deletions on dextran and oligosaccharide synthesis. *Appl. Environ. Microbiol.* **64**, 1644-1649.
- Monchois, V., Vignon, M., Russell, R.R., 2000.** Mutagenesis of asp-569 of glucosyltransferase I glucansucrase modulates glucan and oligosaccharide synthesis. *Appl. Environ. Microbiol.* **66**, 1923-1927.
- Monchois, V., Willemot, R.M., Monsan, P., 1999.** Glucansucrases: mechanism of action and structure–function relationships. *Fems Microbiol. Rev.* **23**, 131.
- Monchois, V., Willemot, R.M., Remaud-Simeon, M., Croux, C., Monsan, P., 1996a.** Cloning and sequencing of a gene coding for a novel dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only $\alpha(1-6)$ and $\alpha(1-3)$ linkages. *Gene* **182**, 23-32.
- Monchois, V., Willemot, R.M., Remaud-Simeon, M., Croux, C., Soucaille, P., Monsan, P., 1996b.** Primary structure of dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299. Comparison with dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *Ann. N Y Acad. Sci.* **799**, 151-156.
- Monsan, P., Remaud-Siméon, M., André, I., 2010.** Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr. Opin. Microbiol.* **13**, 293.
- Montès, H., Cavallé, J.Y., 1999.** Secondary dielectric relaxations in dried amorphous cellulose and dextran. *Polymer* **40**, 2649-2657.
- Montesanti, N., Véronèse, G., Buléon, A., Escalier, P.-C., Kitamura, S., Putaux, J.-L., 2010.** A-type crystals from dilute solutions of short amylose chains. *Biomacromolecules* **11**, 3049-3058.
- Moon, Y., Nam, S., Kang, J., Kim, Y.-M., Lee, J.-H., Kang, H.-K., Breton, V., Jun, W.-J., Park, K.-D., Kimura, A., Kim, D., 2007.** Enzymatic synthesis and characterization of arbutin glucosides using glucansucrase from *Leuconostoc mesenteroides* B-1299CB. *Appl. Microbiol. Biotechnol.* **77**, 559.
- Mooser, G., Iwaoka, K.R., 1989.** Sucrose 6- α -D-glucosyltransferase from *Streptococcus sobrinus*: characterization of a glucosyl-enzyme complex. *Biochemistry* **28**, 443-449.
- Mosi, R., He, S., Uitdehaag, J., Dijkstra, B.W., Withers, S.G., 1997.** Trapping and characterization of the reaction intermediate in cyclodextrin glycosyltransferase by use of activated substrates and a mutant Enzyme. *Biochemistry* **36**, 9927-9934.
- Moulis, C., Arcache, A., Escalier, P.C., Rinaudo, M., Monsan, P., Remaud-Simeon, M., Potocki-Veronese, G., 2006a.** High-level production and purification of a fully active recombinant dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *FEMS Microbiol. Lett.* **261**, 203-210.

Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M., 2006b. Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J. Biol. Chem.* **281**, 31254-31267.

Muthana, S., Cao, H., Chen, X., 2009. Recent progress in chemical and chemoenzymatic synthesis of carbohydrates. *Curr. Opin. Chem. Biol.* **13**, 573.

N

Naessens, M., Cerdobbel, A., Soetaert, W., Vandamme, E.J., 2005. *Leuconostoc* dextranucrase and dextran: production, properties and applications. *J. Chem. Technol. Biotechnol.* **80**, 845.

Ness, J.E., Kim, S., Gottman, A., Pak, R., Krebber, A., Borchert, T.V., Govindarajan, S., Mundorff, E.C., Minshull, J., 2002. Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat. Biotech.* **20**, 1251.

Neubauer, H., Bauche, A., Mollet, B., 2003. Molecular characterization and expression analysis of the dextranucrase DsrD of *Leuconostoc mesenteroides* Lcc4 in homologous and heterologous *Lactococcus lactis* cultures. *Microbiology* **149**, 973-982.

O

Oliphant, A.R., Nussbaum, A.L., Struhl, K., 1986. Cloning of random-sequence oligodeoxynucleotides. *Gene* **44**, 177.

Olsen, M.J., Stephens, D., Griffiths, D., Daugherty, P., Georgiou, G., Iverson, B.L., 2000. Function-based isolation of novel enzymes from a large library. *Nat Biotech* **18**, 1071.

Olvera, C., Fernandez-Vazquez, J.L., Ledezma-Candanoza, L., Lopez-Munguia, A., 2007. Role of the C-terminal region of dextranucrase from *Leuconostoc mesenteroides* IBT-PQ in cell anchoring. *Microbiology* **153**, 3994-4002.

Ostermeier, M., Shim, J.H., Benkovic, S.J., 1999. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat. Biotech.* **17**, 1205.

P

Parikh, M.R., Greene, D.N., Woods, K.K., Matsumura, I., 2006. Directed evolution of RuBisCO hypermorphs through genetic selection in engineered *E. coli*. *Protein Eng. Des. Sel.* **19**, 113-119.

Parikh, M.R., Matsumura, I., 2005. Site-saturation mutagenesis is more efficient than DNA shuffling for the directed evolution of β -fucosidase from β -galactosidase. *Journal of Molecular Biology* **352**, 621.

Peisajovich, S.G., Rockah, L., Tawfik, D.S., 2006. Evolution of new protein topologies through multistep gene rearrangements. *Nat. Genet.* **38**, 168.

Pijning, T., Vujičić-Žagar, A., Kralj, S., Eeuwema, W., Dijkhuizen, L., Dijkstra, B.W., 2008. Biochemical and crystallographic characterization of a glucansucrase from *Lactobacillus reuteri* 180. *Biocatal. Biotransform.* **26**, 12-17.

Pinder, D.N., Swanson, A.J., Hebraud, P., Hemar, Y., 2006. Micro-rheological investigation of dextran solutions using diffusing wave spectroscopy. *Food Hydrocolloids* **20**, 240-244.

Piret, J., Lamontagne, J., Bestman-Smith, J., Roy, S., Gourde, P., Desormeaux, A., Omar, R.F., Juhasz, J., Bergeron, M.G., 2000. *In vitro* and *in vivo* evaluations of sodium lauryl sulfate and dextran sulfate as microbicides against herpes simplex and human immunodeficiency viruses. *J. Clin. Microbiol.* **38**, 110-119.

Pizzut-Serin, S., Potocki-Véronèse, G., van der Veen, B.A., Albenne, C., Monsan, P., Remaud-Simeon, M., 2005. Characterisation of a novel amylosucrase from *Deinococcus radiodurans*. *FEBS Lett.* **579**, 1405.

Ponsard, I., Galleni, M., Soumillion, P., Fastrez, J., 2001. Selection of metalloenzymes by catalytic activity using phage display and catalytic elution. *ChemBioChem* **2**, 253.

Popov, D., Buléon, A., Burghammer, M., Chanzy, H., Montesanti, N., Putaux, J.L., Potocki-Véronèse, G., Riekkel, C., 2009. Crystal structure of A-amylose: a revisit from synchrotron microdiffraction analysis of single crystals. *Macromolecules* **42**, 1167-1174.

Popov, D., Burghammer, M., Buléon, A., Montesanti, N., Putaux, J.L., Riekkel, C., 2006. A-amylose single crystals: unit cell refinement from synchrotron radiation microdiffraction data. *Macromolecules* **39**, 3704-3706.

Potocki-Veronese, G., Putaux, J.-L., Dupeyre, D., Albenne, C., Remaud-Siméon, M., Monsan, P., Buleon, A., 2005. Amylose synthesized *in vitro* by amylosucrase: morphology, structure, and properties. *Biomacromolecules* **6**, 1000-1011.

Potocki de Montalk, G., Remaud-Simeon, M., Willemot, R.-M., Sarçabal, P., Planchot, V., Monsan, P., 2000. Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties. *FEBS Letters* **471**, 219-223.

Potocki De Montalk, G.P., Remaud-Simeon, M., Willemot, R.M., Planchot, V., Monsan, P., 1999. Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J. Bacteriol.* **181**, 375-381.

R

Reetz, M.T., Becker, M.H., Klein, H.W., Stöckigt, D., 1999. A method for high-throughput screening of enantioselective catalysts. *Angewandte Chemie International Edition* **38**, 1758.

Reetz, M.T., Eipper, A., Tielmann, P., Mynott, R., 2002. A practical NMR-based high-throughput assay for screening enantioselective catalysts and biocatalysts. *Adv. Synth. Catal.* **344**, 1008-1016.

Reetz, M.T., Hermes, M., Becker, M.H., 2001a. Infrared-thermographic screening of the activity and enantioselectivity of enzymes. *Applied Microbiology and Biotechnology* **55**, 531.

- Reetz, M.T., Hobenreich, H., Soni, P., Fernandez, L.,** 2008a. A genetic selection system for evolving enantioselectivity of enzymes. *Chem. Commun.*, 5502.
- Reetz, M.T., Kahakeaw, D., Lohmer, R.,** 2008b. Addressing the numbers problem in directed evolution. *Chembiochem* **9**, 1797-1804.
- Reetz, M.T., Kühling, K., Deege, A., Hinrichs, H., Belder, D.,** 2000a. Super-high-throughput screening of enantioselective catalysts by using capillary array electrophoresis. *Angewandte Chemie International Edition* **39**, 3891.
- Reetz, M.T., Kühling, K., Hinrichs, H., Deege, A.,** 2000b. Circular dichroism as a detection method in the screening of enantioselective catalysts. *Chirality* **12**, 479-482.
- Reetz, M.T., Kühling, K.M., Wilensek, S., Husmann, H., Häusig, U.W., Hermes, M.,** 2001b. A GC-based method for high-throughput screening of enantioselective catalysts. *Catalysis Today* **67**, 389.
- Reetz, M.T., Zonta, A., Schimossek, K., Jaeger, K.E., Liebeton, K.,** 1997. Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angewandte Chemie International Edition in English* **36**, 2830.
- Remaud-Simeon, M., Lopez-Munguia, A., Pelenc, V., Paul, F., Monsan, P.,** 1994. Production and use of glucosyltransferases from *Leuconostoc mesenteroides* NRRL B-1299 for the synthesis of oligosaccharides containing α -(1 \rightarrow 2) linkages. *Appl. Biochem. Biotechnol.* **44**, 101-117.
- Remaud-Simeon, M., Willemot, R.M., Sarcabal, P., de Montalk, G.P., Monsan, P.,** 2000. Glucansucrases: molecular engineering and oligosaccharide synthesis. *J. Mol. Catal. B-Enzym.* **10**, 117-128.
- Reymond, J.L., Babiak, P.,** 2007. Screening systems. *Adv Biochem Eng Biotechnol* **105**, 31-58.
- Roberfroid, M.B.,** 2007. Prebiotics: the concept revisited. *The Journal of Nutrition* **137**, 830S-837S.
- Robyt, J.F., Eklund, S.H.,** 1983. Relative, quantitative effects of acceptors in the reaction of *Leuconostoc mesenteroides* B-512F dextranucrase. *Carbohydr. Res.* **121**, 279.
- Roessner, D., Kulicke, W.-M.,** 1994. On-line coupling of flow field-flow fractionation and multi-angle laser light scattering. *J. Chromatogr. A* **687**, 249-258.
- Rolland-Sabaté, A., Mendez-Montealvo, M.G., Colonna, P., Planchot, V.,** 2008. Online determination of structural properties and observation of deviations from power law behavior. *Biomacromolecules* **9**, 1719-1730.
- Russell, R.R., Gilpin, M., Mukasa, H., Dougan, G.,** 1987. Characterization of glucosyltransferase expressed from a *Streptococcus sobrinus* gene cloned in *Escherichia coli*. *Gen. Microbiol.* **133**, 935-944.
- Russell, R.R.B.,** 1990. Molecular genetics of glucan metabolism in oral *streptococci*. *Archives of Oral Biology* **35**, S53.

S

- Sabatié, J., Choplin, L., Doublier, J.L., Arul, J., Paul, F., Monsan, P.**, 1988. Rheology of native dextrans in relation to their primary structure. *Carbohydr. Pol.* **9**, 287.
- Sabatie, J., Choplin, L., Moan, M., Doublier, J.L., Paul, F., Monsan, P.**, 1988. The effect of synthesis temperature on the structure of dextran NRRL B 512F. *Carbohydr. Pol.* **9**, 87.
- Sanchez-Gonzalez, M., Alagon, A., Rodriguez-Sotres, R., Lopez-Munguia, A.**, 1999. Proteolytic processing of dextransucrase of *Leuconostoc mesenteroides*. *FEMS Microbiol. Lett.* **181**, 25-30.
- Santoro, S.W., Wang, L., Herberich, B., King, D.S., Schultz, P.G.**, 2002. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat Biotech* **20**, 1044.
- Seibel, J., Hellmuth, H., Hofer, B., Kicinska, A.M., Schmalbruch, B.**, 2006. Identification of new acceptor specificities of glycosyltransferase R with the aid of substrate microarrays. *ChemBioChem* **7**, 310.
- Seymour, F.R., Knapp, R.D., Bishop, S.H.**, 1979a. Correlation of the structure of dextrans to their ¹H-N.M.R. spectra. *Carbohydr. Res.* **74**, 77-92.
- Seymour, F.R., Knapp, R.D., Bishop, S.H., Jeanes, A.**, 1979b. High-temperature enhancement of ¹³C-N.M.R. chemical-shifts of unusual dextrans, and correlation with methylation structural analysis. *Carbohydr. Res.* **68**, 123-140.
- Shaikh, F.A., Withers, S.G.**, 2008. Teaching old enzymes new tricks: engineering and evolution of glycosidases and glycosyl transferases for improved glycoside synthesis. *Biochem Cell Biol* **86**, 169-177.
- Shao, Z., Zhao, H., Giver, L., Arnold, F.H.**, 1998. Random-priming *in vitro* recombination: an effective tool for directed evolution. *Nucleic Acids Res.* **26**, 681-683.
- Shimamura, A.**, 1989. Use of ¹³C-n.m.r. spectroscopy for the quantitative estimation of 3-O- and 3,6-di-O-substituted -glucopyranosyl residues in α -D-glucans formed by the D-glycosyltransferases of *Streptococcus sobrinus*. *Carbohydr. Res.* **185**, 239-248.
- Shimamura, A., Nakano, Y.J., Mukasa, H., Kuramitsu, H.K.**, 1994. Identification of amino acids residues in *Streptococcus mutans* glycosyltransferases influencing the structure of the glucan product. *J. Bacteriol.* **176**, 4845-4850.
- Shimba, N., Yokoyama, K., Suzuki, E.**, 2002. NMR-based screening method for transglutaminases: rapid analysis of their substrate specificities and reaction rates. *J. Agric. Food Chem.* **50**, 1330-1334.
- Shiroza, T., Ueda, S., Kuramitsu, H.K.**, 1987. Sequence analysis of the gtfB gene from *Streptococcus mutans*. *J. Bacteriol.* **169**, 4263-4270.
- Sieber, V., Martinez, C.A., Arnold, F.H.**, 2001. Libraries of hybrid proteins from distantly related sequences. *Nat. Biotech.* **19**, 456.

Simpson, C.L., Cheetham, N.W., Giffard, P.M., Jacques, N.A., 1995. Four glucosyltransferases, GtfJ, GtfK, GtfL and GtfM, from *Streptococcus salivarius* ATCC 25975. *Microbiology* **141 (Pt 6)**, 1451-1460.

Sinnott, M.L., 1990. Catalytic mechanism of enzymic glycosyl transfer. *Chem. Rev.* **90**, 1171.

Skov, L.K., Mirza, O., Henriksen, A., De Montalk, G.P., Remaud-Simeon, M., Sarcabal, P., Willemot, R.-M., Monsan, P., Gajhede, M., 2001. Amylosucrase, a glucan-synthesizing enzyme from the α -amylase family. *J. Biol. Chem.* **276**, 25273-25278.

Skov, L.K., Mirza, O., Sprogøe, D., Dar, I., Remaud-Simeon, M., Albenne, C., Monsan, P., Gajhede, M., 2002. Oligosaccharide and sucrose complexes of amylosucrase: structural implications for the polymerase activity. *J. Biol. Chem.* **277**, 47741-47747.

Slade, L., Levine, H., 1991. A food polymer science approach to structure-property relationships in aqueous food systems: non-equilibrium behavior of carbohydrate-water systems. *Adv. Exp. Med. Biol.* **302**, 29-101.

Soumillion, P., Jespers, L., Bouchet, M., Marchand-Brynaert, J., Winter, G., Fastrez, J., 1994. Selection of β -lactamase on filamentous bacteriophage by catalytic activity. *J. Mol. Biol.* **237**, 415.

Spee, J.H., de Vos, W.M., Kuipers, O.P., 1993. Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucleic Acids Res.* **21**, 777-778.

Squires, T.M., Quake, S.R., 2005. Microfluidics: fluid physics at the nanoliter scale. *Reviews of Modern Physics* **77**, 977.

Stemmer, W.P., 1994a. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *PNAS* **91**, 10747-10751.

Stemmer, W.P., 1994b. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370**, 389.

Striegel, A., Isenberg, S., Côté, G., 2009. An SEC/MALS study of alternan degradation during size-exclusion chromatographic analysis. *Anal. Bioanal. Chem.* **394**, 1887.

Sun, G.M., Shen, Y.I., Ho, C.C., Kusuma, S., Gerecht, S., 2009. Functional groups affect physical and biological properties of dextran-based hydrogels. *J. Biomed. Materials Res. Part A* **93A**, 1080-1090.

Suwannarangsee, S., Moulis, C., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M., Chulalaksananukul, W., 2007. Search for a dextranase minimal motif involved in dextran binding. *FEBS Lett.* **581**, 4675.

Svensson, B., 1994. Protein engineering in the α -amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Mol. Biol.* **25**, 141.

Sylvestre, J., 2010. Massive Mutagenesis®: high-throughput combinatorial site-directed mutagenesis, In vitro mutagenesis protocols: Third Edition. pp. 233-238.

T

Takahashi, F., Ebihara, T., Mie, M., Yanagida, Y., Endo, Y., Kobatake, E., Aizawa, M., 2002.

Ribosome display for selection of active dihydrofolate reductase mutants using immobilized methotrexate on agarose beads. *FEBS Letters* **514**, 106.

Tawfik, D.S., Griffiths, A.D., 1998. Man-made cell-like compartments for molecular evolution. *Nat Biotech* **16**, 652.

Teraoka, I., 2002. Polymer solutions: an introduction to physical properties. Chapter 1: Models of polymer chains. *New York: John Wiley & Sons*, 1-77.

Tindall, K.R., Kunkel, T.A., 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**, 6008.

U

Ueda, S., Shiroza, T., Kuramitsu, H.K., 1988. Sequence analysis of the gtfC gene from *Streptococcus mutans* GS-5. *Gene* **69**, 101-109.

Uitdehaag, J.C.M., Mosi, R., Kalk, K.H., van der Veen, B.A., Dijkhuizen, L., Withers, S.G., Dijkstra, B.W., 1999. X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α -amylase family. *Nat. Struct. Mol. Biol.* **6**, 432.

V

Valette, P., Pelenc, V., Djouzi, Z., Andrieux, C., Paul, F., Monsan, P., Szylit, O., 1993. Bioavailability of new synthesized glucooligosaccharides in the intestinal tract of gnotobiotic rats. *J. Sci. Food Agric.* **62**, 121-127.

Van Der Veen, B.A., Potocki-Veronese, G., Albenne, C., Joucla, G., Monsan, P., Remaud-Simeon, M., 2004. Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett.* **560**, 91-97.

Van Hijum, S.A.F.T., Kralj, S., Ozimek, L.K., Dijkhuizen, L., van Geel-Schutten, I.G.H., 2006. Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 157-176.

Van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2009. Structural characterization of bioengineered α -D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180. *Biomacromolecules* **10**, 580-588.

Van Leeuwen, S.S., Kralj, S., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008a. Structural analysis of bioengineered α -D-glucan produced by a triple mutant of the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180: generation of (α 1 \rightarrow 4) linkages in a native (1 \rightarrow 3)(1 \rightarrow 6)- α -D-glucan. *Biomacromolecules* **9**, 2251.

van Leeuwen, S.S., Kralj, S., van Geel-Schutten, I.H., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008b. Structural analysis of the α -D-glucan (EPS35-5) produced by the *Lactobacillus reuteri* strain 35-5 glucansucrase GTFA enzyme. *Carbohydr. Res.* **343**, 1251-1265.

Van Leeuwen, S.S., Kralj, S., Van Geel-Schutten, I.H., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008c. Structural analysis of the α -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme. *Carbohydr. Res.* **343**, 1237.

Van Leeuwen, S.S., Leeflang, B.R., J., G.G., Kamerling, J.P., 2008d. Development of a ^1H NMR structural-reporter-group concept for the primary structural characterisation of α -D-glucans. *Carbohydr. Res.* **343**, 1114-1111.

Van Sint Fiet, S., van Beilen, J.B., Witholt, B., 2006. Selection of biocatalysts for chemical synthesis. *PNAS* **103**, 1693-1698.

Van Tomme, S.R., Hennink, W.E., 2007. Biodegradable dextran hydrogels for protein delivery applications. *Expert Rev. Med. Devices* **4**, 147-164.

Vanwetswinkel, S., Avalle, B., Fastrez, J., 2000. Selection of β -lactamases and penicillin binding mutants from a library of phage displayed TEM-1 β -lactamase randomly mutated in the active site Ω -loop. *J. Mol. Biol.* **295**, 527.

Vickerman, M.M., Sulavik, M.C., Nowak, J.D., Gardner, N.M., Jones, G.W., Clewell, D.B., 1997. Nucleotide sequence analysis of the *Streptococcus gordonii* glucosyltransferase gene, *gtfG*. *DNA Seq.* **7**, 83-95.

Viebke, C., Williams, P.A., 2000. The influence of temperature on the characterization of water-soluble polymers using asymmetric flow field-flow-fractionation coupled to multiangle laser light scattering. *Anal. Chem.* **72**, 3896-3901.

Voigt, C.A., Martinez, C., Wang, Z.-G., Mayo, S.L., Arnold, F.H., 2002. Protein building blocks preserved by recombination. *Nat. Struct. Mol. Biol.* **9**, 553.

Vujicic-Zagar, A., Pijning, T., Kralj, S., Lopez, C.A., Eeuwema, W., Dijkhuizen, L., Dijkstra, B.W., 2010. Crystal structure of a 117 kDa glucansucrase fragment provides insight into evolution and product specificity of GH70 enzymes. *PNAS* **107**, 21406-21411.

W

Wang, H.-F., Lim, P.-S., Kao, M.-D., Chan, E.-C., Lin, L.-C., Wang, N.-P., 2001. Use of isomaltoligosaccharide in the treatment of lipid profiles and constipation in hemodialysis patients. *J. Renal Nutr.* **11**, 73.

Williams, G.J., Zhang, C., Thorson, J.S., 2007. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657.

Wittgren, B., Wahlund, K.-G., 1997. Fast molecular mass and size characterization of polysaccharides using asymmetrical flow field-flow fractionation-multiangle light scattering. *J. Chromatogr. A* **760**, 205-218.

Wong, T.S., Roccatano, D., Loakes, D., Tee, K.L., Schenk, A., Hauer, B., Schwaneberg, U., 2008. Transversion-enriched sequence saturation mutagenesis (SeSaM-Tv+): a random mutagenesis method

with consecutive nucleotide exchanges that complements the bias of error-prone PCR. *Biotechnol. J.* **3**, 74.

Wong, T.S., Zhurina, D., Schwaneberg, U., 2006. The diversity challenge in directed protein evolution. *Comb. Chem. High Throughput Screen* **9**, 271-288.

Y

Yalin, Y., Jin, L., Jianhua, W., Da, T., Zigang, T., 2008. Expression and characterization of dextransucrase gene *dsrX* from *Leuconostoc mesenteroides* in *Escherichia coli*. *J. Biotechnol.* **133**, 505-512.

Yi, A.R., Lee, S.R., Jang, M.U., Park, J.M., Eom, H.J., Han, N.S., Kim, T.J., 2009. Cloning of dextransucrase gene from *Leuconostoc citreum* HJ-P4 and its high-level expression in *E. coli* by low temperature induction. *J. Microbiol. Biotechnol.* **19**, 829-835.

Yohannes, G., Jussila, M., Hartonen, K., Riekkola, M.L., 2011. Asymmetrical flow field-flow fractionation technique for separation and characterization of biopolymers and bioparticles. *J. Chromatogr. A* in press.

Yoshioka, Y., Hasegawa, K., Matsuura, Y., Katsube, Y., Kubota, M., 1997. Crystal structures of a mutant maltotetraose-forming exo-amylase cocrystallized with maltopentaose. *J. Mol. Biol.* **271**, 619-628.

Z

Zaccolo, M., Williams, D.M., Brown, D.M., Gherardi, E., 1996. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**, 589.

Zha, D., Eipper, A., Reetz, M.T., 2003a. Assembly of designed oligonucleotides as an efficient method for gene recombination: a new tool in directed evolution. *ChemBioChem* **4**, 34.

Zha, W., Zhu, T., Zhao, H., 2003b. Family shuffling with single-stranded DNA. *Methods in Molecular Biology*, Volume 231, 1, 91-97, DOI: 10.1385/1-59259-395-X:91.

Zhang, H., Hu, Y., Zhu, C., Zhu, B., Wang, Y., 2008. Cloning, sequencing and expression of a dextransucrase gene (*dexYG*) from *Leuconostoc mesenteroides*. *Biotechnol. Lett.* **30**, 1441-1446.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A., Arnold, F.H., 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotech.* **16**, 258.

Zhao, H., Zha, W., 2006. *In vitro* 'sexual' evolution through the PCR-based staggered extension process (StEP). *Nat. Protocols* **1**, 1865.

Zimm, B.H., 1948. The scattering of light and the radial distribution function of high polymer solutions. *J. Chem. Phys.* **16**, 1093-1099.

CHAPITRE 2 : Développement d'une stratégie multi-étapes de sélection et criblage haut-débit de la spécificité des glucane-saccharases

Cette thèse a porté sur l'ingénierie des glucane-saccharases de la famille GH70, afin d'approfondir les connaissances de leurs relations structure-spécificité et de concevoir de nouveaux outils enzymatiques de synthèse d' α -glucanes de structures et propriétés innovantes. Devant le peu d'informations structurales disponibles jusqu'en 2010 sur cette famille d'enzymes, nous avons donc choisi d'appréhender ces questions par une approche d'ingénierie combinatoire, basée sur la création et le criblage des larges banques de variants de l'enzyme DSR-S vardel Δ 4N, la dextrane-saccharase la plus active connue à ce jour. Or, il n'existait aucune méthode de criblage à haut-débit de la spécificité de liaison des transglucosidases.

Le premier objectif de cette thèse a donc été de développer une stratégie innovante de criblage haut-débit, sensible, robuste et reproductible, permettant de discriminer une banque de variants de glucane-saccharases selon leurs spécificités de liaisons. Pour ce faire, nous nous sommes inspirés du concept de « rapporteurs structuraux » obtenus par RMN 1D du proton, correspondant à des déplacements chimiques spécifiques du type de liaisons glucosidiques synthétisées par les glucane-saccharases. En effet, la spectrométrie RMN est un outil puissant pour l'élucidation des structures moléculaires, notamment celle des sucres. Elle permet d'accéder facilement à des informations structurales essentielles et d'identifier rapidement et sans destruction des molécules la nature des liaisons osidiques impliquées dans les polysaccharides.

La méthode qui a été mise au point comprend :

- (i) L'expression des gènes variants et la production des protéines recombinantes par *E. coli*, au format microplaque dans des conditions optimisées.
- (ii) Pour chaque clone, la synthèse enzymatique d'oligosaccharides à partir de saccharose et d'isomalto-oligosaccharides jouant le rôle de molécules acceptrices.
- (iii) L'analyse directe du milieu réactionnel brut par RMN 1D du proton et l'identification des rapporteurs structuraux, via un système de RMN par injection en flux.

Le développement de ce crible a été réalisé en collaboration avec Stéphane Massou (plateforme MetaSyS, LISBP) et la plateforme RMN biologique de l'IPBS (animée par Olivier Saurel et Alain Milon), qui est dotée du système Best-NMR permettant l'analyse en flux des échantillons conditionnés au format micro-plaque.

La méthode a, dans un premier temps, été validée sur trois glucane-saccharases modèles de spécificités connues. Il s'agit des glucane-saccharases DSR-S vardel $\Delta 4N$ [spécifique des liaisons de type $\alpha(1\rightarrow 6)$], ASR C-del [spécifique des liaisons de type $\alpha(1\rightarrow 6)/\alpha(1\rightarrow 3)$ alternées] et DSR-E [spécifique des liaisons de type $\alpha(1\rightarrow 6)$, $\alpha(1\rightarrow 3)$ et $\alpha(1\rightarrow 2)$]. Les résultats ont montré d'une part que la résonance des protons anomériques, caractéristiques des différentes liaisons des produits synthétisés, constitue un marqueur de choix pour l'identification de la spécificité des glucane-saccharases. Cette méthode est de plus suffisamment sensible, reproductible et robuste pour cribler 480 variants/jour, ce qui permet d'explorer de larges banques de variants. Ce concept a été validé en appliquant cette méthode au criblage d'une banque de plus de 4 000 variants de la DSR-S vardel $\Delta 4N$, ce qui a permis d'isoler plusieurs centaines de clones de spécificité altérée par rapport à l'enzyme sauvage.

En amont de ce travail, nous avons également travaillé sur le développement d'une méthode de sélection génétique *in vivo* afin de réaliser un premier tri dans la diversité générée lors de la création de banques de variants, afin d'en limiter la taille pour le criblage de la spécificité. Cette méthode à très haut débit (200 000 clones/semaine) permet d'isoler les variants actifs et discriminer les glucane-saccharases selon leurs activités d'hydrolyse ou de polymérisation à partir du saccharose. Cette méthode est basée sur le fait que des variants capables de cliver le saccharose libèrent du fructose qui induit la croissance de la cellule hôte, incapable de métaboliser le saccharose.

NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity

Romain Irague, Stéphane Massou, Claire Moulis, Olivier Saurel, Alain Milon, Pierre Monsan, Magali Remaud-Siméon, Gabrielle Potocki-Véronèse

Abstract

We report here the development of a straightforward, sensitive, and quantitative NMR-based method for high-throughput characterization of carbohydrate structure and screening of carbohydrate active enzyme (CAZyme) specificity. Automated assays starting from gene library expression to carbohydrate structure determination directly from crude reaction media have been established and successfully used to screen a library of 4032 CAZymes obtained by combinatorial engineering, at a rate of 480 enzyme variants per day. This allowed one to accurately discriminate 303 enzyme variants with altered specificity. The results demonstrate the potential of high-throughput NMR technology in glycomics, to mine artificial and natural enzyme diversity for novel biocatalysts.

Results and discussion

The development of glycoenzyme-based processes for the production of novel glycans, oligosaccharides or glycoconjugates has gained interest for applications in human health, food and feed industries, fine chemistry and white biotechnologies. Transglycosidases and glycosyltransferases, in particular, constitute very efficient tools for glycodiversification (Monsan *et al.*, 2010; Weijers *et al.*, 2008). In addition, protein engineering has shown to be extremely powerful to expand their substrate or product specificities and adapt them to synthetic substrates (Homann and Seibel, 2009; Kittl and Withers, 2010). However, searching for a carbohydrate-active enzymes (CAZymes) showing required specificities in libraries of thousand mutants is still a challenge. This is principally due to (i) the high structural complexity and diversity of carbohydrates (three hexoses can yield more than 30,000 different oligosaccharidic structures in theory) (ii) the difficult acquisition of carbohydrate structural information via rapid and automated method. These limitations often force the biologists to drastically reduce the size of library to be screened, especially when searching for glycoenzymes that catalyze novel carbohydrate substitution patterns. The detailed characterization of carbohydrate structures usually requires the combination of several analytical methods including NMR spectroscopy, mass spectrometry, enzymatic and chemical degradation, methylation and acetylation analysis. Among this panel of techniques, ¹H NMR spectroscopy, which is both non destructive and rapid, can provide key data for the structural characterization of carbohydrates (Duus *et al.*, 2000). Indeed, the concept of the "structural reporter group", which encompasses the assignment of well-resolved ¹H NMR chemical shifts to specific carbohydrate structural features including the type of sugar units, the linkage composition, or the substitution pattern, was introduced thirty years ago (Vliegthart *et al.*, 1983). Recently, this concept was revived and applied to the identification and quantification of the various types of glucosidic linkages found in α -glucans and glucooligosaccharides produced by glucansucrases (Van Leeuwen *et al.*, 2009; Van Leeuwen *et al.*, 2008b; Van Leeuwen *et al.*, 2008c). NMR spectroscopy thus appears very attractive to screen and profile CAZyme specificities. However, to be applicable to large library screening, the technology must allow automated, rapid, sensitive, reproducible and quantitative monitoring of enzymatic reactions from crude and miniaturized reaction media. In that regard, flow injection technology may answer to most of those pre-requisites. It was successfully used for combinatorial chemistry, metabolic profiling, racemic mixture analysis and enzyme inhibitor identification (Dalvit *et al.*, 2003; Dumas *et al.*, 2006; Keifer, 2003; Reetz *et al.*, 2002). However, it was never tested to screen large libraries of enzyme variants directly from crude medium.

Here, we describe a flow-injection ¹H-NMR screening method used to screen libraries of glucansucrases for regiospecificity of glycidic bond formation. The automated and miniaturized method

encompasses (i) the expression of the gene library and production of the protein variants in 96-deepwell microplates, (ii) the enzymatic synthesis of glucooligosaccharides from a non-labelled substrate and (iii) the direct 1D ^1H NMR analysis of the crude reaction medium using automated flow-injection to reveal the key structural features of the produced carbohydrate(s) (Figure 30). The non-destructive, quantitative and straightforward method we developed permits one to sort out enzyme variants with modified specificity (in terms of type and/or amount of osidic linkage content) directly from crude reaction medium at a throughput exceeding 400 samples per day. This approach opens new perspectives for screening complex carbohydrate structures and for sorting highly specific CAzyme biocatalysts.

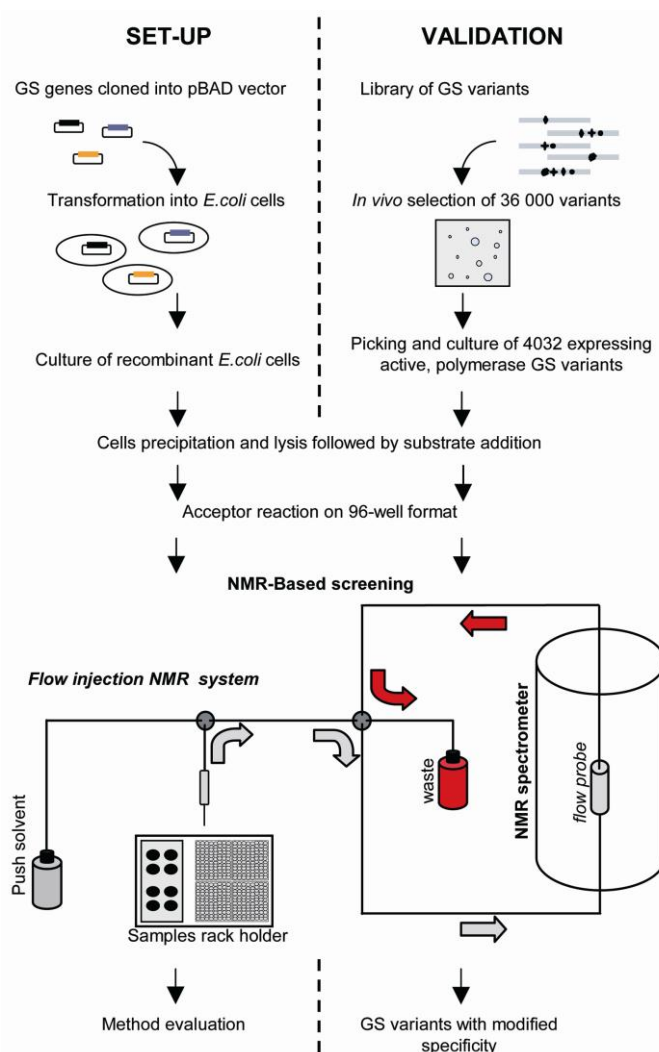


Figure 30: Outline of NMR-based screening for carbohydrate active enzyme specificity. The method was set-up with *E. coli* cells encoding empty plasmid (negative control) and parental glucansucrases with known specificities. Cells were cultured in 96-deepwells format for growth and gene expression. After lysis, recombinant enzymes were incubated with 292 mM sucrose and 50 mM iso-maltooligosaccharides to produce oligosaccharides. (trimethylsilyl)propionic acid-2,2,3,3-d₄ acid (37 mM), used as internal standard, was added to each samples before 1D ^1H NMR analysis carried out on the crude reaction media. The method was validated by screening a library of glucansucrase variants. Mutants were first selected on their abilities to cleave sucrose. Then, for each of the 4,032 selected active glucansucrase variants, the synthesized oligosaccharide structure was screened at a flow rate of one sample per 3 minutes.

Our model enzymes are glucansucrases. These transglucosidases belong to the Glycoside-Hydrolase family 70 (see Carbohydrate Active Enzymes database at <http://www.cazy.org/>) and catalyse the synthesis of α -glucans from sucrose (Cantarel *et al.*, 2009). When glycosides or non-carbohydrate molecules are used as exogenous acceptors, these enzymes catalyze the transfer of the glucosyl moiety from sucrose onto the acceptor molecules leading to oligosaccharide or glucoconjugate production. Depending on the enzyme regioselectivity, various types of glucosidic linkages (*ie* $\alpha(1\rightarrow2)$, $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$, $\alpha(1\rightarrow6)$) occur in the resulting sugar structures.

To set up the method (Figure 30), we first evaluated the ability of NMR to provide relevant structural information within short times – i.e. minutes - and for small scale assays. Three glucansucrases with known and different specificities were selected: DSR-S vardel $\Delta 4N$, which synthesizes $\alpha(1\rightarrow6)$ glucosidic linkages, ii) ASR C-del, which synthesizes both $\alpha(1\rightarrow6)$ and $\alpha(1\rightarrow3)$ glucosidic linkages and iii) DSR-E, which synthesizes $\alpha(1\rightarrow6)$, $\alpha(1\rightarrow3)$ and $\alpha(1\rightarrow2)$ glucosidic linkages (Bozonnet *et al.*, 2002; Joucla *et al.*, 2006; Moulis *et al.*, 2006b). All of the three glucansucrase encoding genes had previously been cloned into pBAD vector. The enzymes were produced using small-scale cultures of recombinant *E. coli* cells in 96-well microtitre plates. Following gene expression and cell disruption, the lysates were directly used as a source of enzyme to catalyze transglucosylation reaction from sucrose donor to isomaltooligosaccharide acceptors (IMOS). The crude reaction mixtures (500 μ L), containing the reaction products, were directly injected into an NMR flow probe. Process was automated using a flow injection NMR system and monitored by IconNMR software. All the experiments, starting from small-scale cultures to NMR analysis, were carried out in triplicate.

The anomeric protons of glucosyl moieties gave clear NMR signals that were assigned as described in previous studies (Figure 31 and Table 7) (Maina *et al.*, 2008; Van Leeuwen *et al.*, 2009; Van Leeuwen *et al.*, 2008b; Van Leeuwen *et al.*, 2008c). The chemical shifts at 4.98, 5.09, 5.19 and 5.34 ppm correspond to the anomeric proton of the $\alpha(1\rightarrow6)$ linked D-Glcp, $\alpha(1\rightarrow2)$ linked D-Glcp, $\alpha(1\rightarrow6)$ linked 2,6-di-O-substituted α -D-Glcp and $\alpha(1\rightarrow3)$ linked D-Glcp units, respectively, and were used as structural-reporter groups. NMR spectra revealed that the glucooligosaccharides produced in the different reaction mixtures contained new types of osidic linkages in addition to the $\alpha(1\rightarrow6)$ linkages of the IMOS acceptor (Figure 31). Indeed, signals corresponding to anomeric protons of units linked by

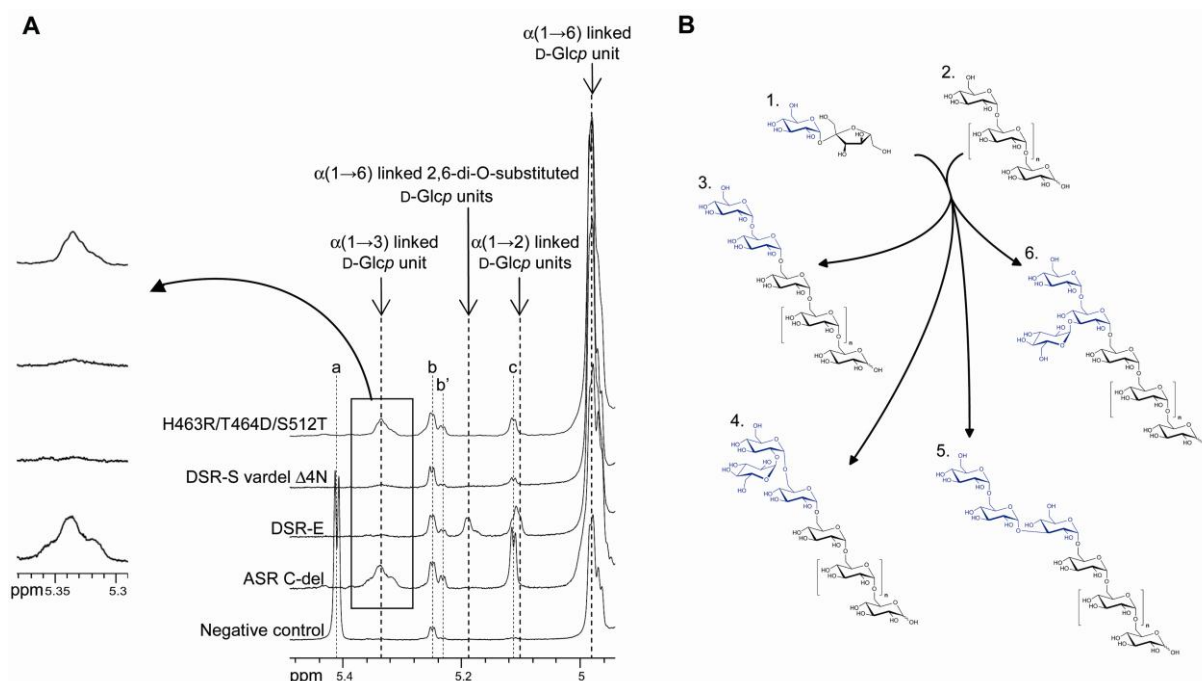


Figure 31: A. “Structural reporter group” used to screen the carbohydrate structures synthesized by the glucansucrases DSR-S vardel Δ 4N, DSR-E, ASR C-del, and by the DSR-S vardel Δ 4N H463R/T464D/S512T variant by 1D ^1H NMR. Spectra are centered at the anomeric region (4.94 to 5.48 ppm). Peaks were referenced to internal TSP-d4 (1H = 0 ppm). (a) α -D-Glcp unit of the sucrose, (b) α -D-Glcp reducing unit, (b') free D-Glcp unit resulting from sucrose hydrolysis reaction and (c) $\alpha(1\rightarrow5)$ -D-Glcp unit of the leucrose resulting from sucrose isomer synthesis. B. Examples of glucooligosaccharides synthesized from (1) sucrose and (2) isomaltooligosaccharides by (3) DSR-S vardel Δ 4N, (4) DSR-E, (5) ASR C-del and (6) DSR-S vardel Δ 4N H463R/T464D/S512T variant (structural representations show the osidic linkage composition of the products), n corresponds to the number of α -D-Glcp unit ranging from 3 to 5.

Table 7: ^1H NMR chemical shifts (in ppm) of the synthesized products.

	Negative control	DSR-S vardel Δ 4N	DSR-E	ASR C-del
β -D-Fruf-(2 \leftrightarrow 1)- α -D-Glcp H1 (sucrose)	5,413			
$\alpha(1\rightarrow3)$ linked D-Glcp unit		5,339 (1)	5,34 (1)	5,336 (10)
α -D-Glcp reducing unit		5,249	5,249	5,249
free D-Glcp unit		5,232	5,231	5,229
$\alpha(1\rightarrow6)$ linked 2,6-di-O-substituted D-Glcp unit			5,187	
α -D-Glcp (1 \rightarrow 5)-D-frup (leucrose)		5,11	5,112	5,113
$\alpha(1\rightarrow2)$ linked D-Glcp unit			5,09 (5)	
$\alpha(1\rightarrow6)$ linked D-Glcp unit	4,976	4,976 (99)	4,977 (94)	4,976 (90)

Anomeric protons were referenced according to internal TSP-d4 (1H = 0 ppm). Numbers into brackets indicates the ratio of $\alpha(1\rightarrow x)$ linked D-Glcp units involved in the synthesized oligosaccharide primary structures, calculated as follows: Ratio of

$\alpha(1\rightarrow x)$ linked D-Glcp units = $\frac{I_x}{I_t} \times 100$, where I_x and I_t correspond to the relative intensities of the anomeric proton

signals of $\alpha(1\rightarrow x)$ linked D-Glcp units and the total $\alpha(1\rightarrow)$ linked D-Glcp units respectively involved in the synthesized oligosaccharide primary structures.

$\alpha(1\rightarrow2)$ or $\alpha(1\rightarrow3)$ arose on the spectra in addition to the signal due to the anomeric proton of the $\alpha(1\rightarrow6)$ linked moiety. Their integration revealed that DSR-S vardel $\Delta 4N$ and ASR C-del produced glucooligosaccharides in which the ratios of $\alpha(1\rightarrow3)$ linkages versus $\alpha(1\rightarrow6)$ linkages are of 1:99 and 10:90, respectively. The resonance complexity of the $\alpha(1\rightarrow3)$ linked D-Glcp residues is probably due to the neighbouring of $\alpha(1\rightarrow6)$ linked D-Glcp residues, as previously described (Cheetham *et al.*, 1991; Van Leeuwen *et al.*, 2008a). Besides, DSR-E enzyme produced glucooligosaccharides containing $\alpha(1\rightarrow2)$, $\alpha(1\rightarrow3)$ and $\alpha(1\rightarrow6)$ linkages in a ratio 5:1:94. These results were fully consistent with the known regioselectivity of the three transglucosidases. A signal-to-noise value of 160 was calculated from the weaker signal detected (ie anomeric proton of the $\alpha(1\rightarrow3)$ linked D-Glcp units of the glucooligosaccharides synthesised by DSR-S vardel $\Delta 4N$) (Figure 31). The corresponding concentration of this residue was 5.5 mM. So considering that a 16 fold lower signal-to-noise ratio would still be acceptable for a correct detection, we estimated that the method sensitivity was in the range of a few hundreds of micromolar. This was satisfactory considering that these analyses were carried out on crude media without any purification steps. Moreover, the standard deviation of the signal intensities between the triplicates was within 3.3 %, which demonstrated the reproducibility of the method. Finally, an entire microtitre plate was analysed in 288 min (one sample every 3 min). The throughput was thus sufficiently high to analyse 480 samples per day and authorize large library screening in a reasonable duration.

Following the workflow depicted on Figure 30, the method was thus used to screen for glucosidic bond specificity, a library of 36 000 *E. coli* clones expressing DSR-S vardel $\Delta 4N$ variants. As described in the SI, this library was obtained by combinatorial mutagenesis of eight residues (D306, F353, N404, W440, D460, H463, T464 and S512) located in the catalytic domain of the parental enzyme, by using the ISOR method (Herman and Tawfik, 2007). A selection pressure previously described (Van Der Veen *et al.*, 2004) and consisting in growing the recombinant clones onto sucrose substrate was first applied to isolate 5 232 clones able to cleave sucrose substrate. Among this selected population, 4 032 clones were retained due to the detection of a polysaccharide bubble formed around the colony indicating a glucansucrase activity. Then, the flow NMR based protocol was applied to this library of variants which were screened for their glucosidic bond specificity in 8 days. The 1D 1H NMR spectra of the oligosaccharides produced by each mutant were standardized by using the TSP-d4 signal intensity, and compared to that obtained for the parental DSR-S vardel $\Delta 4N$. Of the 4 032 screened mutants, 303 (7.5% of the assayed variants) synthesized gluco-oligosaccharides containing a 2 to 6 fold increase of $\alpha(1\rightarrow3)$ linked D-Glcp units compared to those produced by the parental enzyme (Figure 32a).

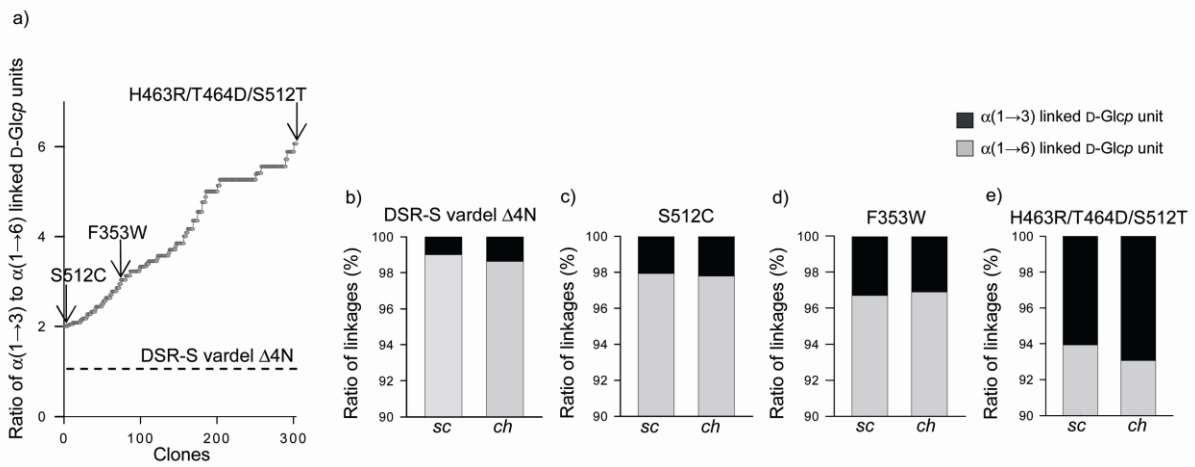


Figure 32: Identification of glucanase mutants with modified specificity by NMR-based screening. a) Ratio of $\alpha(1\rightarrow3)$ to $\alpha(1\rightarrow6)$ linked D-Glcp units in the glucooligosaccharides produced by the 303 isolated variants compared to DSR-S vardel $\Delta 4N$ (black dotted line). These results are also presented in details in supplementary table 3. b) Comparison of synthesized oligosaccharides primary structures determined after screening using flow injection technology (sc) and large-scale characterization (ch) for DSR-S vardel $\Delta 4N$, c) for mutant F353W, d) for mutant S512C and e) for mutant H463R/T464D/S512T.

Of the 303 positive hits, two single mutants (S512C and F353W) and one triple mutant (H463R/T464D/S512T) showing increasing amount of $\alpha(1\rightarrow3)$ linked D-Glcp units were selected to assess the reliability of the method (Figure 32a). They were produced in larger scale and isolated for further characterization. The 1D 1H NMR structural analysis of the oligosaccharides produced by the three variants confirmed their aptitude to synthesise more $\alpha(1\rightarrow3)$ linkages than the parental DSR-S vardel $\Delta 4N$. Integration of the anomeric proton signals of the $\alpha(1\rightarrow3)$ and $\alpha(1\rightarrow6)$ D-Glcp units (5.34 ppm and 4.98 ppm respectively) revealed $\alpha(1\rightarrow3)/\alpha(1\rightarrow6)$ ratios of 1.4:98.6, 2.2:97.8, 3.1:96.9 and 6.9:93.1 for parental DSR-S vardel $\Delta 4N$, S512C, F353W and H463R/T464D/S512T respectively (Figure 33). These values are in agreement with the screening results, showing that the miniaturization and automation of oligosaccharide analysis was reliable (Figure 32b, 32c, 32d and 32e).

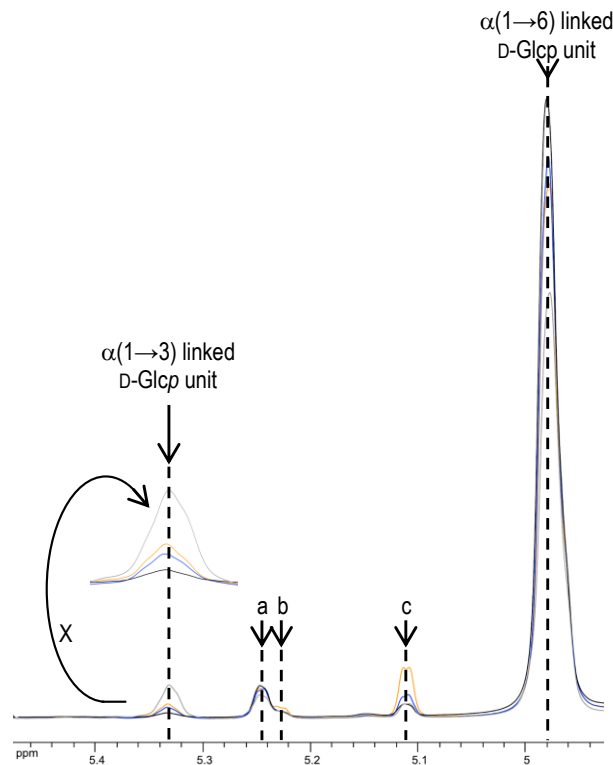


Figure 33: Superimposition of 1D ¹H NMR spectra of the oligosaccharides synthesized by DSR-S vardel Δ4N (black) and the three mutants S512C (blue), F353W (orange) and H463R/T464D/S512T (gray). These mutants show an alteration in their capacity to synthesize α(1→3) glucosidic linkages, visible as disturbance of the relative intensities of α(1→3) linked D-Glcp units anomeric proton signals at 5.33 ppm. Spectra were calibrated according to the internal standard TSP-d4 present in each sample at the same concentration (¹H = 0 ppm, 37 mM). (a) α-D-Glcp reducing unit, (b) free D-Glcp unit and (c) α-D-Glcp (1→5)-D-frup (leucrose).

Structural differences between the oligosaccharides produced by the variants and by the parental enzyme were also confirmed by HPAEC-PAD (Figure 34). Besides, we noticed that the single mutants S512C and F353W also produced higher amounts of leucrose (5-O-α-D glucosyl fructose) compared to the parental enzyme. Indeed, as shown on the HPAEC-PAD profiles, leucrose yield increased from 3 % for DSR-S vardel Δ4N to 6 and 20 % for S512C and F353W respectively. This indicates that the mutations also affected fructose acceptor recognition. This phenomenon was also evidenced from ¹H NMR data, as the intensity of the peak corresponding to the D-Glcp unit anomeric proton of leucrose (δ 5.11 ppm) clearly increased (Figure S-1). These results also proved that the method is useful to detect changes in glycoenzyme stereoselectivity, since it provided more informations than initially expected. Finally, α-glucans were produced by these three variants using sucrose as sole substrate, purified and subjected to 1D ¹H NMR analysis. The NMR data showed that the α(1→3)/α(1→6) ratios in the glucans produced by the parental DSR-S vardel Δ4N and the variants S512C, F353W and

H463R/T464D/S512T were of 4:96, 6:94, 9:91 and 12:88, respectively (Figure 35), thus confirming the change of variant specificity and the potential of the methods to screen for novel bio-polymer synthesising enzymes.

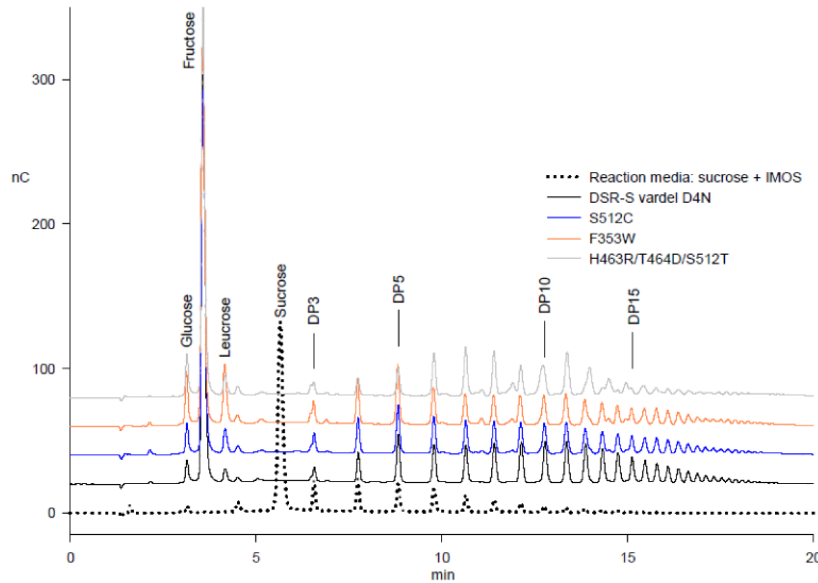


Figure 34: HPAEC-PAD profiles of the reaction products obtained from sucrose and iso-maltooligosaccharides by using parental DSR-S varde1 Δ 4N and variants S512C, F353W and H463R/T464D/S512T. IMOS: Isomalto-oligosaccharides; DP: Polymerisation degree (referenced to IMOS used as acceptors).

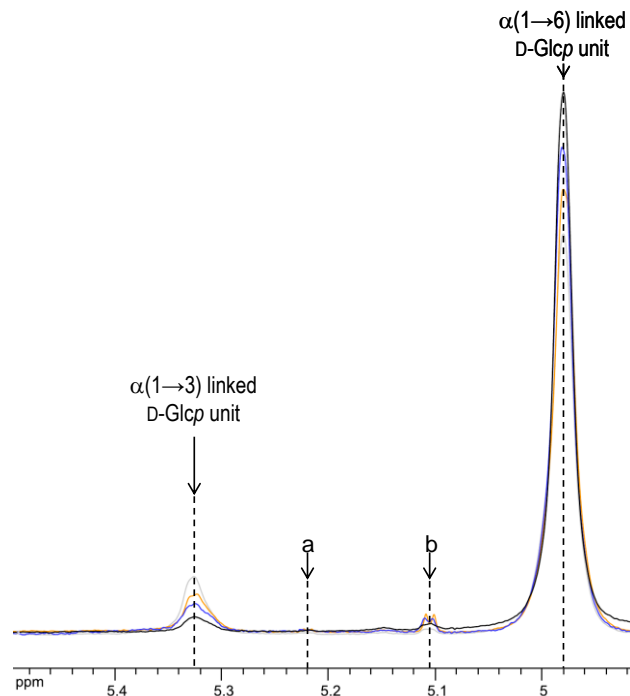


Figure 35: Superimposition of 1D ¹H NMR spectra of the α-glucans synthesized by DSR-S varde1 Δ 4N (black) and the mutants S512C (blue), F353W (orange) and H463R/T464D/S512T (gray) that show an altered specificity. Spectra were calibrated according to the internal standard TSP-d4 present in each sample at the same concentration (¹H = 0 ppm, 37 mM). (a) α-D-Glcp reducing unit, (b) α-D-Glcp (1→5)-D-frup.

In 3 minutes per sample, this flow NMR-based method allowed the detection and quantification of the type of osidic linkages in glucooligosaccharides with a high sensitivity and reproducibility. The method is not only faster than the LC or LC-MS methods used for carbohydrate analysis, but it also provides structural information without any compound separation. Finally, the method is non destructive and can be applied directly on crude cell lysates and natural substrates, as it does not necessitate any labelling strategy. Provided that one is able of identifying an appropriate structural reporter, the methodology can be extended to screen substrate selectivity and linkage specificity of other CAZymes. In a general way, this approach opens promising perspectives in glycomics, to mine artificial and natural enzyme diversity for novel biocatalysts.

Materials and Methods

DNA manipulations. A variant library was constructed by using the ISOR method (Herman and Tawfik, 2007) from the parental gene encoding the dextransucrase DSR-S vardel Δ 4N. An AatII restriction site was first incorporated in the pBAD *Thio-dsrs vardel Δ 4N-His* plasmid (Moulis et al., 2006a) by silent mutation using inverse PCR at position 1881 of the gene *dsr-s vardel Δ 4N*, to generate the pBAD *Thio-dsrs vardel Δ 4N_aatII1881-His* plasmid. Then, a 1130 bp cassette, corresponding to the positions 829 to 1958 of the gene *dsr-s vardel Δ 4N* was PCR amplified from the pBAD *Thio-dsr-s vardel Δ 4N_aatII1881-His* plasmid. Five μ g of purified PCR products were digested with 1U of DNaseI in the supplied buffer at 20°C in a final volume of 50 μ L. The reaction was stopped after 3 min by adding 15 μ L of 0.5 M EDTA and heating 10 min at 75°C. Fragments were separated on a 2% agarose gel and those with a size between 50-100 bp were extracted using QIAquick Gel Extraction Kit (Qiagen). Gene reassembly was carried out with 100 ng of purified fragments in combination with 2 μ M of degenerated oligonucleotides mix, to mutate positions D306, F353, N404, W440, D460, H463, T464 and S512 (DSR-S vardel Δ 4N numbering). The reaction mixture (30 μ L), containing 1U Phusion® High-Fidelity DNA Polymerase (Finnzyme) in the appropriate buffer and 0.4 mM of each dNTPs, was thermocycled according to the following program: 1 denaturation step at 98°C for 30 s; 40 cycles composed of a denaturation step at 98°C for 10 s, 6 successive hybridization steps separated by 4°C each, from 65°C to 41°C for 10 s each and an elongation step at 72°C for 20s; and finally a 2 min step at 72°C. The fully recombined cassettes were isolated from the reassembly products in a last amplification by nested PCR using the primers forK7A3nted (5'-CCACAGTGAATGAACTAGTGAAGATATG-3') and revK7A8nted (5'-ATGGCATCTTTACCATAGCGAACACTT-3'). The purified nested PCR products were digested with *SpeI* and *AatII* restriction enzymes and ligated into the pBAD *Thio-dsrs vardel Δ 4N_aatII1881-His* plasmid to substitute the parental cassette. The ligation products were precipitated by adding 5 volumes

of absolute ethanol and the DNA pellet was rinsed 2 times with 70% ethanol. The resuspended plasmids were transformed into *E. coli* TOP 10 electrocompetent cells and plated on LB agar supplemented with ampicillin (100 µg/ml). The transformants were grown overnight (ON) at 37°C and the colonies were scraped from the plates for plasmid extraction, constituting the glucansucrase DNA libraries.

Selection of glucansucrase active clones. The glucansucrase DNA libraries were transformed into chemiocompetent *E. coli* BL21 AI cells (Invitrogen, Carlsbad, CA) and plated onto LB agar supplemented with 100 µg/ml ampicillin. After overnight growth at 37°C, the cells were scraped, resuspended with physiological water and diluted to an OD_{600nm} of 5.10⁻⁵. The clones were subjected to selection pressure by plating them on 22x22 cm plates containing solid M9 mineral medium (42 mM Na₂HPO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 8.5 mM NaCl, 2.5.10⁻² mM CaCl₂, 1 mM MgSO₄) supplemented with 100 µg/ml ampicillin, 0.02 % arabinose (wt/vol) and 146 mM sucrose as the sole carbon source. The plates were incubated 7 days at 20°C allowing enough growth and expression of the recombinant proteins. Glucansucrase positive variants were picked and transferred to 96 well microplates (Nunc™ Brands Products, Roskilde, Denmark) filled with 250 µL LB per well supplemented with 100 µg/ml ampicillin using a Biomek2000 pipettor (Beckman Coulter, Brea, CA). After overnight growth at 30°C under horizontal shaking at 250 rpm, each starter culture was duplicated into new microplates containing 250 µL LB supplemented with 100 µg/ml ampicillin and 9 % (wt/vol) glycerol. The resulting libraries were stocked at -80°C.

Screening of the active glucansucrase library

Small-scale expression and oligosaccharides synthesis. Starter cultures of active variants were inoculated from storage microplates using the QpixII automate (Genetix, Hampshire, UK). After growth for 24 h at 30°C under agitation, 50 µL of each starter culture were used to inoculate 96-deepwell plates (ABgene, Epsom, UK) containing 500 µL auto-inducing media ZYM-5052 (Studier, 2005) supplemented with 100 µg/ml ampicillin and 0.1 % arabinose (wt/vol). Growth and gene expression were conducted during 48 h at 20°C in an incubator-shaker (INFORS HT, Bottmingen, Switzerland). Plates were centrifuged (5 min, 3700 g, 4°C) and the supernatants removed. Bacterial cell pellets were resuspended in 200 µL of 0.5 mg/ml lysozyme solution, incubated 20 min at 37°C and frozen at -80°C for 12 h. After thawing at room temperature, 800 µL of reaction mixture containing sucrose and iso-maltooligosaccharides (1 kDa, Pharmacosmos, Denmark) diluted in buffered deuterium water (50 mM NaAc, 0.05 g/L CaCl₂, pH 5.2) at a final concentration of 292 mM and 50 mM respectively, were added to each well. Enzymatic reactions were incubated 48 h at 25°C under agitation. After that, 50 µL of

TSP-d4 (trimethylsilyl)propionic acid-2,2,3,3-d4 acid) in D₂O were added to each well at a final concentration of 37 mM. The deepwell plates were then loaded onto the flow injection NMR system rack holders and subjected to 1D ¹H NMR analysis.

NMR spectroscopy. Measurements were performed by using the BEST system combining a Bruker Avance 600 MHz spectrometer with a Gilson Liquid Handler for sample train preparation and injection. A 5 mm TCi (Triple Cryocooler inverse) probe with a 120 µL active volume cryoFIT (CryoProbe Flow Conversion System) was used. Samples were transferred to the probe with a flow rate of 2.6 ml/min and eliminated after analysis. All NMR data were acquired at 298 K and processed with IconNMR software, requiring 3 min per sample. A ZGPR pulse sequence, for presaturation of the residual water signal, was applied and experiments were carried out using the parameters described in tables 8 and 9. The ¹H-signal from D₂O was used for automatic lock and a gradient shimming was performed on each sample. Before Fourier transformation, the FIDs were multiplied by an exponential function with a line broadening of 0.3 Hz. Spectra were processed with a 64 k zero filling, baseline correction and referenced using the TSP-d4 signal at 0 ppm (Tables 8 and 9 relate sample preparation and NMR measurement parameters in detail).

Table 8: Parameters for sample train preparation and automated direct injection of crude reaction media to the NMR spectrometer.

	Sample train	Gap composition	Volume (µL)
	1	Push solvent (H ₂ O)	
Trailing gaps	2	Gas	10
	3	H ₂ O	35
	4	Gas	10
	5	Sample	500
Leading gaps	6	Gas	10
	7	Sample	35
	8	Gas	10
	9	D ₂ O	50
	10	Gas	10

Table 9: NMR measurement parameters.

Temperature	298 K
Time domain	32 k
Number of scans	16
Number of dummy scans	4
Relaxation delay	1 s
Acquisition time	1.95 s
Sweep width	8389 Hz
90° pulse length	8.2 µs
Presaturation radio frequency field	12.4 Hz

Analysis of the products synthesised by variants S512C, F353W and H463R/T464D/S512T

Enzyme production. Variants were grown in flask containing 50 ml of ZYM-5052 medium supplemented with 100 µg/ml ampicillin and 0.1 % (wt/vol) arabinose. Growth and gene expression were conducted during 24 h at 20°C in an incubator-shaker. Bacterial cells were collected by centrifugation (15 min, 4,500 g, 4°C) and resuspended in 3 ml of sodium acetate buffer (50 mM NaAc, 0.05 g.L⁻¹ CaCl₂, pH 5.2). Cells were sonicated and precipitated by centrifugation (3500 g, 30 min, 4°C), supernatants were used as enzymatic extracts.

Oligosaccharide and glucan synthesis. Reactions were carried out at 25°C, using 1 U.ml⁻¹ of enzymes, in sodium acetate buffer (50 mM NaAc, 0.05 g.L⁻¹ CaCl₂, pH 5.2), and 292 mM sucrose for glucan synthesis or 292 mM sucrose and 50 mM iso-maltooligosaccharides for oligosaccharide synthesis, until sucrose depletion. Reactions were stopped by 5 min incubation at 95°C.

HPLC analysis. Oligosaccharide analysis was performed by HPAEC-PAD using a Dionex Carbo-pack PA100 column (250 x 4 mm, Dionex). A gradient of sodium acetate in 150 mM NaOH was applied at 1 ml/min flow rate as following: from 0 to 300 mM in 30 min, 300 to 450 mM in 1 s, 450 to 0 mM in 5 min and 0 mM during 10 min. Detection was performed using an ED40 module with a gold electrode.

NMR spectroscopy. 1D ¹H NMR spectra were recorded on a Bruker Avance 500 MHz spectrometer using a 5 mm z-gradient BBI probe at 298 K, an acquisition frequency of 500.13 MHz and a spectral width of 8012.82 Hz. Oligosaccharide samples were lyophilised and dissolved in 650 µL of D₂O. Glucan samples were precipitated with one volume of absolute ethanol, recovered by centrifugation, washed with distilled water, lyophilised and dissolved in 650 µL of D₂O. Chemical shifts are given in ppm by reference to TSP-d4 (δ 0 ppm).

Acknowledgment

The high throughput screening work was carried out at the Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP, Toulouse, France) with the equipments of the ICEO facility dedicated to the screening and the discovery of new and original enzymes and at the Institut de Pharmacologie et de Biologie Structurale (IPBS, Toulouse, France). ICEO is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund and the Institut National de la Recherche Agronomique (INRA). The NMR spectra were recorded on spectrometers financed with the help of European Structural funds, Région Midi-Pyrénées and Centre National de la Recherche Scientifique (CNRS). They are part of the Integrated Screening Platform of Toulouse (PICT, IBISA).

References

- Bozonnet, S., Dols-Laffargue, M., Fabre, E., Pizzut, S., Remaud-Simeon, M., Monsan, P., Willemot, R.M.**, 2002. Molecular characterization of DSR-E, an α -1,2 linkage-synthesizing dextranucrase with two catalytic domains. *J. Bacteriol.* **184**, 5753-5761.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B.**, 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucl. Acids Res.* **37**, D233-238.
- Cheetham, N.W., Fiala-Beer, E., Walker, G.J.**, 1991. Dextran structural details from high-field proton NMR spectroscopy. *Carbohydr. Polym.* **14**, 149-158.
- Dalvit, C., Ardini, E., Flocco, M., Fogliatto, G.P., Mongelli, N., Veronesi, M.**, 2003. A general NMR method for rapid, efficient, and reliable biochemical screening. *J Am Chem Soc* **125**, 14620-14625.
- Dumas, M.E., Barton, R.H., Toye, A., Cloarec, O., Blancher, C., Rothwell, A., Fearnside, J., Tatoud, R., Blanc, V., Lindon, J.C., Mitchell, S.C., Holmes, E., McCarthy, M.I., Scott, J., Gauguier, D., Nicholson, J.K.**, 2006. Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc. Natl. Acad. Sci.* **103**, 12511-12516.
- Duus, J.O., Gottfredsen, C.H., Bock, K.**, 2000. Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem. Rev.* **100**, 4589.
- Herman, A., Tawfik, D.S.**, 2007. Incorporating synthetic oligonucleotides via gene reassembly (ISOR): a versatile tool for generating targeted libraries. *Protein Eng Des Sel* **20**, 219-226.
- Homann, A., Seibel, J.**, 2009. Chemo-enzymatic synthesis and functional analysis of natural and modified glycostructures. *Nat. Prod. Rep.* **26**, 1555-1571.
- Joucla, G., Pizzut, S., Monsan, P., Remaud-Simeon, M.**, 2006. Construction of a fully active truncated alternansucrase partially deleted of its carboxy-terminal domain. *FEBS letters* **580**, 763.
- Keifer, P.A.**, 2003. Flow NMR applications in combinatorial chemistry. *Curr. Opin. Chem. Biol.* **7**, 388.
- Kittl, R., Withers, S.G.**, 2010. New approaches to enzymatic glycoside synthesis through directed evolution. *Carbohydr. Res.* **345**, 1272.
- Maina, N.H., Tenkanen, M., Maaheimo, H., Juvonen, R., Virkki, L.**, 2008. NMR spectroscopic analysis of exopolysaccharides produced by *Leuconostoc citreum* and *Weissella confusa*. *Carbohydr. Res.* **343**, 1446.
- Monsan, P., Remaud-Siméon, M., André, I.**, 2010. Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr. Opin. Microbiol.* **13**, 293.
- Moulis, C., Arcache, A., Escalier, P.C., Rinaudo, M., Monsan, P., Remaud-Simeon, M., Potocki-Veronese, G.**, 2006a. High-level production and purification of a fully active recombinant dextranucrase from *Leuconostoc mesenteroides* NRRL B-512F. *FEMS Microbiol Lett* **261**, 203-210.
- Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M.**, 2006b. Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J Biol Chem* **281**, 31254-31267.

Reetz, M.T., Eipper, A., Tielmann, P., Mynott, R., 2002. A practical NMR-based high-throughput assay for screening enantioselective catalysts and biocatalysts. *Adv. Synth. Catal.* **344**, 1008-1016.

Studier, F.W., 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234.

Van Der Veen, B.A., Potocki-Veronese, G., Albenne, C., Joucla, G., Monsan, P., Remaud-Simeon, M., 2004. Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett.* **560**, 91-97.

Van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2009. Structural characterization of bioengineered alpha-D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180. *Biomacromolecules* **10**, 580-588.

Van Leeuwen, S.S., Kralj, S., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008a. Structural analysis of bioengineered α -D-glucan produced by a triple mutant of the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180: generation of (α 1 \rightarrow 4) linkages in a native (1 \rightarrow 3)(1 \rightarrow 6)- α -D-glucan. *Biomacromolecules* **9**, 2251.

Van Leeuwen, S.S., Kralj, S., Van Geel-Schutten, I.H., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008b. Structural analysis of the α -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme. *Carbohydr. Res.* **343**, 1237.

Van Leeuwen, S.S., Leeflang, B.R., J., G.G., Kamerling, J.P., 2008c. Development of a ^1H NMR structural-reporter-group concept for the primary structural characterisation of α -D-glucans. *Carbohydr. Res.* **343**, 1114-1111.

Vliegthart, J.F.G., Dorland, L., Van Halbeek, H., 1983. High-resolution, ^1H -nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *Adv. Carbohydr. Chem. Biochem.* **41**, 209.

Weijers, C.A.G.M., Franssen, M.C.R., Visser, G.M., 2008. Glycosyltransferase-catalyzed synthesis of bioactive oligosaccharides. *Biotechnol. Adv.* **26**, 436.

CHAPITRE 3 : Ingénierie combinatoire de la spécificité de la dextrane-saccharase DSR- S

L'identification des déterminants de la spécificité des glucane-saccharases est un enjeu majeur dans la compréhension des relations structure/fonction de ces enzymes mais également pour la conception, à façon, d'oligosaccharides et d' α -glucanes de structures contrôlées. Ainsi, de nombreux travaux d'ingénierie rationnelle et semi-rationnelle réalisés sur les glucane-saccharases de la famille GH70 ont été initiés. En l'absence de données cristallographiques jusqu'en 2010, les travaux se sont focalisés sur des régions bien conservées entre les glucane-saccharases de la famille GH70 et GH13. La région IV, comprenant le stabilisateur de l'intermédiaire covalent, a particulièrement été étudiée et est maintenant connue pour être directement impliquée dans le contrôle de la spécificité de ces enzymes. Cependant, l'absence de méthode de criblage haut-débit était, jusqu'à présent, un frein important dans l'exploration de nouvelles régions et explique en partie la taille limitée des banques générées au cours de ces divers travaux. Ainsi, la mise au point d'une méthodologie qui couple sélection *in vivo* et criblage haut-débit de la spécificité des glucane-saccharases, présentée dans le chapitre précédent, était une étape essentielle pour explorer la diversité de larges banques de mutants.

Dans ce chapitre, nous proposons d'utiliser l'ingénierie combinatoire pour étudier les relations structure-spécificité de la dextrane-saccharase DSR-S vardel $\Delta 4N$, la dextrane-saccharase la plus active connue à ce jour, et pour concevoir de nouvelles glucane-saccharases de spécificité contrôlée. Guidés par l'alignement de séquences de glucane-saccharases de spécificités différentes et l'analyse de la structure cristallographique de la GTF180, nous avons remodelé le domaine catalytique de DSR-S par mutagenèse de saturation et recombinaison simultanée de 8 résidus potentiellement impliqués dans la spécificité de l'enzyme, dont le rôle n'avait, pour certains, jamais été étudié auparavant. La méthode de criblage multi-étapes présentée ci-avant a été mise à profit pour trier les variants obtenus en fonction de leur spécificité.

Afin de mieux comprendre l'effet des mutations sur la nature des produits synthétisés, leur importance dans la reconnaissance du substrat par l'enzyme et dans la réaction de transglucosylation, ce chapitre décrit :

- l'analyse de la distribution des mutations à chaque étape du criblage
- la caractérisation fonctionnelle et cinétique de 7 mutants de spécificité altérée par rapport l'enzyme sauvage, capables de produire des dextrans présentant des taux variables de liaisons $\alpha(1\rightarrow3)$.
- la modélisation de la structure du domaine catalytique de DSR-S et de ses mutants de spécificité altérée

En parallèle à cette étude, nous avons travaillé sur l'ingénierie de la spécificité des glucane-saccharases par une approche totalement aléatoire, le « family shuffling ». En effet, l'équipe CIMES dispose d'un large panel de gènes codant pour des glucane-saccharases de spécificités différentes (DSR-S, DSR-E, ASR et GTF-I). Le family shuffling permet de tirer profit de la diversité naturelle de protéines de fonctions proches. Cependant, la limite de cette approche est liée à la nécessité de disposer de gènes présentant de très fortes similarités (> 70%) pour permettre la recombinaison hétérologue. Or, l'identité entre les séquences codant pour ces quatre glucane-saccharases est comprise entre 25 et 48 %. Ces valeurs atteignent néanmoins 56 et 62 % pour la comparaison des domaines catalytiques, comme l'indique le tableau ci-dessus.

Tableau 10 : Similarité de séquences entre les gènes codant pour les glucane-saccharases DSR-S, DSR-E, GTF-I et ASR.

Protéine entières	DSR-E	DSR-S	GTF-I	Domaines catalytiques	DSR-E*	DSR-S	GTF-I
ASR	25	46	38	ASR	56	61	60
DSR-E		31	32	DSR-E*		59	54
DSR-S			48	DSR-S			62

*la séquence correspond au second domaine catalytique responsable de la synthèse de liaisons de type $\alpha(1\rightarrow2)$.

Nous avons donc entrepris le brassage des gènes codant pour le domaine catalytique de ces 3 enzymes, en utilisant la technique du single-stranded family shuffling pour augmenter la probabilité de recombinaison entre les gènes parentaux. Les travaux ont consisté dans un premier temps à brasser les domaines catalytiques de DSR-S et GTF-I, qui présentent le plus fort pourcentage de similarité (Tableau 10), et de remplacer le domaine catalytique de DSR-S vardel $\Delta 4N$ par les domaines chimères. Malheureusement, seuls 10 % des variants séquencés constituaient des chimères des enzymes parentales, et ne présentaient qu'un point de croisement. La sélection de 768 variants actifs et le criblage de leur spécificité n'a pas permis d'isoler de mutants de spécificité modifiée par rapport à DSR-S. Ces travaux n'ont pas été poursuivis.

Combinatorial engineering of dextranase specificity

Romain Irague, Laurence Roncalli-Tarquis, Isabelle André, Claire Moulis, Sandrine Morel, Pierre Monsan, Gabrielle Potocki-Véronèse, Magali Remaud-Siméon

Abstract

To investigate the relationships between structure and specificity of the dextranase DSR-S from *Leuconostoc mesenteroides* NRRL B-512F and to generate enzyme with altered specificity, semi-rational engineering of the enzyme was carried out. Eight amino acids (D306, F353, N404, W440, D460, H463, T464 and S512) were targeted from sequence and structural analysis of GH70 enzymes, randomized by saturation mutagenesis and simultaneously recombined. Screening of two libraries totalizing $3.6 \cdot 10^4$ clones enabled the isolation of a toolbox comprising 82 variants, which synthesize dextrans with gradually increasing amounts of $\alpha(1 \rightarrow 3)$ linkages, ranging from 4 to 20%. Molecular modelling of DSR-S enzyme revealed that all targeted amino acids were located in the catalytic pocket. In addition, mutant sequence analysis, biochemical characterization and molecular modelling studies revealed for the first time the role of peptide $^{460}\text{DYVHT}^{464}$ in DSR-S linkage specificity. Together with residue S512, this amino acid sequence takes a part in the definition of +2 subsite topology, which is likely to be critical for the enzyme regiospecificity.

Introduction

Bacterial glucansucrases (EC. 2.4.1.) are transglucosidases that synthesize high molecular weight α -glucans, oligosaccharides or glucoconjugates from sucrose, a low cost agroressource, as glucosyl donor. These enzymes exhibit different linkage specificity depending on their structures. With their ability to produce carbohydrates varying in size, structure and physico-chemical properties, they stand as attractive tools for glycodiversification.

The glucansucrases produced by acid lactic bacteria of the genus *Leuconostoc*, *Streptococcus*, *Lactobacillus*, *Exigobacterium* and *Weissella* (André *et al.*, 2010; Van Hijum *et al.*, 2006) are classified into the family 70 of Glycoside-Hydrolases (GH70) (Cantarel *et al.*, 2009), in which 47 glucansucrases have been characterized (http://www.cazy.org/GH70_characterized.html). Dextransucrase DSR-S vardel Δ 4N is the most efficient dextransucrase characterized to date. This enzyme is derived from the native dextransucrase DSR-S from *Ln. mesenteroides* NRRL B-512F, of which the N- and C-terminal regions were partly truncated to enhance its recombinant production by *E. coli* cells, without affecting the dextran synthesis ability or the osidic linkage distribution within the products (Moulis *et al.*, 2006a). Like the native enzyme, dextransucrase DSR-S vardel Δ 4N produces an α -glucan containing 95 % of $\alpha(1\rightarrow6)$ and 5 % of $\alpha(1\rightarrow3)$ linkages from sucrose.

Recently, the crystallographic structures of two glucansucrases from GH family 70, *ie* the truncated forms of the dextransucrase from *Lb. reuteri* 180 (GTF180) (Pijning *et al.*, 2008; Vujicic-Zagar *et al.*, 2010) and the mutansucrase from *S. mutans* (GTF-SI) (Ito *et al.*, 2010; Ito *et al.*, 2011), have been solved. They revealed the same unusual fold organized in five domains (A, B, C, IV and V). Except for domain C, all the domains are built up from discontinuous segments of the polypeptide chain. The catalytic domain A adopts a $(\beta/\alpha)_8$ barrel fold, which is circularly permuted when compared to the related $(\beta/\alpha)_8$ barrel domains of GH13 and GH77 enzymes, which belong to the same GH-H clan. The unique active site of GTF180- Δ N and GTF-SI is defined by a pocket-shaped cavity that allows acceptor glucosylation from their non-reducing end, according to an α -retaining double displacement mechanism, comparable to that of GH13 and GH77 enzymes (Moulis *et al.*, 2006b).

Nevertheless, the low sequence similarity shared by GH70 enzymes and the lack of structural data related to glucansucrases able to form $\alpha(1\rightarrow2)$, $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$ or $\alpha(1\rightarrow6)$ osidic linkages limited over years the identification of structural determinants involved in their linkage specificity. This question has been addressed for many years by rational and semi-rational engineering, uniquely guided by primary structure comparisons of glucansucrases. Altogether, earlier reports have allowed to isolate ~ 70 enzyme variants with altered specificity, and highlight the key role of catalytic domain regions

(Hellmuth *et al.*, 2008; Kralj *et al.*, 2006; Kralj *et al.*, 2005b; Kralj *et al.*, 2008; Monchois *et al.*, 2000; Moulis *et al.*, 2006b; Van Leeuwen *et al.*, 2009; Van Leeuwen *et al.*, 2008a). In particular, conserved regions, similar to the signature regions II, III and IV of GH13 family, surrounding the catalytic residues (D551 and E589, DSR-S numbering) and the transition state stabilizer (D662, DSR-S numbering), have been shown to be crucial for the specificity control of GTF-D (Shimamura *et al.*, 1994), GTF-R (Hellmuth *et al.*, 2008), GTF-180 (Van Leeuwen *et al.*, 2009; Van Leeuwen *et al.*, 2008a) dextranases, GTF-B (Shimamura *et al.*, 1994) and GTF-I mutanases (Monchois *et al.*, 2000), GTF-A reuteranase (Kralj *et al.*, 2006; Kralj *et al.*, 2005b) and ASR alternanase (Moulis *et al.*, 2006b). Moreover, the rational design of chimeric reuteranases (Kralj *et al.*, 2008) and dextranases (Funane *et al.*, 2004) has also allowed to demonstrate the importance of regions located at the N-terminal and the C-terminal parts of the catalytic domain (upstream and downstream of the $(\beta/\alpha)_8$ barrel, respectively) in linkage specificity. However, the lack of automated high-throughput methods to screen for linkage specificity prevented the exploration of large sequence spaces generated by combinatorial engineering approaches. To our knowledge, the largest glucanase library screened for specificity contained ~ 2,000 clones (Hellmuth *et al.*, 2008). To circumvent these limitations, we recently developed a straightforward, sensitive and quantitative NMR-based method for screening carbohydrate active enzyme specificity at a throughput of 480 enzyme variants per day (Irague *et al.*, 2011). This automated method was applied to screen 36,000 *E. coli* clones expressing variants of dextranase DSR-S vardel $\Delta 4N$ and sort out 303 clones producing enzymes with altered specificity.

Starting with DSR-S vardel $\Delta 4N$ encoding gene, the present paper describes the rationally guided combinatorial engineering approach used to generate a toolbox of glucanases able to catalyse the formation of dextrans with gradually increasing amounts of $\alpha(1\rightarrow 3)$ linkages. Sequencing data obtained at each step of our protocol enabled us to analyze the quality of the library generated and to evaluate the tolerance to mutations of the targeted amino acids. Seven variants catalyzing dextrans with increasing $\alpha(1\rightarrow 3)$ linkages content were further analyzed. With the support of molecular modelling of both DSR-S enzyme and variants, their functional and kinetic characterization enabled us to gain insight in the structural determinants involved in substrate recognition and linkage specificity.

Results and discussion

We generate diversity on eight positions corresponding to the amino acids D306, F353, N404, W440, D460, H463, T464 and S512 of the DSR-S vardel Δ 4N enzyme. From sequence alignment and structural analyses, these residues were suspected to play a role in the linkage specificity of DSR-S vardel Δ 4N catalytic domain. To generate sequence diversity, we apply the ISOR method (Herman and Tawfik, 2007), which allows i) the randomization of mutations at several amino acid positions (including neighbouring residues), and ii) the simultaneous recombination of these mutations.

Selection of amino acid targets

Twenty four glucansucrase catalytic domains were clustered on the basis of their linkage specificity and their sequences were aligned. Of the selected enzymes, sixteen dextransucrases exhibit mainly $\alpha(1\rightarrow6)$ linkage specificity, four mutansucrases from *Streptococcus sp.* are specific for $\alpha(1\rightarrow3)$ linkages and two reuteransucrases from *Lb. reuteri* species catalyse both $\alpha(1\rightarrow4)$ and $\alpha(1\rightarrow6)$ linkages (Kralj *et al.*, 2005a; Kralj *et al.*, 2004). We also included in the alignment the sequence of alternansucrase from *Ln. citreum* B-1355 (ASR) which catalyses the formation of alternating $\alpha(1\rightarrow3)/\alpha(1\rightarrow6)$ linkages (Argüello-Morales *et al.*, 2000; Côté and Robyt, 1982) and the sequence of the second catalytic domain of DSR-E (DRS-E CD2) from *Ln. citreum* NRRL B-1299 which was shown to catalyse exclusively the synthesis of $\alpha(1\rightarrow2)$ glucosidic linkages (Fabre *et al.*, 2005). Four highly conserved segments emerged from the alignment among which are found the regions homologous to signature region II, III and IV of GH13 family, which are closed to the catalytic residues (Figure 36). The variations observed in the sequence alignment determined the choice of the amino acids that were mutated.

Within segment S304 - N317 (DSR-S vardel Δ 4N numbering, Figure 36), residue D306 is found well conserved amongst dextransucrases from *Ln. mesenteroides* species. However, this residue is substituted by a glutamine in ASR and by an alanine in the DRS-E CD2. In reuteransucrases, Leucine or Isoleucine are found at this position. Downstream of this region, the segment G352 - N374 contains an aromatic residue (Phenylalanine or Tyrosine) at position 353 (DSR-S vardel Δ 4N numbering) for dextransucrases highly specific for $\alpha(1\rightarrow6)$ linkages. In contrast, dextransucrase GTF180 which is less specific for $\alpha(1\rightarrow6)$ linkages, reuteransucrases and alternansucrase exhibit an alanine, a serine or a glutamine at this position. In the signature regions II and III of GH13 family, residues N404 and W440, downstream the catalytic nucleophile D400 and upstream the general acid/base catalyst E438, respectively, are strictly conserved amongst glucansucrases except for DSR-E CD2. The equivalent residues of GH13 enzymes (H210 and L232 of the Taka-amylase A from *Aspergillus oryzae*) are known

to interact with the substrate aglycon (MacGregor *et al.*, 1996). Moreover, equivalent residues in GTF180 (N1029 and W1065) were identified to be part of the subsite +1 for substrate (sucrose) and acceptor (maltose) molecules (Vujicic-Zagar *et al.*, 2010). Based on these analyses, we considered D306, F353, N404, W440 residues likely to be important for linkage specificity. Another residue retained our attention, the amino acid Ser512. As a part of the tripeptide ⁵¹²SEV⁵¹⁴, it is located just downstream of the second aspartic acid of catalytic triad in the region IV (Kralj *et al.*, 2006; Moulis *et al.*, 2006b; Van Leeuwen *et al.*, 2009). As it was already shown to affect linkage specificity in various reports, it was also selected as a target.

Finally, three additional positions, located in a non-conserved region of GH70 enzymes, were targeted from structural analysis of the first available GH70 crystallographic structure, that of GTF180-ΔN in complex with maltose (PDB code: 3KLL). Structural analysis highlighted water mediated and direct hydrogen bond interactions between four residues (D1085, R1088, N1089 and S1137) of +2 subsite with the reducing end of maltose, thus participating to the correct positioning of maltose in the active site to favour its glucosylation at the C6-hydroxyl group. By analogy, the corresponding residues of DSR-S vardel Δ4N (D460, H463, T464 and S512) can also be suggested to participate in linkage specificity and were selected.

Altogether, this information led us to focus engineering efforts on the 8 residues above described (D306, F353, N404, W440, D460, H463, T464 and S512) in order to redesign the DSR-S active site and explore recombination of mutations at these positions using the ISOR method.

Analysis of the initial library

Two libraries were thus created: library "LibA", corresponding to the randomized mutagenesis targeted on amino acids D306, F353, N404, W440 and the library "LibB", focused on amino acids D460, H463, T464 and S512. The synthetic oligonucleotides used for saturation mutagenesis were designed to contain the conventional degenerated NNS codons at the targeted positions and give access in theory to the 20 possible amino acid residues.

A total of $2.5 \cdot 10^5$ and $3.8 \cdot 10^4$ clones were obtained for LibA and LibB, respectively. Twenty four and fifty one clones were randomly picked up from LibA and LibB, respectively, and sequenced. Analysis of sequencing data for the two libraries indicates that mutations occur at each position with a similar distribution (Table 11). For both libraries, the mean number of mutations per protein sequence is 3. This demonstrates the efficiency of the ISOR method to generate multiple mutations with equivalent frequency on both distant and adjacent residues, such as those found in LibA and LibB, respectively. In addition, the sequences of LibA variants were all different whereas in the case of LibB, three single mutants, encoded by the same nucleotide sequences, were obtained twice at position 512 (S512C,

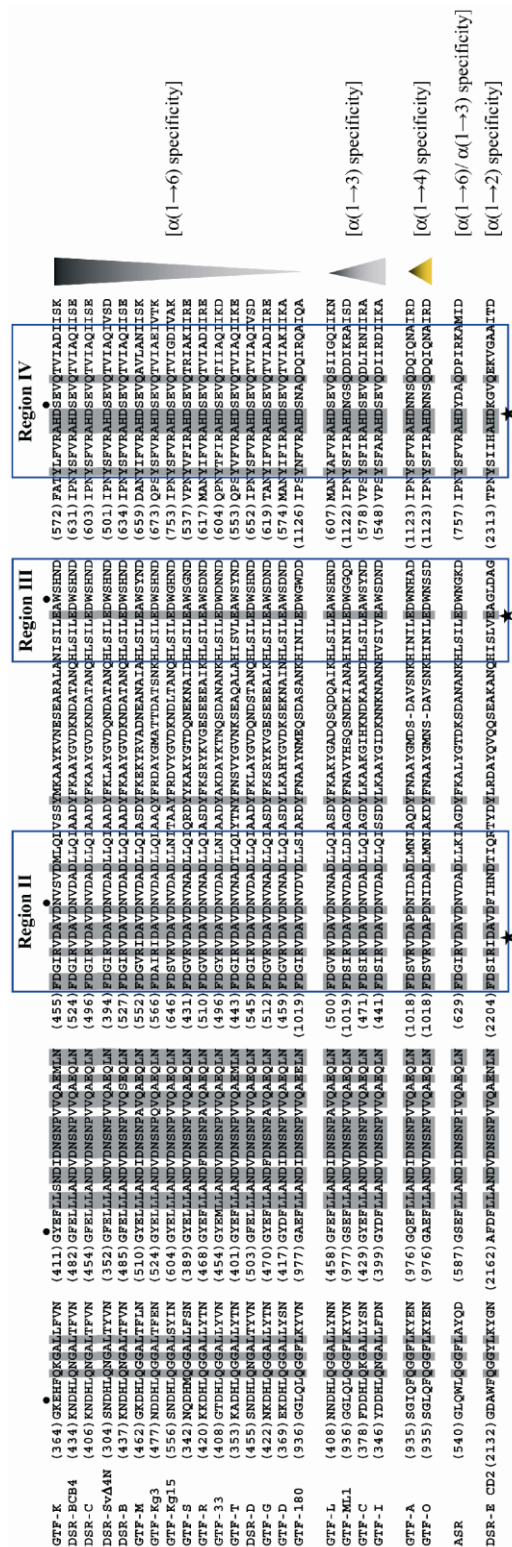


Figure 36: Alignment of GH70 amino acid sequences in the regions chosen for the combinatorial site directed mutagenesis of DSR-S varde1 Δ4N. Sequences used are clustered according to the specificity of the characterized glucansucrases. Dextransucrase sequences are ordered from the higher to lower ratio between the α(1→6) and the α(1→3) linkages contained in the α-glucans produced by these enzymes. Mutansucrase sequences are ordered from the lower ratio between the α(1→3) and the α(1→6) linkages contained in the produced α-glucans. Reuteransucrase sequences are ordered from the higher to lower ratio between the α(1→4) and the α(1→6) linkages. aa: strictly conserved amino acids; ★: catalytic amino acids (D400 and E438, DSR-S varde1 Δ4N numbering) and transition state stabilizer (D511, DSR-S varde1 Δ4N numbering); •: amino acids chosen for mutagenesis.

S512L and S512G). This indicates that a bias may have occurred in the construction of LibB library that is probably due to disequilibrium in the composition of the degenerated oligonucleotides (SI Table 1, p.146).

Selection and screening of active glucansucrase variants

The DNA isolated from the initial libraries LibA and LibB was used to transform *E. coli* BL21 AI cells. A total of $1.5 \cdot 10^4$ and $2.1 \cdot 10^4$ clones, for libraries LibA and LibB respectively, were plated on sucrose selective medium which promotes only the growth of clones producing enzyme variants able to cleave sucrose (Van Der Veen *et al.*, 2004). Moreover, if the produced glucansucrase synthesized α -glucan (polymerase positive clones), a polymer bubble above the colony was easily detected. Altogether, selection and visual screening enabled us to detect 216 clones exhibiting polymerase activity from LibA. They represent less than 2 % of the selected clones. Sequence analysis revealed that 80 % of the mutations were located at position 353, the remaining 20 % being equally distributed over the three other positions. Mutation at positions 306, 404 and 440 seem to drastically affect substrate recognition and/or transglucosylation activity, while position 353 is more tolerant to mutations.

From LibB, 3,816 clones exhibiting polymerase activity were detected. They represent 20 % of the selected variants. Mutational spectrum revealed that 44 % of the mutations occurred at position 512. Residue S512 is thus probably less essential for sucrose accommodation and polymerisation activity than D460, H463 and T464. Noteworthy, the single mutant DSR-S varde1 Δ 4N S512Y was found unable to produce any glucan in earlier report (Moullis *et al.*, 2006b), in consistency with the fact that such mutation is not found among sequenced mutants isolated upon selection and screening.

Variants with altered linkage specificity

As previously described, we subjected the 4,032 polymerase active clones to a secondary high-throughput NMR based screening method, to discriminate glucansucrase variants with altered specificity. Screening efforts enabled the identification of 14 clones from LibA and 289 from LibB displaying altered osidic linkage specificity compared to that of the parental enzyme (Irague *et al.*, 2011). Sequencing of the variants isolated from this second screening step revealed that the mutation distribution was similar to that observed from sequencing of the polymerase active clones (Table 11). The mutants producing oligosaccharides with the highest $\alpha(1\rightarrow3)/\alpha(1\rightarrow6)$ linkage ratio (up to 8 times more than the parental enzyme) were obtained from library LibB (SI Table 2, p.147). We retained 7 of these variants that synthesized increasing amount of $\alpha(1\rightarrow3)$ linkage (Table 12).

Table 11: Sequence analysis of the LibA and LibB variants and evolution of the mutation distribution at each step of the screening strategy.

	Clone nbr.		Sequenced variant nbr.		mutation distribution evolution	
	LibA	LibB	LibA	LibB		
Initial libraries (<i>E. coli</i> TOP 10)	2.5.10 ⁵	3.8.10 ⁴	24 (nt:24, pt:24)*	51 (nt:48, pt:43)	LibA	
Screened libraries (<i>E. coli</i> BL21 AI)	1.5.10 ⁴	2.1.10 ⁴			Initial library	D306 F353 N404 W440
Selected clones					Selected clones	
<i>hydrolase</i>	1.2.10 ³	0			Hits with altered specificity	
<i>polymerase</i>	216	3,816	66 (nt: 52, pt: 30)	102 (nt:64, pt:36)	LibB	D460 H463 T464 S512
Hit clones with altered specificity	14	289	14 (nt:4, pt:3)	253 (nt:126, pt: 79)	Initial library	
					Selected clones	
					Hits with altered specificity	

*number of different sequences at the DNA (nt) and protein (pt) level.

Functional and kinetic characterization of the selected variants

Purified glucans produced from sucrose (292 mM) with the seven variants were analysed by 1D ¹H NMR measurements (Table 12). The estimation of the $\alpha(1\rightarrow3)/\alpha(1\rightarrow6)$ ratio was fully consistent with that obtained for oligosaccharide synthesis (which was used for NMR-based screening) (Irague *et al.*, 2011). The lowest variations of $\alpha(1\rightarrow3)/\alpha(1\rightarrow6)$ linkage ratio were observed for mutants exhibiting only one mutation (F353T, S512C, F353W). All the mutants that synthesized more than 10 % of $\alpha(1\rightarrow3)$ linkages harboured 3 to 4 combined mutations, the highest value of $\alpha(1\rightarrow3)$ linkages (20 %) being obtained for the quadruple mutant D460M/H463Y/T464/S512C. Notably, mutations at these positions had never been explored before and are here revealed to affect linkage specificity. HPSEC analysis of the reaction products was also performed. Like the parental enzymes, the variants produced two populations of glucans referred as High Molar Mass dextrans (molar mass higher than 10⁷ g.mol⁻¹) and Intermediate Molar Mass dextrans (molar mass ranging from 0.5.10³ to 5.7.10³ g.mol⁻¹). In addition, all of them also catalyse the formation of leucrose, glucose and fructose (Figure 37). Except for the two single mutants F353T and F353W, product yields are only slightly affected by the mutations, suggesting that mutations introduced at positions D460, H463, T464, and S512 do not alter enzyme acceptor binding ability. Notably, mutants F353T and F353W displayed a 3 to 4-fold increase in leucrose yield, indicating that residue F353 may interact in the positioning of the fructosyl moiety. Interestingly, single mutant F353T exhibits a 15-fold and 76-fold reduction of *K_m* value compared to the parental enzyme

and F353W mutant, respectively. This suggests that the aromatic ring of F and W residues may affect fructosyl ring accommodation of fructose.

Table 12: Products synthesized by the parental *Thio*-DSR-S vardel $\Delta 4N$ -His enzyme and the seven selected variants, from 292 mM sucrose, as determined by HPSEC and 1D 1H NMR spectroscopy.

Enzyme	Distribution of products from sucrose (%) [*]				HPSEC elution yield (%)	α -glucan linkages (%) ^a	
	HMW glucan	LMW glucan	Leucrose	Glucose		$\alpha(1\rightarrow3)$	$\alpha(1\rightarrow6)$
<i>Thio</i> -DSR-S vardel $\Delta 4N$ -His	60.6	31.1	5.9	2.4	95.6	4	96
F353T	33.9	36.3	22.6	7.1	75.2	3	97
S512C	68.8	18.7	9.3	3.2	89.2	6	94
F353W	47.8	22.5	22.5	7.1	100	9	91
H463R/T464D/S512T	53.6	32.6	8.8	5.0	97.3	12	88
H463R/T464V/S512T	56.4	31.4	7.7	4.5	90.2	15	85
D460A/H463S/T464L	54.3	37.2	6.1	2.4	99.0	15	85
D460M/H463Y/T464M/S512C	54.1	31.7	7.3	6.9	99.4	20	80

^{*}The percentage distribution from sucrose corresponds to the relative amount of glucosyl units incorporated into product from sucrose.

^aRatio of $\alpha(1\rightarrow3)$ linked D-Glcp units = $I_{(\alpha(1\rightarrow3))} / (I_{(\alpha(1\rightarrow3))} + I_{(\alpha(1\rightarrow6))}) \times 100$, where $I_{(\alpha(1\rightarrow3))}$ and $I_{(\alpha(1\rightarrow6))}$ correspond to the relative intensities of the anomeric proton signals of $\alpha(1\rightarrow3)$ and $\alpha(1\rightarrow6)$ linked D-Glcp units, respectively.

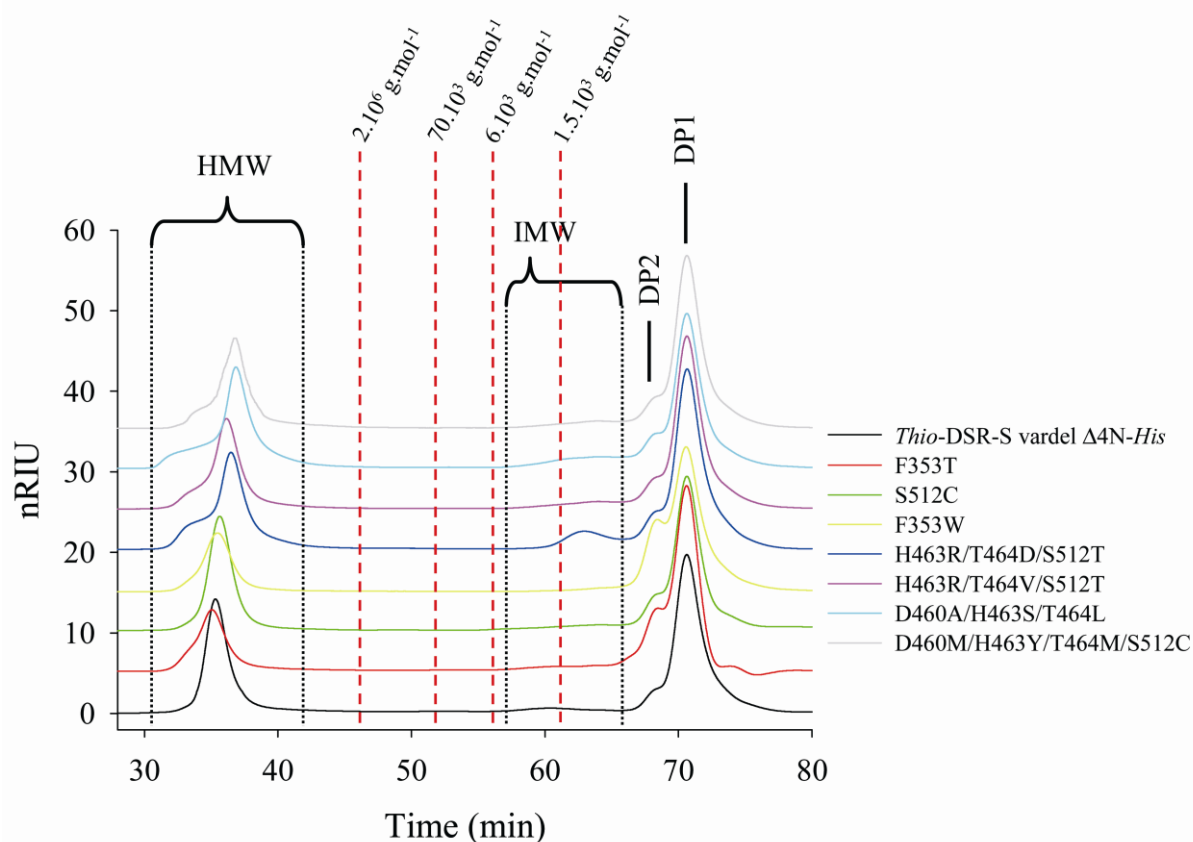


Figure 37: HPSEC chromatogram of products synthesized by *Thio*-DSR-S vardel $\Delta 4N$ -His and the 7 selected variants, from 292 mM sucrose. HMW: High Molecular Weight polymers; IMW: Intermediary Molecular Weight products; DP2: disaccharides; DP1: monosaccharides. Red dotted lines indicate the retention time of standard dextrans.

As previously shown for the parental enzyme (Moulis *et al.*, 2006b), kinetic analyses revealed that sucrose inhibition occurs for all the variants except for F353W and H463R/T464V/S512T mutants. Moreover, ten to thirty fold reduced catalytic efficiencies were observed for 4 mutants (D460M/H463Y/T464M/S512C, D460A/H463S/T464L, H463R/T464D/S512T and F353T) whereas the three other ones were found as efficient as the parental enzyme (Table 13). Reasonable catalytic efficiencies were maintained. Notably, mutant H463R/T464V/S512T is slightly more efficient than the parental enzyme, synthesizes 15 % $\alpha(1\rightarrow3)$ linkages and is not inhibited by sucrose (initial rate at 600 mM sucrose is 3 times higher than that of the parental enzyme). In contrast, variant H463R/T464D/S512T, differing only by the nature of the amino acid substitution at position 464, is found to be 8 times less efficient than parental enzyme. This indicates a critical role of this position.

Table 13: Linkage specificity and kinetic properties of the 8 selected variants, in comparison with those of the WT *Thio*-DSR-S vardele $\Delta 4N$ -His.

Enzyme	Culture	Purification	Kinetic parameters			
	Act. (U.ml ⁻¹)	Spec. Act. (U.mg ⁻¹)	K _m (mM)	K _s (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _m (s ⁻¹ .mM ⁻¹)
<i>Thio</i> -DSR-S vardele $\Delta 4N$ -His	70.3	108.3	7.75*	326*	584*	75.4
F353T	110	24.9	115.8	46.6	338.5	2.9
S512C	18	76	3.9	405.8	207.1	53.1
F353W	44.6	27.6	1.5	-	86	57.3
H463R/T464D/S512T	17.6	10.8	15.5	201.7	172.1	11.1
H463R/T464V/S512T	13	22.6	6.9	-	610.2	88.4
D460A/H463S/T464L	6.8	8.5	5.3	552.7	43.9	8.3
D460M/H463Y/T464M/S512C	2.5	7.5	15.5	175.1	38.2	2.5

*Moulis *et al.* (2006)

3D- modelling of DSR-S catalytic domain

Homology modelling techniques were used to generate a 3D-model of DSR-S vardele $\Delta 4N$ catalytic domain using as template the structure of *Lb. reuteri* GTF180 and the sequence alignment provided in Figure 36. The model was then checked for main chain conformations using the VADAR web server (Willard *et al.*, 2003). The geometry analysis showed that less than 4 % of the Φ - Ψ plots fell into the generously allowed and disallowed regions of the Ramachandran's plot. Most of the residues which were in the generously allowed and disallowed regions belonged to the loop regions in which major insertions and deletions were made. The overall protein fold of DSR-S vardele $\Delta 4N$ is shown in Figure 38. On the basis of the 3D model built for parental DSR-S vardele $\Delta 4N$, we also generated a 3D model for the D460M/H463Y/T464M/S512C mutant, of which the linkage specificity was the most affected compared to parental enzyme. As for GTF180 structure, the three domains, A, B and C of DSR-S vardele

$\Delta 4N$ were defined by non continuous sequence regions (Figure 38). Domain A, the $(\beta/\alpha)_8$ -barrel catalytic domain is composed of three stretches of amino acids 364-614, 745-880 and 917-959. Domain B is inserted between helix $\alpha 1$ and strand $\beta 8$ and consists of regions 296-364, 880-917 and 959-973 and domain C is composed by residues 614-745 and forms the central part of the U-shaped protein.

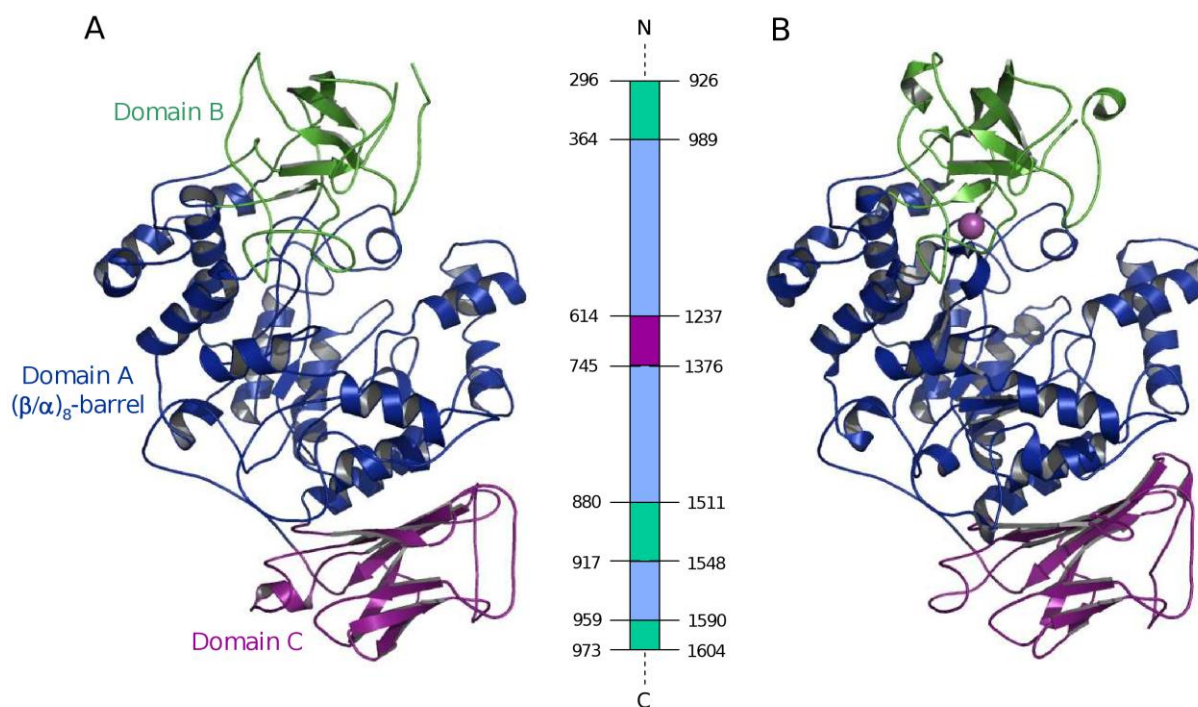


Figure 38: (A) 3D-model of DSR-S varde1 $\Delta 4N$ catalytic domain. (B) Catalytic domain of *Lactobacillus reuteri* N-terminally truncated glucansucrase GTF180 (PDB accession code: 3KLLK).

Catalytic site description and docking of substrate and putative products

The 3D-model of DSR-S varde1 $\Delta 4N$ with sucrose bound in the catalytic site was used to describe binding interactions with the glucosyl ring at subsite -1 and the fructosyl ring at subsite +1 (Figure 39A). Topology of enzyme active site appears to be a shallow pocket with putative catalytic residues located at the bottom, containing the catalytic diade composed of the acid/base E438 (E1063 in GTF180) and nucleophile D400 (D1025 in GTF180), and the transition state stabilizer D511 (D1136 in GTF180). Most amino acids involved in sucrose binding interactions are conserved in GTF180. Glucosyl moiety in subsite -1 stacks onto Y834 and makes hydrogen bonding interactions with H510, D511 and Y880 while β -D-fructosyl moiety in subsite +1 interacts mostly with D511.

The 3D-model of DSR-S varde1 $\Delta 4N$ was used to locate the eight targeted amino acids (Figure 39B). All the mutated positions belong to the first shell of amino acid residues at less than 20 Å away from

catalytic residues, residue S512 being the nearest as it is adjacent to catalytic D511. All targeted residues could thus potentially play a critical role in the binding of sucrose and oligosaccharides in the catalytic site. This is in agreement with the data obtained from the library analysis and kinetic characterization of the 7 mutants with altered specificity.

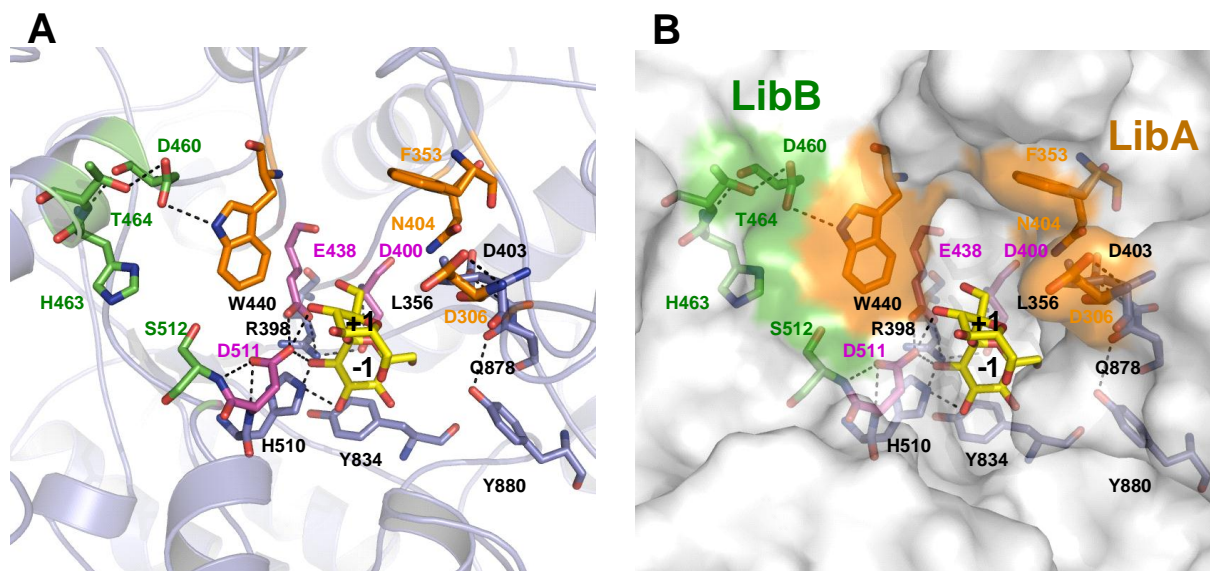


Figure 39: (A) Molecular docking of sucrose in the active site of DSR-S vardele Δ4N. (B) Overview of mutation sites affecting DSR-S vardele Δ4N linkage specificity. Docked sucrose molecule is shown for reference. Targeted residues are shown in orange and green for LibA and LibB respectively.

To get some insight on acceptor binding subsites and influence of mutations on acceptor recognition, we performed docking studies of the tetrasaccharide α -D-Glcp-(1→6)[α -D-Glcp-(1→3)] α -D-Glcp-(1→6)-D-Glcp in D460M/H463Y/T464M/S512C mutant, which synthesizes the highest level of α (1→3) linkages. This tetrasaccharide is a putative product of glucosyl transfer onto isomaltotriose with the formation of α (1→3) linkages. The favoured binding mode of the tetrasaccharide is shown in Figure 40. The glucosyl ring at subsite -1 is bound in a very similar way to the glucosyl unit of sucrose. The glucosyl unit at subsite +1 was also found in productive conformation with respect to acid/base E438 for the formation of an α (1→3) linkage. The +2 subsite accommodates the non reducing end of isomaltotriose which is through van der Waals interactions with W440T. The +2' subsite is formed by a small pocket between loops carrying respectively catalytic residues D400 and E438. Interestingly, the topology of this pocket is also defined by amino acids F353 and N404, targeted in LibA, what might suggest that mutations F353W and N404G identified at these two positions could affect binding pocket topology and thus acceptor recognition. Interestingly, positions carrying mutations affecting the linkage specificity, such as in D460M/H463Y/T464M/S512C mutant, are located in close spatial vicinity to W440

(+2 subsite). In particular, H463Y mutation could introduce an improved stacking platform to interact with longer oligosaccharide chains and even favour $\alpha(1\rightarrow6)$ oligosaccharide binding. Presence of a polar side chain at position 512 (Ser in the parental enzyme or Cys in the mutant) also appears to be required for establishing hydrogen bonding interactions with hydroxyl groups of acceptor glucosyl unit bound at +2 subsite, confirming the putative role of S512 in oligosaccharide binding suggested in an earlier report (Moulis *et al.*, 2006b).

The introduced mutations at subsite +2 seem to promote productive binding of $\alpha(1\rightarrow6)$ linked glucooligosaccharides in the active site and favour the $\alpha(1\rightarrow3)$ branching transglucosylation of glucan chains. However, the branching pattern of products synthesized by mutant D460M/H463Y/T464M/S512C will have to be characterized in more details by analytical techniques to better understand how the protein accommodates oligosaccharide molecules in the catalytic pocket, compared to parental DSR-S vardele $\Delta 4N$.

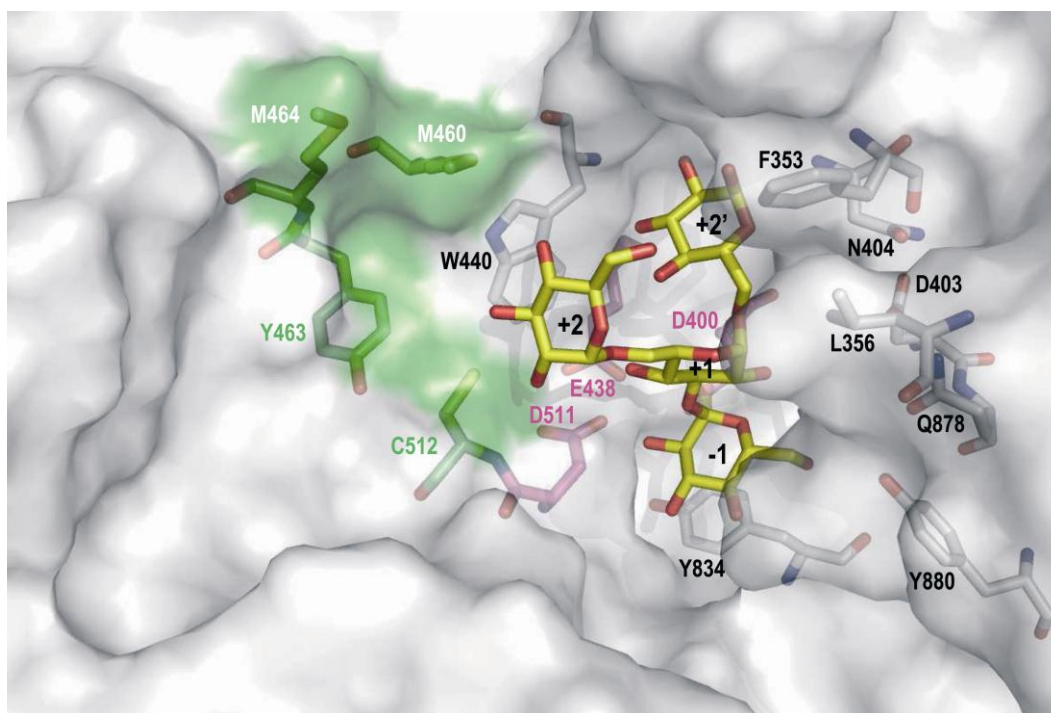


Figure 40: Molecular docking of α -D-Glcp-(1 \rightarrow 6)[α -D-Glcp-(1 \rightarrow 3)] α -D-Glcp-(1 \rightarrow 6)-D-Glcp tetrasaccharide in the active site of D460M/H463Y/T464M/S512C mutant.

Conclusions

In this work, we demonstrate the efficiency of combinatorial engineering and high-throughput screening to investigate relationships between structure and specificity of glucansucrases, and for isolating novel enzymatic tools for glycodiversification. Focusing on the DSR-S dextranucrase catalytic domain, we identified key positions involved in the control of dextranucrase specificity. In particular, this is the first time that the role of peptide ⁴⁶⁰DYVHT⁴⁶⁴ was investigated and demonstrated to be critical for $\alpha(1\rightarrow3)$ linkages formation. Modifications of acceptor binding at subsite +2 are likely to be responsible for the promotion of $\alpha(1\rightarrow3)$ linkage formation. In addition, the semi rational design of catalytic site-directed library of variants combined with efficient screening protocols enabled the isolation of eighty-two variants producing dextrans with gradually increasing amounts of $\alpha(1\rightarrow3)$ linkages. Physico-chemical properties of these original bio-sourced products will be investigated in near future, to evaluate their potential as novel biomaterials and texturants.

Materials and methods

Sequence alignment

The sequences of the glucansucrase from *S. salivarius* ATCC 25975 GTF-K (GenBank accession no. AAA26896.1), *Ln. mesenteroides* NRRL B-1299CB4 DSR-BCB4 (GenBank accession no. ABF85832.1), *Ln. citreum* B-1355 DSR-C (GenBank accession no. CAB76565.1), *Ln. citreum* NRRL B-1299 DSR-B (GenBank accession no. AAB95453.1), *S. salivarius* ATCC 25975 GTF-M (GenBank accession no. AAC41413.1), *Lb. fermentum* Kg3 GTF-kg3 (GenBank accession no. AAU08008.1), *Lb. sakei* Kg15 GTF-kg15 (GenBank accession no. AAU08011.1), *S. downei* MFE 28 GTF-S (GenBank accession no. AAA26898.1), *S. oralis* ATCC10557 GTF-R (GenBank accession no. BAA95201.1), *Ln. citreum* NRRL B-1299 DSR-A (GenBank accession no. AAB40875.1), *Lb. parabuchneri* 33 GTF-33 (GenBank accession no. AAU08006.1), *S. sobrinus* B13N /OMZ176 GTF-T (GenBank accession no. AAX76986.1), *Ln. mesenteroides* Lcc4 DSR-D (GenBank accession no. AAG61158.1), *S. gordonii* str. Challis substr. CH1 GTF-G (GenBank accession no. AAC43483.1), *S. mutans* GS 5 GTF-D (GenBank accession no. AAA26895.1), *Lb. reuteri* 180 GTF-180 (GenBank accession no. AAU08001.1), *S. salivarius* ATCC 25975 GTF-L (GenBank accession no. AAC41412.1), *Lb. reuteri* ML1 GTF-ML1 (GenBank accession no. AAU08004.1), *S. downei* Mfe 28 GTF-I (GenBank accession no. AAC63063.1), *Lb. reuteri* 121 GTF-A (GenBank accession no. AAU08015.1), *Lb. reuteri* ATCC 55730 GTF-O (GenBank accession no. AAY86923.1), *Ln. citreum* B-1355 ASR (GenBank accession no. CAB65910.2), *Ln. citreum* NRRL B-1299 DSR-E (GenBank accession no. CAD22883.1) and DSR-S vardel Δ 4N were aligned using AlignX (Vector NTI Advance 10, Invitrogen) and the multiple alignment editor Jalview (Waterhouse *et al.*, 2009).

Bacterial strains, plasmids and growth conditions

One shot® *E. coli* TOP 10 (Invitrogen) was used for mutant library construction. One shot® *E. coli* BL21 AI™ (Invitrogen) was used for library selection, screening and larger scale production of the selected variants. Plasmid pBAD harbouring the gene *dsr-s* vardel Δ 4N (Moulis *et al.*, 2006a) fused to thioredoxine and polyhistidine tags was used as template for library construction. Bacterial strains were grown on LB medium supplemented with 100 μ g/ml ampicillin at 37°C.

DNA manipulations

Restriction endonucleases and DNA-modifying enzymes were purchased from New England Biolabs and used according to manufacturer's instructions. DNA plasmid isolation was performed using Qiaprep® Spin miniprep kit (Qiagen). PCR products extraction from agarose gel was carried out using

Quiquick® gel extraction kit (Qiagen). DNA sequencing was carried out using the dideoxy chain termination procedure (Sanger *et al.*, 1977) by GATC Biotech SARL (Mulhouse, France).

Generation of DSR-S vardel Δ 4N mutant libraries

An *AatII* restriction site was first incorporated into the *dsr-s* vardel Δ 4N gene, at the position 1881, by silent mutation using inverse PCR and pBAD *Thio-dsrs* vardel Δ 4N-*His* plasmid as template (SI Table 1). Then, a 1130bp cassette, corresponding to positions 829 to 1958 of the gene *dsr-s* vardel Δ 4N, was PCR amplified using primers for *_DSRS_Alpha3* and *rev_DSR-S_Alpha8* (SI Table 1). 5 μ g of purified PCR products were digested with 1 U of DNaseI in the supplied buffer at 20°C in a final volume of 50 μ L. The reaction was stopped after 3 min by adding 15 μ L of 0.5 M EDTA and heating 10 min at 75°C. Fragments were separated on a 2 % agarose gel and those with a size between 50-100 bp were extracted. Gene reassembly was carried out with approximately 100 ng of purified fragments in combination with 2 μ M of degenerated oligonucleotides (SI Table 1), to mutate positions D306, F353, N404 and W440 for library LibA, and D460, H463, T464, S512 for library LibB (DSR-S vardel Δ 4N numbers). The reaction mixture (30 μ L), containing 1 U Phusion® High-Fidelity DNA Polymerase (Finnzyme) in the appropriate buffer and 0.4 mM of each dNTPs, was thermocycled according to the following program: 1 denaturation step at 98°C for 30 s; 40 cycles composed of a denaturation step at 98°C for 10 s, 6 successive hybridization steps separated by 4°C each, from 65 to 41°C for 10 s each and an elongation step at 72°C for 20 s; and finally a 2 min step at 72°C. The fully recombined cassettes were isolated from the reassembly products in a last amplification by nested PCR using the primers for *_K7a3_nted* and *rev_K7a8_nted*. The purified nested PCR products were digested with *SpeI* and *AatII* restriction enzymes and ligated into the pBAD *Thio-dsrs* vardel Δ 4N-*His* plasmid to substitute the parental cassette. The ligation products were precipitated by adding 5 volume of absolute ethanol and the DNA pellet was rinsed 2 times with 70 % ethanol. The resuspended plasmids were transformed into *E. coli* TOP 10 electrocompetent cells and plated on LB agar supplemented with ampicillin (100 μ g/ml). The transformants were grown overnight at 37°C and the colonies were scraped from the plates for plasmid extraction constituting the DNA libraries.

Selection of glucansucrase active clones

The DNA libraries were transformed into chemo-competent *E. coli* BL21 AI cells and plated onto LB agar supplemented with ampicillin (100 μ g/ml). After overnight growth at 37°C, the cells were scraped, resuspended with physiological water and diluted to an OD_{600nm} of 5.10⁻⁵. The clones were subjected to selection pressure by plating them on 22x22 cm plates containing solid M9 mineral medium (42 mM Na₂HPO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 8.5 mM NaCl, 2.5.10⁻² mM CaCl₂, 1 mM MgSO₄)

supplemented with ampicillin (100 µg/ml), 0.02 % arabinose (wt/vol) and sucrose (50 g/L) as the sole carbon source. The plates were incubated 7 days at 20°C allowing sufficient expression of the recombinant genes encoding active glucansucrase and cell growth.

Screening for glucansucrase specificity

Active polymerase positive variants (visible when the colonies were recovered by a polymer bubble) were picked and transferred into 96 wells microplates (Nunc™ Brands Products, Roskilde, Denmark) filled with 250 µL LB per well supplemented with ampicillin (100 µg/ml) using a Biomek2000 pipettor (Beckman Coulter, Brea, CA). After overnight growth at 30°C under horizontal shaking at 250 rpm, 50 µL of each starter culture were used to inoculate 96-deepwell plates (ABgene, Epsom, UK) containing 500 µL auto-inducing media ZYM-5052 (Studier, 2005) supplemented with ampicillin (100 µg/ml) and 0.1 % arabinose (w/v). Growth and gene expression were conducted during 48 h at 20°C in an incubator-shaker (INFORS HT, Bottmingen, Switzerland). Plates were centrifuged (5 min, 3,700 g, 4°C) and the supernatants removed. Bacterial cell pellets were resuspended in 200 µL of lysozyme solution (0.5 mg/ml), incubated 20 min at 37°C and frozen at -80°C for 12 h. After thawing 1 h at room temperature, 800 µL of reaction mixture containing sucrose and dextranT1 (isomalto-oligosaccharides of 1 kDa in average) diluted in buffered deuterium water (NaAc 50 mM, CaCl₂ 0.05 g/L, pH 5.2) at a final concentration of 292 mM and 50 mM respectively, were added to each well. Enzymatic reactions were carried out during 48 h at 25°C under agitation at 700 rpm. Primary structures of the synthesized products were determined according to the NMR-based screening method previously described using the flow injection system (Bruker BioSpin GmbH) combining a Bruker Avance 600 MHz spectrometer with a Gilson Liquid Handler (Irague *et al.*, 2011).

Production and purification of wild-type enzyme and mutants with altered specificities

E. coli BL21 AI cells carrying the pBAD plasmids encoding the *Thio*-DSR-S *vardel* Δ 4N-*His* and selected tagged variants were grown 24 h at 20°C, in flask containing ZYM-5052 medium supplemented with ampicillin (100 µg/ml) and 0.1 % arabinose (w/v). Cells were centrifuged (4,500 g, 15 min, 4°C) and the pellets were resuspended to a final OD_{600nm} of 80 in 50 mM sodium acetate buffer pH 5.2 containing 0.05 g/L of CaCl₂. After sonication, the extracts were centrifuged (15,000 g, 30 to 60 min, 4°C) and the supernatant was harvested for subsequent purification. Variants were purified by affinity chromatography using Probond™ Nickel-Chelating resin (Invitrogen) according the previously described conditions for *Thio*-DSR-S *vardel* Δ 4N-*His* purification (Moulis *et al.*, 2006a). Protein concentrations were determined by spectrometry using a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

Electrophoretic analysis

Proteins from cell-free extracts or purified proteins were separated by electrophoresis using 3-8% Tris-acetate gels (Invitrogen) in denaturing conditions. After migration, the gel was stained with Colloidal Blue (total proteins coloration) or washed three times with 50 mM sodium acetate buffer pH 5.2 and incubated overnight in the same buffer supplemented with 100 g/L sucrose. The active bands were then detected by the formation of dextran inside the gel and coloration with Schiff reagent (zymogram) (SI Figure 1).

Western-blots were performed using the NuPAGE system, detection of the thioredoxin and 6xhis tags with appropriate primary antibodies and secondary antibody coupled with alkaline phosphatase (Invitrogen) (SI Figure 1).

Glucansucrase activity assays

One unit of glucansucrase activity corresponds to the amount of enzymes that catalyses the formation of 1 μmol of fructose per minute at 25°C in 50 mM sodium acetate buffer pH 5.2 containing 0.05 g/L of CaCl_2 and 100g/L sucrose. Concentration of reducing sugars was determined using the dinitrosalicylic acid method (Sumner and Howell, 1935) , using fructose as standard.

Kinetic studies were performed using different initial sucrose concentrations varying from 10 to 600 mM. The reactions were initiated by enzyme addition at a final concentration of 0.25 $\text{U}\cdot\text{mL}^{-1}$. At regular time intervals, samples were withdrawn and immediately heated at 95°C for 5 min to stop the reaction. The concentration of released fructose was determined by HPLC, on a Dionex system equipped with an Aminex HPX-87C Carbohydrate column (300 x 7.7 mm; Bio-rad, Hercules, CA, USA), using ultrapure water as eluent at 0.6 $\text{mL}\cdot\text{min}^{-1}$. The temperature of the column oven was set at 80°C. The initial velocities were directly calculated from fructose production rates and expressed in μmol of fructose produced per minute and per gram of enzyme. Curve fitting of the data and kinetic parameters determination were performed with the SigmaPlot "Enzyme kinetics" module (version 3.1).

α -glucan synthesis and purification

Cell-free extracts preparations of the parental *Thio*-DSR-S varde1 $\Delta 4\text{N-His}$ enzyme and selected tagged variants were incubated in sodium acetate buffer 50 mM, pH 5.2 supplemented with 0.05 g/L of CaCl_2 and 292 mM sucrose. Polymer syntheses were carried out at 25°C using 1 $\text{U}\cdot\text{mL}^{-1}$ of enzymes. Sucrose depletion was monitored by HPLC (see above). Reactions were stopped by 5 min incubation at 95°C. Polymers were recovered directly from the crude synthesis mediums by addition of one volume of pre-chilled ethanol at 4°C and centrifuged (20 min, 16,000 g, 4°C). The process of precipitation was repeated two times with 50 % (v/v) pre-chilled ethanol at 4°C to remove any trace impurities or free

reducing sugars. Finally, polymers were resuspended in ultrapure water, equivalent to the initial volume of reaction media, and freeze-dried.

HPLC analysis

Unpurified polymers analyses were performed by High-Performance Size Exclusion Chromatography (HPSEC) using two Shodex OH-Pack SB-805 and SB-802.5 columns in series. Products were eluted using a 0.45 M NaNO₃ and 1 % (v/v) ethylene glycol solution as eluent at a flow rate of 0.3 ml/min (Moulis *et al.*, 2006a). The column oven temperature was set at 70°C.

Glucose, fructose and leucrose concentrations were determined by HPLC using an Aminex HPX-87C Carbohydrate column as described above. The percentages of glucosyl moieties incorporated into free glucose (%G_{Glucose}) and leucrose (%G_{Leucrose}) were calculated as follows:

$$\%G_{Glucose} = \frac{[Glucose]_f}{[Sucrose]_0} \times 180/342 \times 100 \quad \text{and} \quad \%G_{Leucrose} = \frac{[Leucrose]_f}{[Sucrose]_0} \times 100 \quad (12)$$

where $[Sucrose]_0$ correspond to the initial substrate concentration (in mM), and $[Glucose]_f$ and $[Leucrose]_f$ to the final concentration of Glucose and Leucrose (in mM) at the end of the reaction.

The percentages of glucosyl moieties incorporated into HMW α -glucans were determined as follows:

$$\%G_{Glucan} = \frac{[Glucan]_f}{[sucrose]_0} \times 162/342 \times 100 \quad (13)$$

where $[Glucan]_f$ correspond to the concentration of glucan determined on HPSEC chromatogram at the end of the reaction and $[sucrose]_0$ to that of substrate at initial time. The proportion of glucose incorporated into Intermediate Molecular Weight (IMW) products (or oligosaccharides) was determined as follows:

$$\%G_{IMW} = 100 - \%G_{Glucan} - \%G_{Leucrose} - \%G_{Glucose} \quad (14)$$

NMR spectroscopy

Freeze-dried polymers samples (15 mg) were exchanged twice with 99.9 atom % D₂O, lyophilized and dissolved in 600 μ L of D₂O. 1D ¹H NMR spectra were recorded on a Bruker Avance 500 MHz spectrometer using a 5 mm z-gradient TBI probe at 298 K, an acquisition frequency of 500.13 MHz and a spectral width of 8012.82 Hz. The ¹H-signal from D₂O was used for automatic lock and a gradient shimming was performed on each sample. Before Fourier transformation, the FIDs were multiplied by an exponential function with a line broadening of 0.3 Hz. Spectra were processed with a 64 k zero filling, baseline correction and referenced using the TSP-d4 signal at 0 ppm. All NMR data were processed acquired and processed using TopSpin 2.1 software. The various signals were assigned as described by Seymour *et al.* (Seymour *et al.*, 1979), van Leeuwen *et al.* (Van Leeuwen *et al.*, 2009; Van Leeuwen

et al., 2008a; Van Leeuwen *et al.*, 2008b) and Maina *et al.* (Maina *et al.*, 2008). The percentages of $\alpha(1\rightarrow3)$ and $\alpha(1\rightarrow6)$ linkages in α -glucans were calculated from the relative intensities of the corresponding anomeric proton signals by integration of peak areas.

Three-dimensional molecular modelling

Sequence of DSR-S vardel Δ 4N catalytic domain (region N119 to K973, DSR-S vardel Δ 4N numbering) was submitted to the I-TASSER server, an online service for automated protein structure and function prediction (Roy *et al.*, 2010; Zhang, 2008, 2009) using as restraint the structure of *Lactobacillus reuteri* N-terminally truncated glucansucrase GTF180 (PDB accession code: 3KLLK). The 3D model of DSR-S vardel Δ 4N predicted by I-TASSER was then further refined using the CFF91 forcefield implemented within the DISCOVER module of InsightII software suite (Accelrys, San Diego, CA, USA). For the minimization, the CFF91 cross terms, a harmonic bond potential, and a dielectric of 1.0 were used. An initial minimization with a restraint on the protein backbone was performed using a steepest descent algorithm followed by conjugated gradient minimization steps until the maximum RMS was less than 0.5. In a subsequent step, the system was fully relaxed.

A 3D model of D460M/H463Y/T464M/S512C mutant was constructed by replacing *in silico* the parental enzyme with the corresponding amino acids observed in isolated improved mutant using the Biopolymer module of the InsightII software package (Accelrys). The conformation of the mutated residue side chains was optimized by manually selecting a low-energy conformation from a side-chain rotamer library. Steric clashes (van der Waals overlap) and non-bonded interaction energies (Coulombic and Lennard-Jones) were evaluated for the different side-chain conformations. The mutant was subsequently minimized following the same energy minimization procedure described for the parental enzyme.

Sucrose was manually docked in the active site of DSR-S vardel Δ 4N parental and mutant enzyme using the structure of *Lactobacillus reuteri* N-terminally truncated glucansucrase GTF180 in complex with sucrose (PDB code: 1HZ3). Complexes were then optimized using the same procedure above described.

Putative product of the enzyme, the α -D-Glcp-(1 \rightarrow 6)[α -D-Glcp-(1 \rightarrow 3)]- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp tetrasaccharide, was docked in enzyme mutant active site using the automated flexible docking program FlexX (Biosolveit) (Kramer *et al.*, 1999; Rarey *et al.*, 1996). Molecule was covalently docked in the active site using covalent constraint with catalytic D400. All parameters were set to the standard values as implemented in Version 3.1.1. The docking region was defined to encompass all protein amino acids for which at least one heavy atom was located within a 6.5 Å radius sphere, whose origin was located at

the centre of mass of catalytic D400. For each docking, the top 30 solutions corresponding to the best FlexX scores were retained.

All drawings were performed using PyMOL software (DeLano Scientific).

Acknowledgments

We cordially thank Joana Lima and Marta Leal for technical support.

This work was funded by the French National Institute for Agricultural Research (INRA) and the French National Center for Scientific Research (CNRS).

Supporting Information

Tables section

SI Table 1 : Primers and degenerated oligonucleotides used for combinatorial site-directed mutagenesis of *dsr-s* vardel $\Delta 4N$. Underlined nucleotides correspond to the introduced restriction sites.

Inverse PCR	
for_DSR-S_dAatII	5'-TCCGTTGATAGTAATGACGCTTAAACAAGTGTTGCTAT-3'
rev_DSR-S_dAatII	5'-ATAGCGAACACTTGTTAAGACGTCATTACTATCAACGGA-3'
Cassette amplification	
for_DSR-S_Alpha3	5'-CAAAGTACAGAGTGGTTAAAGGATGCAATG-3'
rev_DSR-S_Alpha8	5'-ATACCTTCAGTACGCGTCTCAGATGTTC-3'
Libraries construction	
LibA	
OligoD306X	5'-GAAGATATGAGCAATNNSCATCTGCAGAATGGCGCATT-3' <i>PstI</i>
OligoF353X	5'-AATCAAAGGTGGTNNSGAATTGTTGCTAGCCAATGAC-3' <i>NheI</i>
OligoN404X	5'-GTAGATGCAGTCGACNNSGTGGATGCAGATCTGTTACAAATT-3' <i>BglII</i>
OligoW440X	5'-CATCTTCAATTCTAGAAGATNNSAGTCACAATGATCCT-3' <i>XbaI</i>
LibB	
OligoXXX	5'-CAAGAAGCAATCAGTTAACCATGGATNNSSTATGTGNNNSNNSCAATTAATCTGGTCTCTA-3' <i>HpaI</i>
OligoS512X	5'-AGCTTTGTACGTGCGCACGACNNSGAAGTGCAAACG-3' <i>FspI</i>
Nested PCR	
for_K7a3_nted	5'-CCACAGTGAATGAAACTAGTGAAGATATG-3'
rev_K7a8_nted	5'-ATGGCATCTTACCATAGCGAACACTT-3'

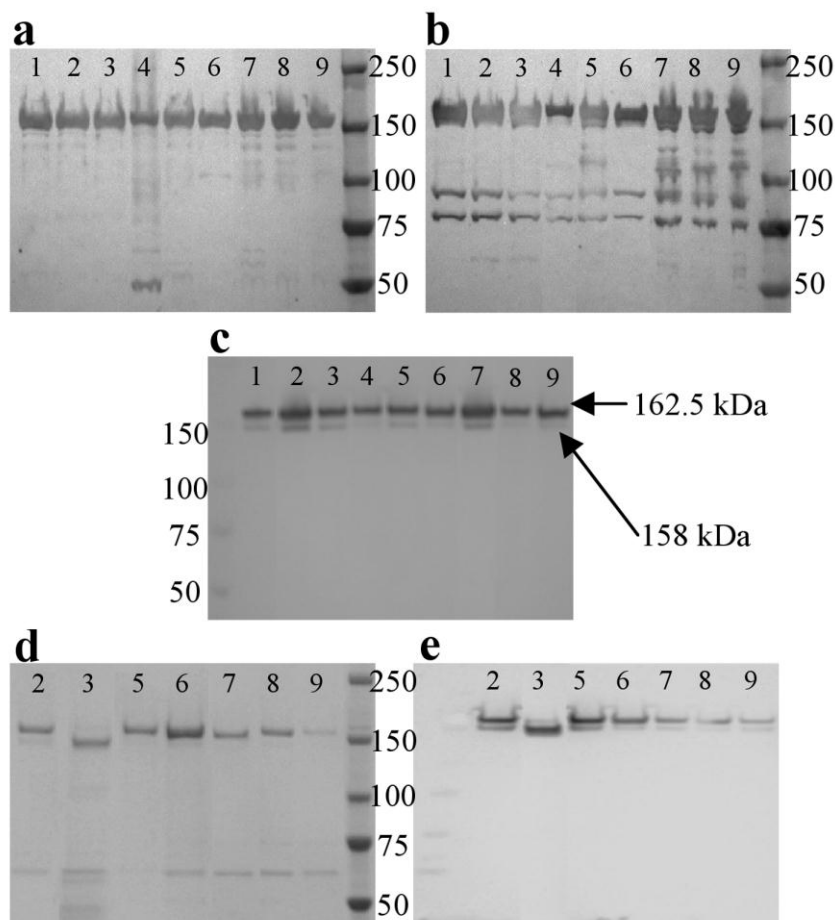
SI Table 2: Mutants from LibA and LibB with altered specificity, identified by NMR-based screening.

LibA					Chemical shift (%)	
					$\alpha(1\rightarrow3)$	$\alpha(1\rightarrow6)$
Parental enzyme	D306	F353	N404	W440	1.0	99.0
mutants		F353M			2.0	98.0
		F353W			3.0	97.0
	D306G	F353L	N405G	W440T	2.2	97.8

LibB					Chemical shift (%)	
					$\alpha(1\rightarrow3)$	$\alpha(1\rightarrow6)$
Parental enzyme	D460	H463	T464	S512	1.0	99.0
monomutants				S512V	5.5	94.5
				S512D	3.7	96.3
				S512C	2.0	98.0
				S512M	2.6	97.4
				S512I	5.0	95.0
				S512G	2.4	97.6
				S512E	3.3	96.7
				S512R	3.7	96.3
				S512T	3.4	96.6
		D460G			3.2	96.8
		D460E			2.0	98.0
		D460M			3.0	97.0
	doublemutants			T464I		2.5
		H463E	T464I		3.2	96.8
D460L			T464A		2.0	98.0
		H463R	T464D		2.3	97.7
		H463Q	T464W		2.7	97.3
		H463S	T464S		3.6	96.4
		H463Q	T464I		2.9	97.1
		H463S	T464V		4.2	95.8
	D460L	T464V		2.8	97.2	
	D460R	H463G		3.6	96.4	
Triplemutants	D460A	H463A	T464S		2.3	97.7
	D460V	H463M	T464E		2.5	97.5
	D460A	H463F	T464D		3.0	97.0
	D460E	H463S	T464G		2.4	97.6
	D460E	H463Q	T464F		3.1	96.9
	D460D	H463S	T464K		4.0	96.0
	D460G	H463Q	T464W		3.2	96.8
	D460V	H463Q	T464F		3.1	96.9
	D460S	H463Q	T464F		3.0	97.0
	D460A	H463Q	T464W		2.7	97.3
	D460A	H463A	T464A		2.5	97.5
	D460V	H463Q	T464C		2.6	97.4
	D460E	H463A	T464G		2.3	97.7
		H463Y	T464V	S512A	5.3	94.7
	D460L	H463Q	T464V		4.4	95.6
	D460M	H463Q	T464R		3.2	96.8
	D460V	H463Y	T464E		2.0	98.0

	D460I	H463E	T464Q		3.1	96.9
		H463G	T464N	S512C	3.4	96.6
	D460T	H463F	T464D		3.3	96.7
	D460M	H463P	T464C		4.2	95.8
	D460E	H463A	H464Q		3.6	96.4
	D460M	H463A	T464F		3.8	96.2
	D460A	H463L	T464E		4.6	95.4
	D460M	H463L	T464E		3.8	96.2
		H463R	T464D	S512T	6.0	94.0
	D460M	H463Y	T464A		2.1	97.9
	D460S	H463W	T464C		2.7	97.3
	D460Q	H463S	T464D		3.8	96.2
	D460H	H463M	T464V		3.8	96.2
	D460R	H463Y	T464E		4.8	95.2
	D460S	H463Q	T464K		3.7	96.3
	D460S	H463S	T464G		3.4	96.6
	D460E	H463F	T464S		4.6	95.4
	D460V	H463S	T464F		4.4	95.6
	D460E	H463F	T464S		4.2	95.8
		H463R	T464V	S512T	5.9	94.1
	D460V	H463V	T464K		2.6	97.4
		H463S	T464W	S512A	3.1	96.9
	D460G	H463Y	T464E		4.0	96.0
	D460G	H463F	T464M		4.6	95.4
	D460V	H463Y	T464S		3.4	96.6
	D460E		T464S	S512V	3.8	96.2
	D460A	H463A	T464V		3.2	96.8
	D460A	H463S	T464L		6.1	93.9
quadruple mutant	D460V	H463Y	T464E	S512C	3.8	96.2
	D460D	H463F	T464S	S512G	2.8	97.2
	D460A	H463Y	T464S	S512A	2.3	97.7
	D460V	H463Y	T464V	S512A	2.1	97.9
	D460V	H463W	T464Y	S512C	4.6	95.4
	D460L	H463L	T464L	S512V	2.7	97.3
	D460V	H463Q	T464C	S512A	3.4	96.6
	D460G	H463F	T464D	S512L	5.6	94.4
	D460G	H463F	T464G	S512C	4.2	95.8
	D460M	H463Y	T464M	S512C	8.3	91.7
	D460P	H463Y	T464G	S512V	6.2	93.8
	D460P	H463S	T464N	S512C	4.6	95.4

Figure section



SI Figure 1: (a) Western blot on unpurified enzymes using anti-thioredoxin antibody. (b) Western blot on unpurified enzymes using anti-6 x his. (c) Zymograms on unpurified enzymes. (d) SDS-PAGE gel and colloidal blue coloration on purified enzymes. (e) Western blot on purified enzymes using anti-6 x his antibody. 1, *Thio*-DSR-S vardel Δ 4N-*His*; 2, *Thio*-F353W-*His*; 3, *Thio*-F353T-*His*; 4, *Thio*-D306G/F353L/N404G/W440T-*His*; 5, *Thio*-S512C-*His*; 6, *Thio*-H463R/T464D/S512T-*His*; 7, *Thio*-D460M/H463Y/T464N/S512C-*His*; 8, *Thio*-H463R/T464V/S512T-*His* and 9, *Thio*-D460A/H463S/T464L-*His*.

Thio-DSR-S vardel Δ 4N-*His* and the selected *Thio*-LibA-*His*, *Thio*-LibB-*His* enzymes were produced in large scale format. A SDS-page with colloidal blue coloration of the cell-free extracts revealed the well expression of the soluble proteins (data not shown). A western blot using anti-thioredoxin antibody and anti-6 x his antibody were also performed (SI Figure 1a and 1b). Two bands at 162.5 and 50 kDa were detected for variant *Thio*-D306G/F353L/N404G/W440T-*His* on western blot using anti-6 x his antibody (SI Figure 1a) indicating that the protein was produced in intact (molecular weight of the tagged enzymes predicted at 161.5 kDa) and N-terminally degraded form. Western blot using anti-thioredoxin antibody revealed degradations at the C-Terminal extremity for all the proteins produced. Zymogram essay indicated that two forms of the proteins were actives (SI Figure 1c). One at 162.5 kDa corresponding to the entire tagged proteins and one at 158 kDa corresponding to a deletion of the

polyhistidine tag at the C-terminal part. The purification was achieved by affinity to a nickel chelating resin. The estimated yields were very weak, from 37 % for *Thio*-DSR-S varde1 Δ 4N-*His* to a total loss of protein for mutant *Thio*-D306G/F353L/N404G/W440T-*His*. The degraded active forms at 158 kDa were retained during the purification step and constituted the major form for variant *Thio*-F353T-*His* visible on zymogram (SI Figure 1e).

Due to the instability of the mutant *Thio*-D306G/F353L/N404G/W440T-*His* after purification, this latest was not retained for further characterization.

References

André, I., Potocki Véronèse, G., Morel, S., Monsan, P., Remaud-Siméon, M., 2010. Sucrose-utilizing transglucosidases for biocatalysis. *Rauter AP, Queneau Y, Vogel P eds.; In Topics in Current Chemistry: Carbohydrates in sustainable development.*

Argüello-Morales, M.A., Remaud-Simeon, M., Pizzut, S., Sarcabal, P., Willemot, R.M., Monsan, P., 2000. Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. *FEMS Microbiol. Lett.* **182**, 81-85.

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

Côté, G.L., Robyt, J.F., 1982. Isolation and partial characterization of an extracellular glucansucrase from *Leuconostoc mesenteroides* NRRL B-1355 that synthesizes an alternating (1-6), (1-3)- α -D-glucan. *Carbohydr. Res.* **101**, 57-74.

Fabre, E., Bozonnet, S., Arcache, A., Willemot, R.M., Vignon, M., Monsan, P., Remaud-Simeon, M., 2005. Role of the two catalytic domains of DSR-E dextranucrase and their involvement in the formation of highly α -1,2 branched dextran. *J. Bacteriol.* **187**, 296-303.

Funane, K., Ishii, T., Terasawa, K., Yamamoto, T., Kobayashi, M., 2004. Construction of chimeric glucansucrases for analyzing substrate-binding regions that affect the structure of glucan products. *Biosci. Biotechnol. Biochem.* **68**, 1912-1920.

Hellmuth, H., Wittrock, S., Kralj, S., Dijkhuizen, L., Hofer, B., Seibel, J., 2008. Engineering the glucansucrase GTFR enzyme reaction and glycosidic bond specificity: toward tailor-made polymer and oligosaccharide products. *Biochemistry* **47**, 6678-6684.

Herman, A., Tawfik, D.S., 2007. Incorporating synthetic oligonucleotides via gene reassembly (ISOR): a versatile tool for generating targeted libraries. *Protein Eng. Des. Sel.* **20**, 219-226.

Irague, R., Massou, S., Moulis, C., Saurel, O., Milon, A., Monsan, P., Remaud-Siméon, M., Portais, J.C., Potocki-Véronèse, G., 2011. NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity. *Anal. Chem.* **83**, 1202-1206.

Ito, K., Ito, S., Shimamura, T., Kawarasaki, Y., Abe, K., Misaka, T., Kobayashi, T., Iwata, S., 2010. Crystallization and preliminary X-ray analysis of a glucansucrase from the dental caries pathogen *Streptococcus mutans*. *Acta Crystallogr., Sect. F: Struct.* **66**, 1086-1088.

Ito, K., Ito, S., Shimamura, T., Weyand, S., Kawarasaki, Y., Misaka, T., Abe, K., Kobayashi, T., Cameron, A.D., Iwata, S., 2011. Crystal structure of glucansucrase from the dental caries pathogen *Streptococcus mutans*. *J. Mol. Biol.* **In Press**.

Kralj, S., Eeuwema, W., Eckhardt, T.H., Dijkhuizen, L., 2006. Role of asparagine 1134 in glucosidic bond and transglycosylation specificity of reuteransucrase from *Lactobacillus reuteri* 121. *Febs J.* **273**, 3735-3742.

Kralj, S., Stripling, E., Sanders, P., van Geel-Schutten, G.H., Dijkhuizen, L., 2005a. Highly hydrolytic reuteransucrase from probiotic *Lactobacillus reuteri* strain ATCC 55730. *Appl. Environ. Microbiol.* **71**, 3942-3950.

- Kralj, S., van Geel-Schutten, G.H., van der Maarel, M.J., Dijkhuizen, L.,** 2004. Biochemical and molecular characterization of *Lactobacillus reuteri* 121 reuteransucrase. *Microbiology* **150**, 2099-2112.
- Kralj, S., van Geel-Schutten, I.G., Faber, E.J., van der Maarel, M.J., Dijkhuizen, L.,** 2005b. Rational transformation of *Lactobacillus reuteri* 121 reuteransucrase into a dextransucrase. *Biochemistry* **44**, 9206-9216.
- Kralj, S., Van Leeuwen, S.S., Valk, V., Eeuwema, W., Kamerling, J.P., Dijkhuizen, L.,** 2008. Hybrid reuteransucrase enzymes reveal regions important for glucosidic linkage specificity and the transglucosylation/hydrolysis ratio. *FEBS J.* **275**, 6002.
- Kramer, B., Rarey, M., Lengauer, T.,** 1999. Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. *Proteins Struct. Funct. Bioinf.* **37**, 228.
- MacGregor, E.A., Jespersen, H.M., Svensson, B.,** 1996. A circularly permuted alpha-amylase-type alpha/beta-barrel structure in glucan-synthesizing glucosyltransferases. *FEBS Lett.* **378**, 263-266.
- Maina, N.H., Tenkanen, M., Maaheimo, H., Juvonen, R., Virkki, L.,** 2008. NMR spectroscopic analysis of exopolysaccharides produced by *Leuconostoc citreum* and *Weissella confusa*. *Carbohydr. Res.* **343**, 1446.
- Monchois, V., Vignon, M., Russell, R.R.,** 2000. Mutagenesis of asp-569 of glucosyltransferase I glucansucrase modulates glucan and oligosaccharide synthesis. *Appl. Environ. Microbiol.* **66**, 1923-1927.
- Moulis, C., Arcache, A., Escalier, P.C., Rinaudo, M., Monsan, P., Remaud-Simeon, M., Potocki-Veronese, G.,** 2006a. High-level production and purification of a fully active recombinant dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *FEMS Microbiol. Lett.* **261**, 203-210.
- Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M.,** 2006b. Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J. Biol. Chem.* **281**, 31254-31267.
- Pijning, T., Vujičić-Žagar, A., Kralj, S., Eeuwema, W., Dijkhuizen, L., Dijkstra, B.W.,** 2008. Biochemical and crystallographic characterization of a glucansucrase from *Lactobacillus reuteri* 180. *Biocatal. Biotransform.* **26**, 12-17.
- Rarey, M., Kramer, B., Lengauer, T., Klebe, G.,** 1996. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470.
- Roy, A., Kucukural, A., Zhang, Y.,** 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725-738.
- Sanger, F., Nicklen, S., Coulson, A.R.,** 1977. DNA sequencing with chain-terminating inhibitors. *PNAS* **74**, 5463-5467.
- Seymour, F.R., Knapp, R.D., Bishop, S.H.,** 1979. Correlation of the structure of dextrans to their ¹H-N.M.R. spectra. *Carbohydr. Res.* **74**, 77-92.

Shimamura, A., Nakano, Y.J., Mukasa, H., Kuramitsu, H.K., 1994. Identification of amino acids residues in *Streptococcus mutans* glucosyltransferases influencing the structure of the glucan product. *J. Bacteriol.* **176**, 4845-4850.

Studier, F.W., 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207-234.

Sumner, J.B., Howell, S.F., 1935. A method for determination of saccharase activity. *J. Biol. Chem.*, 51-54.

Van Der Veen, B.A., Potocki-Veronese, G., Albenne, C., Joucla, G., Monsan, P., Remaud-Simeon, M., 2004. Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett.* **560**, 91-97.

Van Hijum, S.A.F.T., Kralj, S., Ozimek, L.K., Dijkhuizen, L., van Geel-Schutten, I.G.H., 2006. Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 157-176.

Van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2009. Structural characterization of bioengineered α -D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180. *Biomacromolecules* **10**, 580-588.

Van Leeuwen, S.S., Kralj, S., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008a. Structural analysis of bioengineered α -D-glucan produced by a triple mutant of the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180: generation of (α 1 \rightarrow 4) linkages in a native (1 \rightarrow 3)(1 \rightarrow 6)- α -D-glucan. *Biomacromolecules* **9**, 2251.

Van Leeuwen, S.S., Kralj, S., Van Geel-Schutten, I.H., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2008b. Structural analysis of the α -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme. *Carbohydr. Res.* **343**, 1237.

Vujicic-Zagar, A., Pijning, T., Kralj, S., Lopez, C.A., Eeuwema, W., Dijkhuizen, L., Dijkstra, B.W., 2010. Crystal structure of a 117 kDa glucansucrase fragment provides insight into evolution and product specificity of GH70 enzymes. *PNAS*.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191.

Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D., Wishart, D.S., 2003. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* **31**, 3316-3319.

Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40.

Zhang, Y., 2009. I-TASSER: fully automated protein structure prediction in CASP8. *Proteins* **77 Suppl 9**, 100-113.

CHAPITRE 4 : Caractérisation physico-
chimique d' α -glucanes produits par des
mutants de dextrane-saccharase

Les travaux décrits dans les chapitres précédents ont été consacrés au développement d'une stratégie d'ingénierie combinatoire visant à mieux comprendre les relations structure/spécificité des glucane-saccharases de la famille GH70. Grâce à cette approche, nous avons isolé 82 variants de la dextrane-saccharase DSR-S vardel $\Delta 4N$ de spécificité altérée par rapport à l'enzyme parentale. Sept d'entre eux, capables de synthétiser efficacement des dextrans originaux présentant un taux croissant de liaisons $\alpha(1\rightarrow3)$, ont été caractérisés plus profondément. Le chapitre IV est quant à lui consacré à la caractérisation des propriétés physico-chimiques et à l'évaluation du potentiel applicatif de ces nouveaux polysaccharides, qui viennent élargir la gamme des biopolymères synthétisés à partir du saccharose.

En effet, aujourd'hui, ces derniers sont utilisés de façon très marginale dans l'industrie. Pourtant, ils présentent l'avantage d'être biodégradables et leur production par voie enzymatique ou microbienne est très simple. Les propriétés des α -glucanes produits par les enzymes de la famille GH70 méritent donc d'être étudiées de façon approfondie, notamment pour envisager des applications en tant que texturants peu calorique pour l'industrie agro-alimentaire et comme nouveaux matériaux bio-sourcés pour les éco-industries. D'un point de vue générique, il s'agit aussi d'identifier les liens existants entre structure et propriétés physico-chimiques de ces macromolécules pour concevoir à façon les structures optimales pour l'application envisagée.

Nous avons entrepris ce travail en collaboration avec le Dr. Alain Buléon, Agnès Rolland-Sabaté et le Dr. Jean Louis Doublier de l'unité Biopolymères, Interactions, Assemblages de l'INRA de Nantes. Bénéficiant ainsi d'un panel d'expertises incontournable pour appréhender ces questions, nous avons donc pu combiner des techniques qui n'avaient jamais été appliquées à la caractérisation des polymères de haute masse molaire produits par les glucane-saccharases. Le travail a porté en particulier sur la caractérisation structurale, de la mobilité moléculaire, des propriétés texturantes et filmogènes des α -glucanes synthétisés par l'enzyme parentale DSR-S vardel $\Delta 4N$ et ses 7 variants décrits dans le chapitre précédent.

Characterization of bio-engineered α -D-glucans synthesized by dextransucrase mutants

Romain Irague, Agnès Rolland-Sabaté, Laurence Roncalli-Tarquis, Jean Louis Doublier, Claire Moulis, Pierre Monsan, Magali Remaud-Siméon, Gabrielle Potocki-Véronèse, Alain Buléon

Abstract

Seven dextran types, displaying from 3 to 20 % amounts of $\alpha(1\rightarrow3)$ glucosidic linkages, were *in vitro* synthesized from sucrose by mutants of dextransucrase DSR-S from *Leuconostoc mesenteroides* NRRL B-512F. Structural and physico-chemical properties of these original biopolymers were characterized. Using asymmetrical flow field flow fractionation coupled with multiangle laser light scattering, weight average molar masses and radii of gyration were determined to be comprised between 0.76 to $5.95 \cdot 10^8$ g.mol⁻¹ and 55 to 210 nm, respectively. v_G structural parameter values indicated that dextrans Gcn6 and Gcn7, which contain 15 and 20 % $\alpha(1\rightarrow3)$ linkages, are highly branched and contain long ramifications, while Gcn1, is rather linear with only 3 % $\alpha(1\rightarrow3)$ linkages. Others display intermediary molecular structures. Rheological properties investigation showed that all these polymers present classical non-Newtonian pseudoplastic behaviour. However, Gcn_Dv Δ 4N, Gcn2, Gcn3 and Gcn7 form weak gel while others displayed viscoelastic behaviour that is typical of entangled polymer solutions. Finally, glass transition temperature T_g were measured by differential scanning calorimetry. Interestingly, Gcn1 and Gcn5 displayed T_g values of 19.0 and 29.8°C. According to these low T_g values, these two original dextrans were able to form rubber and flexible films at ambient temperature without any plasticizer addition. The mechanical parameters determined for Gcn1 films from tensile tests are very promising regarding the biofilms obtained by other polysaccharides extracted from plants, algae or microbial fermentation. These data open perspectives for using these dextrans as innovative bio-sourced materials.

Introduction

Biopolymers such as polysaccharides are widely used as thickening, stabilizing or gelling agents. They are generally extracted from plants (e.g. starch, pectin, cellulose) or algae (e.g. agar, carrageenan) but some of them are also produced by microbial fermentations (xanthan, gellan, β -glucans, bacterial cellulose, pullulan, for instance). Polysaccharides are also of interest for the production of biodegradable films which could be used for packaging (Lichtenthaler and Peters, 2004). This recent field of application answers to environmental problems related to the partial or total persistence of synthetic films such as polyethylene terephthalate (PET), polyvinylchloride (PVC), polyethylene (PE) or polypropylene (PP) (Siracusa *et al.*, 2008). In particular, plant and algae polysaccharide ability to form films has been explored these last years (Fabra *et al.*, 2008; Garcia *et al.*, 2006; Lafargue *et al.*, 2007; Müller *et al.*, 2009). Microbial exopolysaccharides such as pullulan, kefiran or gellan were also considered for this application (Alves *et al.*, 2011; Kristo *et al.*, 2007; Piermaria *et al.*, 2009; Yang and Paulson, 2000).

In this context, α -glucans synthesized from sucrose by the glucansucrases (GS) from lactic acid bacteria of the genus *Leuconostoc*, *Streptococcus*, *Lactobacillus*, *Exigobacterium* and *Weissella* could offer an alternative source of biopolymers (André *et al.*, 2010; Van Hijum *et al.*, 2006). Glucansucrases (GS) are transglucosidases classified into the family 70 of the Glycoside-Hydrolases (Cantarel *et al.*, 2009), that are able to polymerize the D-glucosyl units of sucrose to form high molecular weight homopolysaccharides. Depending on the reaction conditions, these versatile biocatalysts are able to synthesize a wide range of α -glucans and oligosaccharides varying in terms of size and osidic linkage types. As they use sucrose, a renewable resource as substrate, these enzymes are very suitable tools for glycodiversification (André *et al.*, 2010). However, the polymers they synthesize are still marginally exploited and represent a small fraction of the biopolymers currently commercialized. To date, only dextrans synthesized by the glucansucrase DSR-S from *Leuconostoc mesenteroides* NRRL B-512F are commercialized, mainly for analytical and therapeutic applications. The native polymer directly produced by the bacterial strain is of very high molecular weight (up to 10^6 Da) and composed at 95% and 5 % of $\alpha(1\rightarrow6)$ and $\alpha(1\rightarrow3)$ linkages, respectively. It is considered as a novel food ingredient by the European Union since 2001. Indeed, it was shown that incorporation of 5% of this dextran in bakery products improved softness, crumb texture and loaf volume (Vandamme *et al.*, 2003). However, diversification of applications of the large panel of α -glucans that are produced by GS is nowadays limited by the lack of data related to their physico-chemical properties.

In this paper, we describe the investigation of the structural and physico-chemical properties of seven high molecular weight α -glucans with original structures, to evaluate their applicative potential as bio-

sourced materials. These dextrans, that contain between 3 and 20 % $\alpha(1\rightarrow3)$ linkages, were synthesized from sucrose by seven mutants of the DSR-S enzyme, that were obtained by combinatorial engineering to modify the linkage specificity of the parental enzyme (Irague *et al.*, 2011a; Irague *et al.*, 2011b). To this end, asymmetrical flow field flow fractionation coupled with multi angle laser light scattering (AFFFF-MALLS) was used in order to determine the molar mass of these new polymers, and to characterize their conformations in solution. Rheological behaviour and glass transition properties were also investigated, as well as their ability to form innovative films.

Materials and methods

Production of parental and mutant enzymes

E. coli BL21 AI cells carrying a pBAD plasmid encoding the *Thio*-DSR-S *vardel* $\Delta 4N$ -*His* or the tagged variants were grown 24 h at 20°C in flask containing ZYM-5052 medium supplemented with ampicillin (100 $\mu\text{g}\cdot\text{mL}^{-1}$) and 0.1 % (w/v) arabinose. Cells were centrifuged (4,500 g, 15 min, 4°C) and the pellets were resuspended to a final $\text{OD}_{600\text{nm}}$ of 80 in 50 mM sodium acetate buffer pH 5.2 containing 0.05 $\text{g}\cdot\text{L}^{-1}$ of CaCl_2 , before being sonicated and centrifuged (15,000 g, 30 to 60 min, 4°C). Supernatants (cell free extracts) were harvested for polymer production.

Glucansucrase activity assays

One unit of GS activity corresponds to the amount of enzymes that catalyses the formation of 1 μmol of fructose per minute at 25°C in 50 mM sodium acetate buffer pH 5.2 containing 0.05 $\text{g}\cdot\text{L}^{-1}$ CaCl_2 and 292 mM sucrose. The concentration of reducing sugars was determined using the dinitrosalicylic acid method (Sumner and Howell, 1935), using fructose as standard.

Glucan synthesis

Cell-free extracts preparations of the parental *Thio*-DSR-S *vardel* $\Delta 4N$ -*His* and variant enzymes were incubated in sodium acetate buffer 50 mM, pH 5.2 supplemented with 0.05 $\text{g}\cdot\text{L}^{-1}$ of CaCl_2 and 292 mM sucrose, in a final volume of 20 mL. Polymer syntheses were carried out at 25°C with 1 $\text{U}\cdot\text{mL}^{-1}$ of enzymes until total sucrose depletion (8h of reaction).

α -Glucan purification

After total sucrose depletion, α -glucans were recovered directly from the crude synthesis media by addition of one volume of pre-chilled ethanol at 4°C and centrifuged (20 min, 16,000 g, 4°C). The

precipitation procedure was repeated twice with 50 % (v/v) pre-chilled ethanol at 4°C to remove mono- and oligosaccharides. Finally, polymers were resuspended in pure water (40 mL) and freeze-dried.

NMR spectroscopy

Freeze-dried polymers samples (15 mg) were exchanged twice with 99.9 atom % D₂O, lyophilized and dissolved in 600 µL of D₂O. 1D ¹H NMR spectra were recorded on a Bruker Avance 500 MHz spectrometer using a 5 mm z-gradient TBI probe at 298 K, an acquisition frequency of 500.13 MHz and a spectral width of 8012.82 Hz. The ¹H-signal from D₂O was used for automatic lock and a gradient shimming was performed on each sample. Before Fourier transformation, the FIDs were multiplied by an exponential function with a line broadening of 0.3 Hz. Spectra were processed with a 64 k zero filling, baseline correction and referenced using the TSP-d4 signal at 0 ppm. All NMR data were acquired and processed using TopSpin 2.1 software. The various signals were assigned as described by Seymour *et al.* (Seymour *et al.*, 1979), van Leeuwen *et al.* (Van Leeuwen *et al.*, 2009; Van Leeuwen *et al.*, 2008a; Van Leeuwen *et al.*, 2008b) and Maina *et al.* (Maina *et al.*, 2008). The percentages of α(1→3) and α(1→6) linkages in α-glucans were calculated from the relative intensities of the corresponding anomeric proton signals by integration of peak areas.

Asymmetrical Flow Field Flow Fractionation coupled with Multi Angle Laser Light Scattering (AFFFF-MALLS) analyses

Dextran size and structure were analysed by AFFFF-MALLS, by using synthesis reaction media. Samples were diluted 250 times using Millipore water containing 0.2 g.L⁻¹ sodium azide and filtered through 0.45 µm Durapore™ membranes (Millipore, Bedford, MA). The solubilisation recoveries (Sr) were calculated as follow:

$$Sr (\%) = 100 \times \frac{C_{cf}}{C_{cm}} \quad (15)$$

Where C_{cf} and C_{cm} correspond to the carbohydrate concentrations after and before filtration respectively, determined using the sulphuric acid-orcinol colorimetric method (Planchot *et al.*, 1997).

The equipment used for AFFFF-MALLS analyses was previously described by Rolland-Sabaté *et al.* (Rolland-Sabaté *et al.*, 2007). The two on-line detectors comprised a MALLS instrument (Dawn® HELEOS™) fitted with a K5 flow cell and a He-Ne laser ($\lambda = 658$ nm) (Wyatt Technology, Santa Barbara, CA), and a ERC-7515A refractometer (Erma, Tokyo, Japan). The carrier (Millipore water containing 0.2 g.L⁻¹ sodium azide) used was degassed and filtered through a Durapore (0.1 µm) membrane (Millipore).

As the studied α -glucans were particularly large, a flow method using very low cross flow during elution was carried on. Moreover, the injected masses were adjusted (6 to 8 μg of polymer were injected) in order to avoid overloading of the AFFF channel and to limit aggregation phenomena during the focusing step.

The crossflow was initially set at 1 $\text{mL}\cdot\text{min}^{-1}$ and the channel flow rate at 0.2 $\text{mL}\cdot\text{min}^{-1}$ for the sample introduction and relaxation/focusing period. The samples were injected at 0.2 $\text{mL}\cdot\text{min}^{-1}$ for 300 s. After injection pump was stopped, the samples were allowed to relax and focus for 60 s. For elution, the flow rate was set at 0.84 $\text{mL}\cdot\text{min}^{-1}$ and the crossflow was maintained at 0.1 $\text{mL}\cdot\text{min}^{-1}$ for 2 274 s. The elution recoveries (E_r) were estimated as follow:

$$\text{Estimated } E_r (\%) = \frac{C_G}{C_{\text{sucrose}} \times \%G \times \%G_{\text{Glucan}} \times d \times \%Sr} \quad (16)$$

Where C_G correspond to the concentration of eluted glucans [determined by integration of the differential refractive index (DRI) signal], C_{sucrose} the concentration of sucrose used for the glucan synthesis, $\%G$ the percentage of total glucose units in the reaction mediums, $\%G_{\text{Glucan}}$ the percentage of glucose units incorporated into glucans [determined by HPSEC in previous work (Irague *et al.*, 2011b)], d the dilution factor of the sample.

The weight average molar mass \bar{M}_w ($\text{g}\cdot\text{mol}^{-1}$) and the z-average radius of gyration \bar{R}_G (nm) were established using ASTRA[®] software from Wyatt Technology (version 5.3.2.13 for PC) as previously described (Rolland-Sabaté *et al.*, 2007). A value of 0.145 $\text{mL}\cdot\text{g}^{-1}$ was used as the refractive index increment (dn/dc) for glucans. M_i and R_{Gi} (molar mass and radius of gyration of the i th slice, respectively) were obtained at each slice of the chromatogram peak using the Berry extrapolation (with a first order polynomial fit) of the light scattered to zero angle:

$$\sqrt{\left(\frac{Kc}{R_\theta}\right)_i} = \sqrt{\frac{1}{M_i} \left(1 + \frac{16\pi^2 n^2}{3\lambda^2} R_{Gi}^2 \sin^2(\theta/2)\right)} \quad (17)$$

where c is the concentration, K the optical constant, R_θ the excess Rayleigh ratio of the solute, λ the wavelength of the incident laser beam and θ the angle of observation. The normalization of photodiodes was achieved using a low molar mass P20 pullulan standard.

Rheological measurements

Aqueous solutions of 5 % (w/v) α -glucans were prepared by dissolving freeze-dried polysaccharides with distilled water and gently stirring at 25 °C overnight. Rheological properties were analysed using a strain controlled rheometer (RFS II, Rheometrics Inc., Piscataway, NJ) with cone-plate geometry (5 cm diameter; 0.05 rad cone angle). Flow and dynamic measurements were performed at 20°C.

Flow curves were determined for shear rates ranging from 10^{-2} to 10^2 s $^{-1}$ and thixotropic loops from 0 s $^{-1}$ to 10^2 s $^{-1}$ and back to 0. Viscoelastic measurements (G' and G'' as a function of frequency) were performed in dynamic conditions (oscillatory shear) with frequency ranging from 10^{-2} to 10^2 rad.s $^{-1}$ and a shear strain amplitude of 5 %.

Determination of glass transition temperature T_g

Freeze-dried glucans were equilibrated at 57 % relative humidity (RH) using a saturated sodium bromide solution for one week at ambient temperature under vacuum. Glass transition temperatures (T_g) were determined with a differential scanning calorimeter Q100 system (TA instrument, France). The instrument was calibrated with indium. The measurements were made with 2 – 30 mg sample mass, using hermetically sealed aluminum pans (TA instruments, Guyancourt, France), heated from 0 to 120°C at a rate of 3°C.min $^{-1}$. Two heating scans separated by a cooling stage at about 10°C.min $^{-1}$ were used in order to prevent any signature resulting from previous physical ageing. An empty pan was used as a reference. Glass transition temperature was taken as the inflection point of the heat capacity change (Bizot *et al.*, 1997). Water content was determined after DSC measurement by thermogravimetric analysis (TGA).

Thermogravimetric analysis (TGA)

The water content was determined using a TGA2050 (T.A. Instruments, New Castle, DE). It was taken as the mass difference when heating samples up to 130°C at a rate of 10°C.min $^{-1}$ and keeping them for 40 min at the final temperature.

Film preparation and tensile tests

Aqueous solutions of 5 % (w/v) α -glucans were prepared as described for rheological measurements. Films were obtained by casting 1 mL of the aqueous solutions on Teflon plate at 60°C during 3 h. The films were equilibrated at 57 % relative humidity (RH) using a saturated sodium bromide solution for 3 days at ambient temperature under vacuum. For tensile tests, 5A type dumbbell-shape specimens according to the ISO-527-2 standard were cut from the films. Film thicknesses were measured with a hand-held micrometer. Tensile tests were performed with a TST350 instrument (Linkman scientific instruments, Surrey, UK). Elongation (ϵ) and tensile stress (σ) were measured using the Linksys32 software. The distance between the supports was 20 mm and the crosshead speed was fixed at 50 μ m.s $^{-1}$. Measurements were performed twice, at 25°C, until sample failure.

Results and discussion

Dextranucrase DSR-S vardel $\Delta 4N$ (Moulis *et al.*, 2006b), a truncated mutant of the enzyme DSR-S from *Leuconostoc mesenteroides* NRRL B-512F, is the most efficient glucanucrase known to date. Recently, DSR-S vardel $\Delta 4N$ combinatorial engineering was performed to investigate the structure-specificity relationships of this enzyme, and to generate a toolbox of novel biocatalysts for the synthesis of novel α -glucan structure (Irague *et al.*, 2011a; Irague *et al.*, 2011b). A library of 82 mutants displaying altered specificity compared to the parental enzyme was obtained. Among them, seven (named Gcn1 to Gcn7) were selected for their ability to efficiently produce dextrans with various amounts of $\alpha(1\rightarrow 3)$ linkages, ranging from 3 to 20 % (Figure 41A and 41B). Their structural and physico-chemical properties were investigated by using a various panel of techniques.

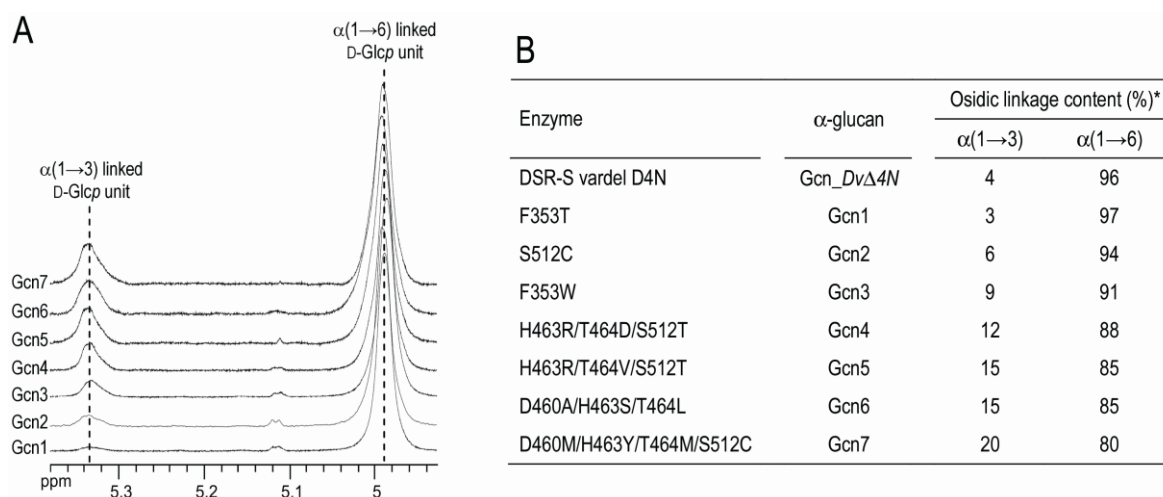


Figure 41: (A) ^1H NMR spectra of the α -glucans synthesized by the seven selected variants. (B) Osidic linkage content of the α -glucans produced by the parental enzyme DSR-S vardel $\Delta 4N$ and the seven selected mutants.

α -glucan size and conformation in solution

After filtration of the seven α -glucans in their synthesis reaction media, the solubilisation recovery rates (Table 14) were ranging from 87.1 to 100 %, excepted for Gcn1 for which the recovery was slightly lower (77.3 %). The elution recovery, which represents the percentage of macromolecules eluted through the AFFFF channel, ranged from 77.3 to 100 %. These recovery values indicate that the majority of each sample could be analyzed by the procedure used. The lower elution recovery values obtained for Gcn_Dv $\Delta 4N$, Gcn4 and Gcn5 may be due to the very high sizes and molar masses of the molecules and to aggregation phenomena. Consequently, the determined \bar{M}_w and \bar{R}_G values that were determined for these polymers do not represent the total α -glucan population. The light scattering at

90° (LS90) and the DRI signals obtained at different elution volumes are shown in Figure 42A and B. They indicate that the various glucans show different size distributions, which were also evidenced by HPSEC measurement in a previous work (Irague *et al.*, 2011b). Samples Gcn_DvΔ4N, Gcn1, Gcn2, Gcn3, Gcn6 and Gcn7 eluted in one LS90 single peak (Figures 42A, 42C, 42D and 42E) and one DRI peak (Figure 42B). LS90 peak tailing was observed which could be due to the presence of a second population of higher size and molar mass (HMM). However, the asymmetry was not observed on DRI responses indicating that the proportion of the putative second population was negligible (Figure 42B). A more pronounced asymmetric shape of the LS90 peaks was observed for samples Gcn4 and Gcn5 (Figure 42D and 42E), which was also revealed by DRI responses at elution volume around 25 mL (Figure 42B and 42F). Data treatment for Gcn4 and Gcn5 was thus carried by considering that two populations of α -glucans of low and high size and molar mass (LMM and HMM) were present.

Table 14: AFFFF-MALLS analysis of α -glucans produced by the parental enzyme DSR-S vardeI Δ4N and its mutants. (\bar{M}_w) weight average Molar Masses, (\bar{R}_G) z-average Radii of Gyration, apparent density (d_{Gapp}) and the slope of the log-log plot of the radius of gyration versus the molar mass (v_G).

Sample	Solubilization recovery (%)	Estimated elution recovery (%)	\bar{M}_w ($\times 10^6$ g.mol ⁻¹)	\bar{R}_G (nm)	v_G	d_{Gapp} (g.mol ⁻¹ .nm ⁻³)
Gcn_DvΔ4N	87.1	71.4	4.13	157.5	0.52	29.3
Gcn1	77.3	94.1	5.95	210.0	0.50	21.1
Gcn2	93.7	n.a.	3.40	133.7	0.60	45.8
Gcn3	97.1	91.4	2.37	121.0	0.53	35.2
Gcn4	97.2	80.6	1.19 (8.42) ^a	78.5 (183.0) ^a	0.53 (0.33) ^a	75.9 (53.7) ^a
Gcn5	99.1	83.2	1.79 (9.22) ^b	88.0 (209.0) ^b	0.52 (0.31) ^b	71.4 (35.0) ^b
Gcn6	96.8	100.0	1.16	67.0	0.43	102.9
Gcn7	100.0	100.0	0.76	55.0	0.47	119.9

^a These values were determined from the second populations at elution volume between 24 and 34 mL (Figure 42D).

^b These values were determined from the second populations at elution volume between 25 and 35 mL (Figure 42E).

n.a.: not available.

As seen in Figure 42, the distribution of molar masses (M_i) versus the elution volume reveals a change of slopes for Gcn6 (Figure 42C), Gcn1 (Figure 42D) and Gcn5 and Gcn7 (Figure 42E), which corresponds to the peak asymmetry and could be due to the presence of the two suspected populations of α -glucans. Regarding samples Gcn4, two changes of M_i slopes are observed that could reflect the presence of three different populations partially overlapped: LMM, HMM and very low molar mass (VLMM) population (Figures 42C and 42F).

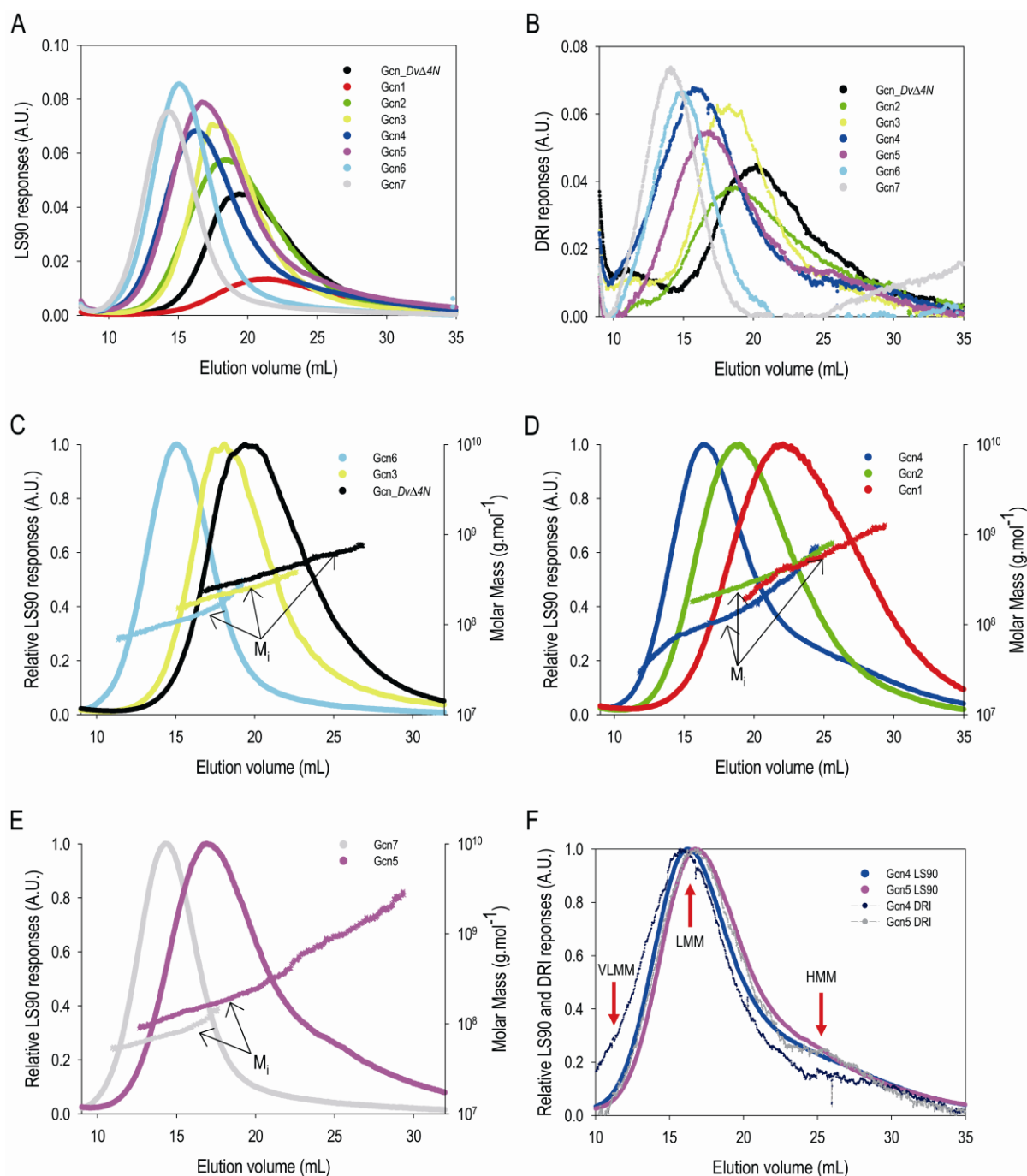


Figure 42. (A) and (B) Elugrams (Light Scattering at 90° and DRI responses) of α -glucans versus the elution volume. (C), (D) and (E) Elugrams (Relative Light Scattering at 90° responses) of α -glucans and their molar masses versus the elution volume. (F) Elugrams (Relative Light Scattering at 90° and DRI responses) of Gcn4 and Gcn5 versus the elution volume. The red arrows indicate the VLMM (very low molecular mass), LMM (low molecular mass) and HMM (high molecular mass) observed for Gcn4.

\bar{M}_w and \bar{R}_G values are reported in Table 1. These parameters were determined from the whole elution peak of each sample except for Gcn4 and Gcn5 for which a second population of HMM was considered. Very high \bar{M}_w values were determined for all glucans, the highest values were found at 4.13 and 5.95

$\times 10^8 \text{ g.mol}^{-1}$ for Gcn_Dv Δ 4N and Gcn1, respectively. Gcn4, Gcn5, Gcn6 and Gcn7 displayed the smallest \bar{M}_w values ranging from 0.76 to $1.79 \times 10^8 \text{ g.mol}^{-1}$. Intermediate values were found for the other glucans. The \bar{M}_w values of Gcn4 and Gcn5 high molar mass populations (which represent 32.4 % and 37.1% of the total recovered mass for Gcn4 and Gcn5, respectively) attain 8.42 and $9.22 \times 10^8 \text{ g.mol}^{-1}$ which is extremely high. The \bar{R}_G values followed the same trend as \bar{M}_w ones with the highest values determined at 210 nm for Gcn1 and the smallest one at 55 nm for Gcn7. The \bar{R}_G values of Gcn4 and Gcn5 HMM populations increased compared to that of the LMM population. However, the increase is less pronounced than that observed for \bar{M}_w values, suggesting that the HMM population is more compact. For each sample, the logarithm of R_{Gi} were plotted against the logarithm of the corresponding M_i (Figure 43) to determine ν_G values from equation (18).

$$R_{Gi} = K_G M_i^{\nu_G} \quad (18)$$

The theoretical ν_G values calculated for a spherical polymer, a linear random coil or a rod conformation are of 0.33, 0.5 - 0.6 and 1, respectively. The ν_G values obtained for Gcn_Dv Δ 4N, Gcn1, Gcn2, and Gcn3 (Table 14), but also for the LMM population of Gcn4 (Gcn4_LMM) and Gcn5 (Gcn5_LMM) were between 0.5 and 0.6 that are representative of linear or quasi-linear random coil conformation.

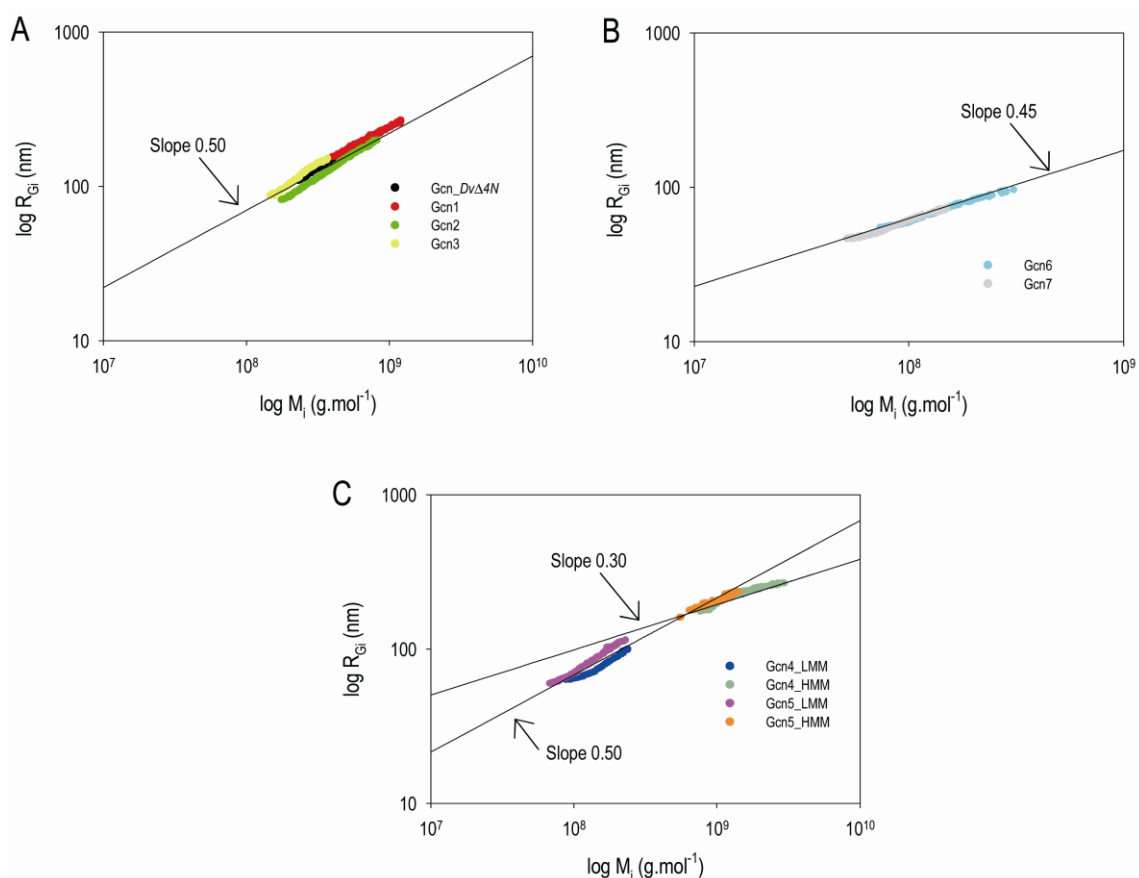


Figure 43. Logarithm of the radii of gyration (R_{Gi}) versus the logarithm of the molar masses (M_i) of the various α -glucans.

Gcn1 to 3 have indeed a slight proportion of α (1 \rightarrow 3) linkages (less than 10 %). This is in agreement with the v_G value of 0.56 recently reported for a linear (1 \rightarrow 6)- α -D-glucan in 0.1 M aqueous NaCl at 25°C (Cui *et al.*, 2008). Moreover, the quasi-linear random coil conformation could be produced either by short chains connected to a linear backbone or by a few very long chains connected to the backbone. Regarding Gcn4_LMM and Gcn5_LMM, the v_G values indicate that the ramifications are probably formed mainly by short chains. The v_G values of glucans Gcn6 and Gcn7 are around 0.45 (Table 14). Interestingly, these glucans possess the highest amount of α (1 \rightarrow 3) osidic linkages (Figure 41B). These v_G values probably reflect a more compact conformation formed by more branched structures containing long ramifications. Finally, the v_G values of the HMM populations in samples Gcn4 (Gcn4_HMM) and Gcn5 (Gcn5_HMM) were evaluated to be around 0.3 (Table 14), indicating very compact and spherical structures, probably some aggregates. The hypothesis of the presence of aggregates in these two samples is in agreement with the lower elution recoveries that were obtained (Table 14).

The determination of apparent particle density (d_{Gapp}) could be another way to investigate the molecule conformation. Apparent particle density was calculated on the basis of a smeared uniform density in the particle and based on the following equation for equivalent homogeneous spheres:

$$d_{Gapp} = M_w / (4\pi/3) R_G^3 \quad (19)$$

Gcn_Dv Δ 4N, Gcn1, Gcn2 and Gcn3 exhibit the smallest densities (21 to 46 g.mol⁻¹.nm⁻³, Table 14), which is in accordance with the quasi-linear conformation deduced from their v_G values. Moreover, comparing these densities at the same molar mass (as the density increases slightly when the molar mass decreases) it could be concluded that Gcn2 is denser than Gcn3. Gcn2 may then contain a higher proportion of long chain branches compared to Gcn3. The highest densities were obtained for both Gcn6 and Gcn7 (higher than 100 g.mol⁻¹.nm⁻³), thus confirming the more compact structure of these two glucans. Gcn4_LMM and Gcn5_LMM show intermediate and similar density values (around 70 g.mol⁻¹.nm⁻³). Their branching pattern is thus probably different from the other glucans.

Flow behaviour

All solutions of glucans (5%, w/v) displayed non-Newtonian shear-thinning behaviours as the apparent viscosity of the polymer solution decreased with the increase of shear rate (Figure 44A). Moreover, the superimposition of the up and down curves demonstrated a non-thixotropic behaviour. Gcn5 exhibited two distinct regions (1) a Newtonian plateau region at low shear rates and (2) a shear-thinning region at higher shear rates. This is classical behavior of macromolecular solutions whose the properties arise from an entanglement network. At low shear rate, the formation of new entanglements counterbalances

the disruption of existing entanglements by shear deformation and so, the apparent viscosity remains constant. At high shear rate, the ratio of entanglements formation/disruption decreases, leading to a decrease of apparent viscosity (Lee *et al.*, 2005). The shear-thinning behaviour was also demonstrated by the variations of shear stress with shear rate (data not shown). Gcn4 and Gcn6 displayed similar flow properties but the low shear rate Newtonian zone was not reached and would appear at much lower shear rates. The other samples (Gcn_Dv Δ 4N, Gcn2 and Gcn3) displayed a reverse tendency at low shear rate, with a yield stress. Gcn_Dv Δ 4N, Gcn2, Gcn3 showed the highest apparent viscosity at low shear rate, with 2 to 3 times higher orders of magnitude compared to Gcn5.

Viscoelastic behaviour

The mechanical spectra of Gcn_Dv Δ 4N, Gcn2 and Gcn3 solutions showed a strong elastic behaviour as the storage (elastic) moduli G' were superior to the loss (viscous) moduli G'' over the range of studied frequencies (Figure 44B). The slight dependency of the G' on the frequency and the relative large value of $\tan \delta$ ($\tan \delta = G''/G' > 0.3$) is typical of so-called weak gels (Tokita *et al.*, 2009). In a previous work, dextran synthesized by a crude extract of DSR-S vardel Δ 4N did not show such a gel-like behaviour (Moulis *et al.*, 2006a). However, compared to this work, the Gcn_Dv Δ 4N sample used here was purified by ethanol precipitation and freeze-dried before rheological analysis. This treatment may affect the viscoelastic properties of the polymer, as already observed by Sabatié *et al.* (Sabatié *et al.*, 1986). Regarding Gcn7, G' moduli and G'' moduli values are significantly lower than those of Gcn_Dv Δ 4N, Gcn2 and Gcn3, suggesting the presence of a more tenuous interaction network. The mechanical spectra of Gcn1, Gcn4 and Gcn5 (Figure 44C) were typical of macromolecular entangled polymer solutions. Both parameters G' and G'' are strongly frequency-dependent and increase with frequency. At low frequencies, the viscous modulus G'' dominates the elastic modulus G' (liquid-like behaviour). As the frequency increases a crossover point is reached ($\sim 1 \text{ rad}\cdot\text{s}^{-1}$ for Gcn1 and Gcn4) and then G' predominates at high frequencies (solid-like behaviour). The crossover point was not completely reached for Gcn5 solution within the tested frequency range of 10^{-2} to $10^2 \text{ rad}\cdot\text{s}^{-1}$, but it presumably occurs at higher frequencies. The behaviour Gcn6 remains similar to macromolecular entangled polymer solutions at high frequencies, but differ slightly at low frequencies (Figure 44D). This finding confirms the unusual behaviour of Gcn6 in solution, for which G' tends to level off at low frequencies even if no crossover was reached in the tested frequency range. This behaviour was previously observed for hydrolyzed oat β -glucans (Doublier and Wood, 1995) and was attributed to possible polymer self-aggregation, due to successive $\beta(1\rightarrow4)$ linked glucosyl units. These results support the

presence of aggregates in Gcn7 solution for which inflexion of the \bar{M}_w were observed (Figure 42B and 42D).

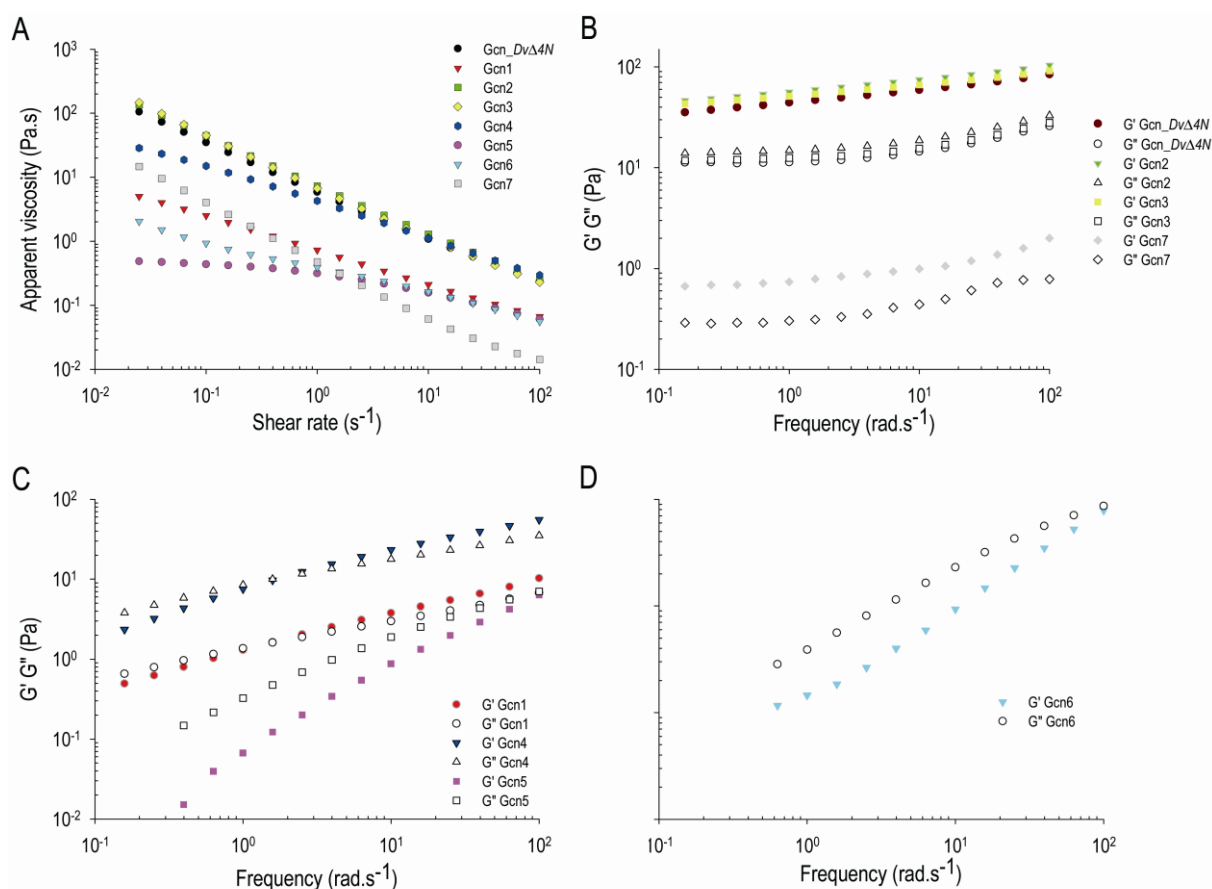


Figure 44: (A) Apparent viscosity as a function of the shear rate of the 5% (w/v) polymer solutions. (B), (C) and (D) Storage modulus (G') and loss modulus (G'') as a function of angular frequency of the 5% (w/v) polymer solutions.

Moisture contents and glass transition

Water content of the different glucans is shown in Table 15. Gcn_Dv Δ 4N, Gcn2, Gcn3, Gcn4, Gcn5 and Gcn6 displayed water content in a range of 11 – 13 % when equilibrated at a_w 0.57. These values are classically found when commercial dextrans, pullulans, amylopectins or starches are equilibrated at the same relative humidity (Bizot *et al.*, 1997; Borde *et al.*, 2002; Diab *et al.*, 2001; Icoz *et al.*, 2005). Surprisingly, the water content determined for Gcn1 was found to be \sim twice lower than for the other polymers, indicating a lower moisture sorption capacity. The sorption capacity is generally attributed to the number of available sites (i.e. hydroxyl groups) for water binding (Bizot *et al.*, 1997). Gcn1 exhibited a high \bar{M}_w as determined by AFFFF-MALLS with a linear conformation (low amount of $\alpha(1\rightarrow3)$ linkages and high v_G value) and should display a comparable behaviour than other glucans. One reason for this peculiar behaviour could be attributed to polymer-polymer interactions between the linear chains of

Gcn1, probably enhanced by freeze-drying packing stress that decreases the accessibility of available hydroxyl groups.

Table 15: Water content, glass transition temperature (T_g) and changes in heat capacity at T_g .

Samples	Water content (%)	$T_{g \text{ midpoint}}$ (°C)	ΔC_p (J/g.°C)
Gcn_Dv Δ 4N	12.74	47.04	0.571
Gcn1	7.07	19.00	0.144
Gcn2	12.59	51.82	0.543
Gcn3	12.86	51.29	0.606
Gcn4	12.61	56.53	0.679
Gcn5	10.96	29.80	0.192
Gcn6	12.69	53.24	0.567
Gcn7	n.d.	n.d.	n.d.

n.d.: Not determined

The temperatures of glass transition reported in Table 15 correspond to the temperature at the midpoint of the transition heat capacity between the glassy and rubbery states measured during the second DSC scan (Figure 45). Gcn2, Gcn3 and Gcn6 displayed comparable T_g values in a range of 51 – 53°C. With the same water content, Gcn_Dv Δ 4N and Gcn4 displayed a slightly different T_g of 47.04 and 56.53°C, respectively.

Gcn1 and Gcn5 displayed a 1.5 to 2.5 times lower T_g values compared to other polymers. These results are very astonishing with regard to the water content values of these polymers. Indeed, water is known to have plasticizing effect (increase of the molecular mobility in the material and shift the T_g to a lower value) for many carbohydrates like mono-, di- and tri-saccharides (Noel *et al.*, 2000), dextran (Icoz *et al.*, 2005), pullulan (Diab *et al.*, 2001), galactan (Mitsuiki *et al.*, 1998), and starches (Bizot *et al.*, 1997).

These data suggest that molecular mobility in Gcn1 and Gcn5 is higher than in the others polymers. Therefore, their T_g being close to ambient temperature without adding any plasticizer, Gcn1 and Gcn5 present mechanical properties particularly suitable for material making.

Finally, as no direct correlations can be found between molecular mass, $\alpha(1\rightarrow3)$ linkage content, moisture content and T_g values, these results indicate significant differences in the molecular structure (branching degree, ramification length and organization) of these seven dextrans especially branching degree, distribution of branching points and ramification length. Indeed, it has been shown that chain length, branching scheme and the nature of α -linkages have a strong impact on the mechanical properties of polymers (Bizot *et al.*, 1997). For example Gcn1 which has the highest average molecular weight and the lowest T_g could contain few very long branchings. Also, when comparing Gcn4 and Gcn5

which have almost the same average molecular weight but very different T_g , the HMM fraction in Gcn5 (37.1% and v_G characteristic of more branched structures with long ramifications) could have a predominant role. Finally, some $\alpha(1\rightarrow3)$ linkages could also be present in the dextran backbone which would behave as a more linear polymer with a higher T_g . In any case, a more detailed knowledge of the structure would be necessary to establish direct relationships between the different data obtained.

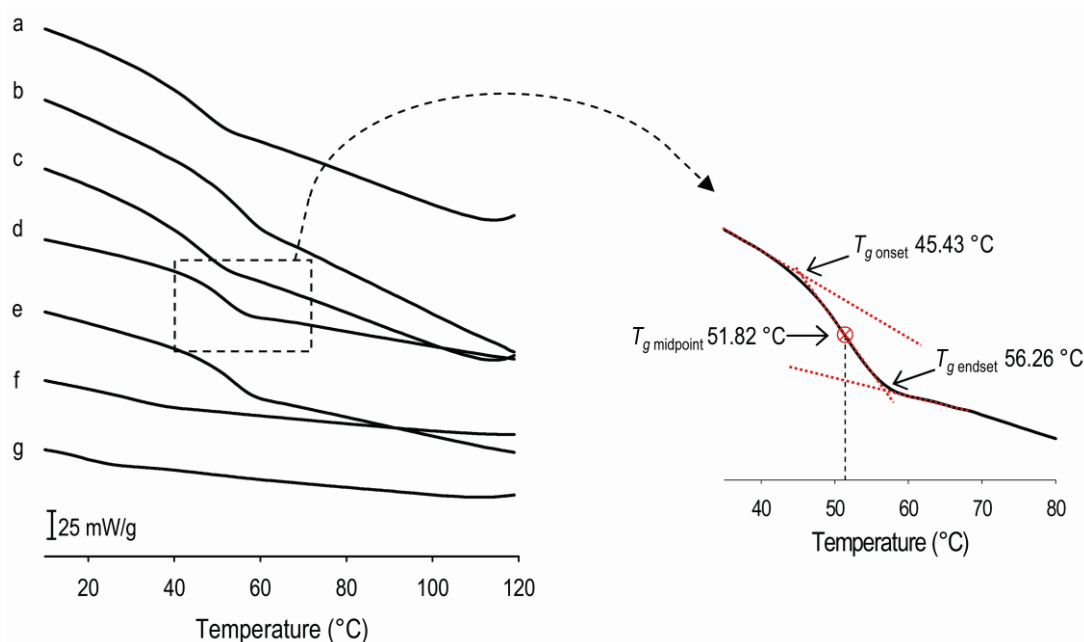


Figure 45: Differential scanning calorimetry thermograms for polymers (a) Gcn_Dv Δ 4N, (b) Gcn4, (c) Gcn3 (d) Gcn2 (e) Gcn6, (f) Gcn5 and (g) Gcn1.

Mechanical behaviour of the films

Translucent films were obtained for all tested α -glucans (Figure 46A). Films from dextrans displaying high T_g ($T_g \gg T_{room}$) were rigid and brittle while films obtained with Gcn1 and Gcn5 ($T_g \cong T_{room}$) were flexible. These observations were in accordance with the DSC results. Below T_g , materials were in the glassy state and mobility was mainly restricted to local chain segments (β relaxation) and hydroxyl groups (γ relaxation) leading to a lower amplitude of motion, explaining the rigid and brittle shape (Montès and Cavallé, 1999). Around and above T_g , amplitude and frequency of movements increased (α relaxation) leading to flexible, rubber materials (Champion *et al.*, 2000).

Unfortunately, the rigid films were very difficult to handle and were crumbled when cut for tensile tests. Gcn6 was also fragile and we did not obtain any dumbbell-shape specimen. For these reasons, mechanical testing was performed only on Gcn1 film. A tensile stress-elongation curve of Gcn1 film is presented in Figure 46B. This polymer displayed a ductile behaviour, as the rupture occurred in the

plastic domain, with a Young's modulus (E) of 0.019 GPa (calculated from the initial slope of the curve), a tensile strength at break (σ_R) of 3.55 MPa and an elongation at break (ϵ_R) of 193.2 %. The low Young's modulus and high elongation at break values revealed a film with a high elasticity. However the strength at break was low and could be related to the low ability to entangle also revealed by the viscoelastic behaviour. A comparison of these mechanical properties with those of other biopolymers and synthetic polymers is given in Table 16. The Gcn1 film displayed a very low Young's modulus and a high deformation capacity compared to most of the biopolymers, even when plasticizers, known to increase the flexibility, were added to complement the polymer. These results are remarkable but remain far from synthetic polymers which displayed 3 to 5 higher deformation capacities and 6 to 13 higher strength at break point (Gorna *et al.*, 2008). However, this is the first time that films are prepared from dextran. The large varieties of T_g determined for the series of studied dextrans could allow to design materials with specific mechanical behaviour either by adding a low amount of plasticizer or by mixing dextrans with different T_g . It could also be interesting to use Gcn1 in mixture with other biopolymers displaying complementary properties to produce biofilms closer to synthetic polymers.

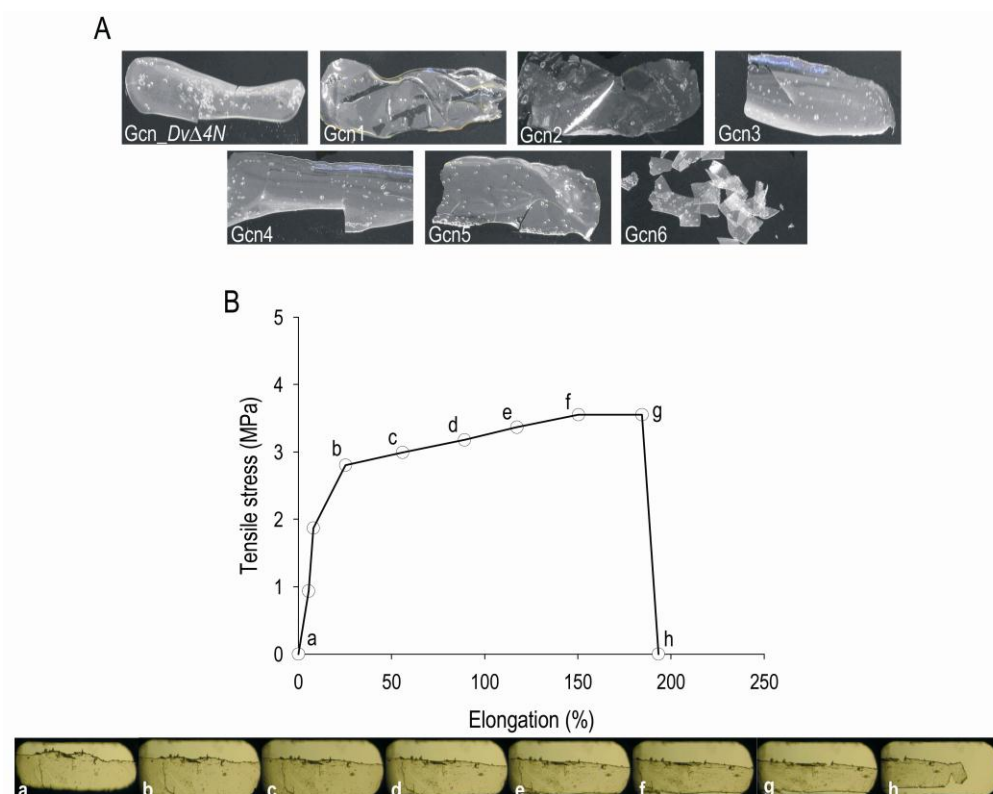


Figure 46: (A) Films obtained after casting and conditioning at a_w 0.57. (B) Tensile stress-elongation curve of Gcn1 films.

Table 16: Mechanical properties of Gcn1 in comparison with those of other polymers referred in the literature.

Film composition	E (GPa)	σ_R (MPa)	ϵ_R (%)	Water content (%)	Reference
Gcn1	0.019	3.55	193.2	-	present study
Pullulan	1.2	80		10	(Diab <i>et al.</i> , 2001)
Pullulan plasticized with sorbitol	0.565	11.85	454	-	(Kristo <i>et al.</i> , 2007)
EPS of <i>Pseudomonas oleovorans</i> NRRL B-14682	1.738	51	9.5	10	(Alves <i>et al.</i> , 2011)
Chitosan	-	60.7	3.3	15	(Garcia <i>et al.</i> , 2006)
Com starch	-	47.4	3.6	9	(Garcia <i>et al.</i> , 2006)
Com starch plasticized with glycerol	-	4 - 5	65 - 85	-	(Ghanbarzadeh <i>et al.</i> , 2010)
Cassava starch	21	1.59	83	12	(Müller <i>et al.</i> , 2009)
κ -carrageenan	1.11	57	6.8	20	(Lafargue <i>et al.</i> , 2007)
Kefiran	-	40.92	2.7	14.8	(Piermaria <i>et al.</i> , 2009)
Gellan gum plasticized with PEG	0.0045	25 - 30	5 -10	-	(Yang and Paulson, 2000)
Gellan gum plasticized with glycerol	0.0025	30	30	-	(Yang and Paulson, 2000)
Soy protein	-	2.8	165.7	25	(Rhim <i>et al.</i> , 2007)
Cellophane	-	85.8	14.4	-	(Shellhammer and Krochta, 1997)
PP ^a	0.798	47	933	-	(Gorna <i>et al.</i> , 2008)
HDPE ^b	0.631	38	531	-	(Gorna <i>et al.</i> , 2008)
LLDPE ^c	0.161	22	686	-	(Gorna <i>et al.</i> , 2008)

^a Polypropylene

^b High density polyethylene

^c Linear low density polyethylene

Conclusion

The structural and physico-chemical properties of seven original α -glucans, enzymatically synthesized in one step from sucrose were investigated. AFFFF-MALLS results enabled to clusterize these polymers in three groups, regarding their size and branching patterns: (i) Gcn1, which displays the highest \bar{M}_w and a linear conformation, (ii) Gcn6 and Gcn7, the smallest α -glucans, that are also probably the most branched, and (iii) the other α -glucans displaying intermediary structures.

In solution, all these α -glucans displayed non-Newtonian and non-thixotropic behaviours. However, analyses in dynamic conditions suggested that they adopt different conformation arrangement. For the first time, high molecular weight dextran glass transition was also studied, and revealed a particularly low T_g for Gcn1 and Gcn5. Gcn1 films displayed very interesting mechanical properties compared to other biopolymers extracted from plants, algae or those produced by microbial fermentation. To our knowledge, this is the first time that films were obtained from dextrans. These results open novel perspectives to dextran applications, as bio-sourced materials and biodegradable films.

Finally, further structural investigations will be necessary to characterize accurately the ramification length and organization and the supra-molecular organisation of the most promising dextrans, and to establish reliable correlations between structure and physical properties of α -glucans.

Acknowledgments

We cordially thank Marion de Carvalho and Laurent Chaunier for technical support.

This work was funded by the French National Institute for Agricultural Research (INRA), the French National Center for Scientific Research (CNRS) and the French National Research Agency (Project ANR-09-CP2D-07-03).

References

- Alves, V.D., Ferreira, A.R., Costa, N., Freitas, F., Reis, M.A.M., Coelho, I.M.,** 2011. Characterization of biodegradable films from the extracellular polysaccharide produced by *Pseudomonas oleovorans* grown on glycerol byproduct. *Carbohydr. Polym.* **83**, 1582-1590.
- André, I., Potocki-Véronèse, G., Morel, S., Monsan, P., Rемаud-Siméon, M.,** 2010. Sucrose-utilizing transglucosidases for biocatalysis, Carbohydrates in Sustainable Development I, Topics in Current Chemistry. Springer Berlin / Heidelberg, p. 25.
- Bizot, H., Le Bail, P., Leroux, B., Davy, J., Roger, P., Buleon, A.,** 1997. Calorimetric evaluation of the glass transition in hydrated, linear and branched polyanhydroglucose compounds. *Carbohydr. Polym.* **32**, 33-50.
- Borde, B., Bizot, H., Vigier, G., Buleon, A.,** 2002. Calorimetric analysis of the structural relaxation in partially hydrated amorphous polysaccharides. I. Glass transition and fragility. *Carbohydr. Polym.* **48**, 83-96.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B.,** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucl. Acids Res.* **37**, 233-238.
- Champion, D., Le Meste, M., Simatos, D.,** 2000. Towards an improved understanding of glass transition and relaxations in foods: molecular mobility in the glass transition range. *Trends Food Sci. Technol.* **11**, 41-55.
- Cui, H., Liu, Q., Tao, Y., Zhang, H., Zhang, L., Ding, K.,** 2008. Structure and chain conformation of a (1→6)- α -D-glucan from the root of *Pueraria lobata* (Willd.) Ohwi and the antioxidant activity of its sulfated derivative. *Carbohydr. Polym.* **74**, 771-778.
- Diab, T., Biliaderis, C.G., Gerasopoulos, D., Sfakiotakis, E.,** 2001. Physicochemical properties and application of pullulan edible films and coatings in fruit preservation. *J. Sci. Food Agric.* **81**, 988-1000.
- Doublier, J.L., Wood, P.J.,** 1995. Rheological properties of aqueous solutions of (1→3)(1→4)- β -D-glucan from oats (*Avena sativa* L.). *Cereal Chem.* **72**, 335-340.
- Fabra, M.J., Talens, P., Chiralt, A.,** 2008. Effect of alginate and l-carrageenan on tensile properties and water vapour permeability of sodium caseinate-lipid based films. *Carbohydr. Pol.* **74**, 419-426.
- Garcia, M.A., Pinotti, A., Zaritzky, N.E.,** 2006. Physicochemical, water vapor barrier and mechanical properties of corn starch and chitosan composite films. *Starch* **58**, 453-463.
- Ghanbarzadeh, B., Almasi, H., Entezami, A.A.,** 2010. Improving the barrier and mechanical properties of corn starch-based edible films: effect of citric acid and carboxymethyl cellulose. *Ind. Crops Prod.* **33**, 229-235.
- Gorna, K., Hund, M., Vucak, M., Gröhn, F., Wegner, G.,** 2008. Amorphous calcium carbonate in form of spherical nanosized particles and its application as fillers for polymers. *Mater. Sci. Eng., A* **477**, 217-225.
- Icoz, D.Z., Moraru, C.I., Kokini, J.L.,** 2005. Polymer-polymer interactions in dextran systems using thermal analysis. *Carbohydr. Pol.* **62**, 120-129.

- Ioan, C.E., Aberle, T., Burchard, W.**, 2001. Structure properties of dextran. 3. Shrinking factors of individual clusters. *Macromolecules* **34**, 3765-3771.
- Irague, R., Massou, S., Moulis, C., Saurel, O., Milon, A., Monsan, P., Remaud-Siméon, M., Portais, J.C., Potocki-Véronèse, G.**, 2011a. NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity. *Anal. Chem.* **83**, 1202-1206.
- Irague, R., Roncalli-Tarquis, L., André, I., Moulis, C., Morel, S., Monsan, P., Potocki-Véronèse, G., Remaud-Siméon, M.**, 2011b. Combinatorial engineering of dextransucrase specificity. *In preparation*.
- Kristo, E., Biliaderis, C.G., Zampraka, A.**, 2007. Water vapour barrier and tensile properties of composite caseinate-pullulan films: Biopolymer composition effects and impact of beeswax lamination. *Food Chem.* **101**, 753-764.
- Lafargue, D., Lourdin, D., Doublier, J.-L.**, 2007. Film-forming properties of a modified starch/ κ -carrageenan mixture in relation to its rheological behaviour. *Carbohydr. Pol.* **70**, 101-111.
- Lee, S., Kinney, M.P., Inglett, G.E.**, 2005. Rheological characterization of a new oat hydrocolloid and its application in cake baking. *Cereal Chem.* **82**, 717-720.
- Lichtenthaler, F.W., Peters, S.**, 2004. Carbohydrates as green raw materials for the chemical industry. *Comptes Rendus Chimie* **7**, 65-90.
- Maina, N.H., Tenkanen, M., Maaheimo, H., Juvonen, R., Virkki, L.**, 2008. NMR spectroscopic analysis of exopolysaccharides produced by *Leuconostoc citreum* and *Weissella confusa*. *Carbohydr. Res.* **343**, 1446-1455.
- Mitsuiki, M., Yamamoto, Y., Mizuno, A., Motoki, M.**, 1998. Glass transition properties as a function of water content for various low-moisture galactans. *J. Agric. Food Chem.* **46**, 3528-3534.
- Montès, H., Cavallé, J.Y.**, 1999. Secondary dielectric relaxations in dried amorphous cellulose and dextran. *Polymer* **40**, 2649-2657.
- Moulis, C., Arcache, A., Escalier, P.C., Rinaudo, M., Monsan, P., Remaud-Simeon, M., Potocki-Veronese, G.**, 2006a. High-level production and purification of a fully active recombinant dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *FEMS Microbiol. Lett.* **261**, 203-210.
- Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M.**, 2006b. Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J. Biol. Chem.* **281**, 31254-31267.
- Müller, C.M.O., Laurindo, J.B., Yamashita, F.**, 2009. Effect of cellulose fibers on the crystallinity and mechanical properties of starch-based films at different relative humidity values. *Carbohydr. Pol.* **77**, 293-299.
- Noel, T.R., Parker, R., Ring, S.G.**, 2000. Effect of molecular structure and water content on the dielectric relaxation behaviour of amorphous low molecular weight carbohydrates above and below their glass transition. *Carbohydr. Res.* **329**, 839-845.

- Piermaria, J.A., Pinotti, A., Garcia, M.A., Abraham, A.G.**, 2009. Films based on kefiran, an exopolysaccharide obtained from kefir grain: Development and characterization. *Food Hydrocolloids* **23**, 684-690.
- Planchot, V., Colonna, P., Saulnier, L.**, 1997. Séparation et dosage des glucides et des amylases. In: *Godon B., Loisel, W. (eds), Guide pratique d'analyse dans les industries des céréales. Lavoisier, Paris*, 341-398.
- Rhim, J.-W., Lee, J.H., Ng, P.K.W.**, 2007. Mechanical and barrier properties of biodegradable soy protein isolate-based films coated with polylactic acid. *LWT - Food Sci. Technol.* **40**, 232-238.
- Rolland-Sabaté, A., Colonna, P., Mendez-Montealvo, M.G., Planchot, V.**, 2007. Branching features of amylopectins and glycogen determined by asymmetrical flow field flow fractionation coupled with multiangle laser light scattering. *Biomacromolecules* **8**, 2520-2532.
- Sabatié, J., Choplin, L., Paul, F., Monsan, P.**, 1986. The effect of synthesis temperature on the rheological properties of native dextran. *Biotechnol. Lett.* **8**, 425-430.
- Seymour, F.R., Knapp, R.D., Bishop, S.H.**, 1979. Correlation of the structure of dextrans to their ¹H-N.M.R. spectra. *Carbohydr. Res.* **74**, 77-92.
- Shellhammer, T.H., Krochta, J.M.**, 1997. Whey protein emulsion film performance as affected by lipid type and amount. *J. Food Sci.* **62**, 390-394.
- Siracusa, V., Rocculi, P., Romani, S., Rosa, M.D.**, 2008. Biodegradable polymers for food packaging: a review. *Trends Food Sci. Technol.* **19**, 634-643.
- Sumner, J.B., Howell, S.F.**, 1935. A method for determination of saccharase activity. *J. Biol. Chem.* **108**, 51-54.
- Tokita, M., Nishinari, K., Nishinari, K.**, 2009. Some thoughts on the definition of a gel, Gels: Structures, Properties, and Functions, Prog. Colloid Polym. Sci. Springer Berlin / Heidelberg, p. 87.
- Van Hijum, S.A.F.T., Kralj, S., Ozimek, L.K., Dijkhuizen, L., van Geel-Schutten, I.G.H.**, 2006. Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 157-176.
- Van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P.**, 2009. Structural characterization of bioengineered α -D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180. *Biomacromolecules* **10**, 580-588.
- Van Leeuwen, S.S., Kralj, S., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P.**, 2008a. Structural analysis of bioengineered α -D-glucan produced by a triple mutant of the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180: generation of (α 1 \rightarrow 4) linkages in a native (1 \rightarrow 3)(1 \rightarrow 6)- α -D-glucan. *Biomacromolecules* **9**, 2251.
- Van Leeuwen, S.S., Kralj, S., Van Geel-Schutten, I.H., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P.**, 2008b. Structural analysis of the α -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme. *Carbohydr. Res.* **343**, 1237.

Vandamme, E., Renard, C., Arnaut, F., Vekemans, N., Tossut, P., 2003. Process for obtaining improved structure build-up of baked products. US patent 6627235,

Yang, L., Paulson, A.T., 2000. Mechanical and water vapour barrier properties of edible gellan films. *Food Res. Int.* **33**, 563.

CONCLUSION

Les travaux effectués aux cours de cette thèse avaient pour objectif de générer de nouveaux outils enzymatiques pour la synthèse d' α -glucanes et de gluco-oligosaccharides de structures et de propriétés originales, à partir de saccharose. Cette démarche s'inscrit dans un contexte de diversification de polymères bio-sourcés. L'enjeu était double. D'une part, il s'agissait d'approfondir la compréhension des relations structure/fonction des glucane-saccharases de la famille GH70, afin d'identifier de nouveaux déterminants de la spécificité de ces enzymes et de concevoir, par ingénierie des protéines, des transglucosidases efficaces de spécificités variées. D'autre part, l'objectif était de mettre en œuvre ces enzymes pour élargir la gamme d' α -glucanes disponibles, et de caractériser les propriétés physico-chimiques de ces nouveaux biopolymères, afin d'évaluer leur potentiel applicatif.

Au début de ces travaux de thèse, seuls quelques résidus du domaine catalytique des glucane-saccharases de la famille GH70, également conservés dans les enzymes de la famille GH13, étaient connus pour affecter la nature des liaisons osidiques synthétisées par certaines de ces enzymes (Funane *et al.*, 2004; Kralj *et al.*, 2008). De plus, seule une vingtaine de séquences codant pour des glucane-saccharases caractérisées étaient disponibles. Enfin, la structure cristallographique de la première glucane-saccharase de la famille 70 (GTF180) avait été récemment résolue, mais les données structurales n'avaient pas encore été relâchées dans la Protein Data Bank (Vujičić Žagar, 2007). Dans ces conditions, il n'était pas concevable d'appréhender l'ingénierie de ces enzymes par une approche rationnelle, c'est pourquoi nous avons envisagé d'appliquer les approches d'ingénierie aléatoire et semi-aléatoire pour créer des banques de variants ciblées sur le domaine catalytique des glucane-saccharases. Néanmoins, aucune méthode de criblage fonctionnel à haut-débit permettant de trier les glucane-saccharases sauvages et mutantes sur la base de la spécificité de liaisons, n'avait été décrite, ce qui représentait un frein majeur à l'exploration de larges banques de variants issues de l'ingénierie combinatoire. Enfin, si l'analyse structurale de certains α -glucanes synthétisés par les glucane-saccharases est bien décrite, leurs propriétés physico-chimiques ont été très peu étudiées, hormis pour les dextrans commerciaux et les produits amylopolysaccharidiques.

Pour lever ces verrous qui limitent autant la compréhension des relations structure-spécificité des glucane-saccharases que l'exploitation des produits qu'elles synthétisent, les travaux réalisés au cours de cette thèse ont consisté à:

- développer une méthodologie à la fois rapide, sensible et robuste permettant d'explorer de larges banques de variants et de trier les glucane-saccharases en fonction de leur activité et de leur spécificité de liaisons

- mettre en œuvre une stratégie d'ingénierie raisonnée pour l'identification de nouveaux déterminants structuraux impliqués dans la spécificité de liaisons des glucane-saccharases.
- caractériser les propriétés physico-chimiques des néo-biopolymères synthétisés par les variants enzymatiques les plus originaux.

Nous proposons à présent de dresser les différentes conclusions que nous pouvons tirer de notre étude et d'en dégager des perspectives à court et moyen termes.

Développement d'une stratégie multi-étape pour le criblage de la spécificité des glucane-saccharases

Afin de sélectionner des mutants actifs au sein de très larges banques de glucane-saccharases, et donc de réduire le nombre de mutants à cribler ultérieurement, nous avons adapté une méthode de sélection *in vivo*, réalisée sur milieu minimum solide additionné de saccharose. Cette méthode repose sur le fait que la production d'une glucane-saccharase active par un clone recombinant permet le clivage du saccharose et la libération du fructose, qui est utilisé comme source de carbone pour la croissance de la souche hôte d'*E. coli* BL21 AI, incapable de métaboliser le saccharose. Cette méthode permet ainsi d'isoler après 7 jours de croissance à 20°C les clones positifs, à une cadence de 200 000 clones par semaine. De plus, les variants de glucane-saccharase ayant conservé une activité de polymérisation des unités glucosyle sont facilement identifiables par la présence d'une bulle de polymère recouvrant la colonie correspondante.

Cette méthode de sélection *in vivo* est très simple à mettre en œuvre et permet d'explorer rapidement des banques de variants de 10⁵ à 10⁶ clones. Cependant, elle est uniquement qualitative et ne permet donc pas d'évaluer avec précision le niveau d'activité d'un variant par rapport à un autre, bien qu'il soit possible d'apprécier les différences de cinétique de croissance des clones affectant la taille des colonies et la formation des bulles de polymères. De plus, il est très difficile d'identifier directement à partir de ce test sur milieu solide les différences de structures ou de propriétés rhéologiques des polymères observés. En effet, même si l'aspect des bulles de polymères formées par des colonies exprimant l'enzyme DSR-S [spécifique de liaisons $\alpha(1\rightarrow6)$] et DSR-E [spécifique de liaisons $\alpha(1\rightarrow6)$, $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow2)$] est différent, celui des colonies produisant des variants synthétisant des taux variables de liaisons $\alpha(1\rightarrow3)$ est beaucoup moins contrasté.

Cette procédure initiale de sélection/criblage constitue un premier « filtre » servant à réduire considérablement la taille des banques de glucane-saccharases à cribler. Elle est un pré-requis au

criblage de la spécificité de liaison des glucane-saccharases et pourra, à l'avenir, être utilisée dans ce même but et ainsi faciliter l'(les) étape(s) de criblage lui succédant.

La discrimination de la spécificité de liaison des glucane-saccharases sauvages ou de leurs variants repose sur l'identification de la nature des liaisons osidiques des produits synthétisés. Pour développer un crible efficace, nous avons sélectionné la RMN 1D du proton et nous nous sommes basés sur le concept de « rapporteurs structuraux » ou signaux spécifiques correspondant aux déplacements chimiques des protons anomériques des unités glucosyle, en fonction de la nature de la liaison osidique. Les glucane-saccharases ont été mises en œuvre en présence de saccharose et d'accepteurs isomaltooligosaccharides (DP moyen 7). Les glucooligosaccharides présentent l'avantage d'être beaucoup moins visqueux que les polymères et sont généralement plus solubles, ce qui facilite l'analyse par RMN. De plus, la spécificité des glucane-saccharases est généralement conservée entre les réactions de polymérisation et d'accepteurs.

Pour répondre aux exigences de criblage à haut débit, une procédure d'analyse totalement automatisée par ¹H-RMN a été établie. Elle repose sur l'emploi d'un robot de pipettage compatible au format de plaques 96 puits, et d'un système d'injection directe dans la cryosonde du spectromètre RMN permettant de faire l'analyse en flux des échantillons. La mise au point a été réalisée à partir de clones recombinants exprimant des enzymes sauvages de spécificités variées (DSR-S vardel Δ4N, DSR-E et ASR C-del) cultivés au format micro-plaque, lysés et mis en présence de saccharose et d'accepteurs. Elle a montré que la régiospécificité des glucane-saccharases pouvait être déterminée à une cadence de 480 clones par jour, de façon quantitative et qualitative, par le biais de la mesure de l'intensité des signaux rapporteurs directement corrélée à la spécificité de liaison. Cette méthode de criblage est aujourd'hui la seule qui soit adaptée au criblage haut débit de la régiospécificité des glucane-saccharases.

Certes, la méthode nécessite de disposer d'un spectromètre RMN équipé d'un module d'analyse des échantillons en flux, équipement lourd pas toujours à la portée de tout expérimentateur. De plus, la cadence de 480 échantillons analysés par jour, même si elle est exceptionnelle pour ce genre d'appareillage, reste modeste comparée à des méthodes telles que la compartimentation *in vitro* par exemple (Agresti *et al.*, 2010), d'où l'intérêt de la première étape de sélection. Malgré ces inconvénients, la méthode présente de nombreux atouts.

- × le criblage peut être réalisé directement à partir de milieux réactionnels bruts, sans étape de purification des produits à analyser.
- × aucun substrat ou analogue de substrat marqué par modification chimique n'est nécessaire. Cela diminue les coûts de la procédure, et surtout élimine le risque d'isolement de

faux positifs (enzymes de spécificité modifiée pour la reconnaissance du substrat marqué mais de régiospécificité identique vis-à-vis du substrat naturel).

× La méthode est non destructive, reproductible et sensible, puisque qu'elle permet de détecter les liaisons osidiques de produits à des concentrations de l'ordre de la centaine de micro-molaire.

Enfin, à condition que l'on puisse identifier un(des) rapporteur(s) structural(aux) approprié(s), cette méthode générique peut être étendue au criblage de la spécificité de liaison d'autres CAZymes de type transglycosidases et glycosyl-transférases. D'une façon générale, l'approche ouvre des perspectives prometteuses pour l'exploration de la diversité fonctionnelle des CAZymes naturelles et artificielles.

Ingénierie des glucane-saccharases de la famille GH70 par family shuffling

En l'absence de structure tridimensionnelle disponible pour les enzymes de la famille GH70, la première stratégie d'ingénierie que nous avons envisagée pour étudier les relations structure-spécificité des glucane-saccharases était basée sur le brassage aléatoire de gènes codant pour des enzymes de spécificités différentes, à savoir DSR-S, DSR-E, ASR et GTF-I, par une approche de type family shuffling. Le pourcentage d'identité entre ces différents gènes étant très faible (<50%), nous nous sommes focalisés sur le brassage des régions géniques de plus fortes identités (54 à 62 %), correspondant aux domaines catalytiques de ces enzymes. Afin de faciliter la recombinaison hétérologue, nous avons opté pour la méthode de family shuffling par brassage d'ADN simple brin (Zha *et al.*, 2003). Une banque de $3,4 \cdot 10^4$ clones a été obtenue après brassage des gènes codant pour les DSR-S et GTF-I, et remplacement du domaine catalytique de DSR-S vardel $\Delta 4N$ par les séquences chimères. Malheureusement, le séquençage de 10 clones de la banque initiale a révélé un très faible taux de recombinaison hétérologue, un seul clone portant lui-même un seul point de recombinaison ayant été isolé. Malgré la faible diversité générée au sein de cette banque, nous l'avons soumise aux approches de sélection et de criblage que nous avons développées. Parmi 768 clones actifs, aucun clone produisant un variant de spécificité de liaison altérée par rapport à DSR-S vardel $\Delta 4N$ n'a pu être identifié.

Le faible pourcentage d'identité est apparu comme une véritable limite à la génération d'évènements de recombinaison entre les séquences codant pour différentes glucane-saccharases. La stratégie ciblant le brassage des régions les plus conservées et l'emploi d'une méthode optimisée pour augmenter théoriquement les points de croisements n'ont pas été suffisants pour s'affranchir de cet obstacle. Aujourd'hui, à la lumière des deux structures cristallographiques résolues récemment, il apparaît que

l'approche aléatoire avait peu de chances de succès. En effet, il semble très ambitieux d'imaginer reconstruire la structure en U si caractéristique des enzymes de la famille GH70, et donc un domaine catalytique actif, à partir de l'enchevêtrement aléatoire de régions peptidiques N et C-terminales issues de protéines distinctes, caractérisées par des boucles de différentes longueurs. A l'avenir, une approche de brassage plus raisonnée ciblant des points de recombinaison identifiés par alignement structural, pourra être envisagée.

Le temps consacré à ces premiers essais et les difficultés rencontrées pour l'obtention de cette première banque nous ont incités à abandonner cette approche, pour nous consacrer à l'exploitation des banques conçues de façon semi-rationnelle.

Ingénierie semi-rationnelle de la dextrane-saccharase DSR-S

En parallèle à la stratégie d'ingénierie aléatoire, nous avons travaillé sur une approche semi-rationnelle de l'ingénierie de la spécificité de liaison de la dextrane-saccharase DSR-S. Dans un premier temps, nous avons d'abord analysé les alignements de séquences des enzymes de la famille GH70 caractérisées sur le plan fonctionnel. Cette analyse a été complétée par l'analyse de la structure tridimensionnelle de la GTF180, qui a permis de localiser une région non conservée dans la famille GH70 qui pourrait jouer un rôle dans le positionnement des oligosaccharides dans le site actif. Au total, huit positions (D306, F353, N404, W440, D460, H463, T464 et S512) ont été ciblées. Deux banques de variants ciblant chacune quatre positions et comprenant $2,5 \cdot 10^5$ et $3,8 \cdot 10^4$ clones ont été construites. Une méthode adaptée de la méthode ISOR permettant à la fois la mutation à saturation des acides aminés ciblés et la recombinaison des mutations a été adoptée pour générer la diversité. Après sélection et criblage de la spécificité par RMN, 82 mutants présentant une spécificité de liaison altérée par rapport à l'enzyme parentale ont été isolés. Ceux-ci constituent une boîte à outils enzymatiques d'intérêt pour la synthèse contrôlée de gluco-oligosaccharides et de dextrans de structure originale, contenant une teneur croissante en liaisons $\alpha(1 \rightarrow 3)$.

Le séquençage des clones (510 au total) à chaque étape du processus a permis de montrer que les positions D306, N404 et W440 jouent un rôle clef pour la reconnaissance du substrat et/ou dans l'activité de transglucosylation de l'enzyme, puisque seulement 20 % des clones actifs séquencés ont été mutés à ces positions. Après criblage, le séquençage de 267 clones produisant des variants de spécificité altérée a montré que, mutée individuellement, la position S512 joue un rôle important dans la spécificité de l'enzyme, ce qui avait déjà été observé pour la GTF-A (Kralj *et al.*, 2006). L'analyse des variants a aussi révélé le rôle clef des positions D460, H463 et T464 pour le contrôle de la spécificité de DSR-S. En effet, les triple-mutants H463R/T464D/S512T, H463R/T464V/S512T, D460A/H463S/T464L

ainsi que le quadruple-mutant D460M/H463Y/T464M/S512C synthétisent des oligosaccharides avec un taux de liaisons $\alpha(1\rightarrow3)$ de 6 à 8 fois supérieur à celui observé dans les oligosaccharides produits par l'enzyme parentale, ce qui représente les changements de spécificités les plus drastiques.

Parmi les 82 mutants de spécificité altérée, sept mutants (F353W, F353T, S512C, H463R/T464D/S512T, H463R/T464V/S512T, D460A/H463S/T464L et D460M/H463Y/T464M/S512C) ont été choisis. Ils synthétisent des oligosaccharides et polysaccharides contenant des taux variables de liaisons $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow6)$ et sont représentatifs du panel de variants obtenus. Après production et purification (à l'échelle de la dizaine de mg), l'étude biochimique de ces variants complétée par des travaux de modélisation moléculaire a été réalisée pour mieux comprendre les effets des mutations. Enfin, les polymères formés à partir de saccharose ont été analysés par un large panel de techniques d'analyse structurale et des propriétés physico-chimiques des polysaccharides.

Compréhension du rôle des mutations

En présence de saccharose, les variants sélectionnés synthétisent deux populations de produits glucosidiques, en plus du leucrose, du glucose et du fructose libérés dans le milieu réactionnel : une population de haute masse molaire ($>10^7$ g.mol⁻¹), et une de masse molaire intermédiaire (0,5 à 5,7.10³ g.mol⁻¹). L'analyse par RMN du proton des α -glucanes de haute masse molaire purifiés a montré, en accord avec les résultats de criblage, que les mutants sélectionnés synthétisent des taux de liaisons $\alpha(1\rightarrow3)$ variant entre 3 et 20 %. De plus, nous avons montré que les mutants F353W et F353T synthétisent environ 3 fois plus de leucrose que DSR-S vardel Δ 4N, au détriment de la synthèse de polymère. L'analyse du comportement cinétique de ces deux variants a montré que leur affinité vis-à-vis du saccharose est drastiquement modifiée par rapport à celle de DSR-S vardel Δ 4N. De plus, le mutant F353W présente un mécanisme Michaelien, alors que l'enzyme parentale est inhibée par excès de substrat. Ainsi, la position F353 jouerait un rôle important dans le mécanisme de reconnaissance de l'unité fructosyle du saccharose. L'absence d'inhibition par excès de saccharose a aussi été observée pour le mutant H463R/T464V/S512T, qui conserve néanmoins des valeurs de K_m et k_{cat} proches de celles de l'enzyme parentale. Ce mutant présente donc un intérêt applicatif particulier, puisqu'il est 3 fois plus rapide que l'enzyme sauvage à haute concentration de saccharose, ce qui représente un avantage pour l'industrialisation des procédés de synthèse d'oligo- et polysaccharides.

Enfin, la modélisation de la structure tridimensionnelle du domaine catalytique de la DSR-S et du mutant D460M/H463Y/T464M/S512C a permis de localiser les résidus ciblés dans le cadre de cette étude. En accord avec les résultats de criblage et caractérisation fonctionnelle des mutants obtenus, il s'avère que les résidus mutés participent à la topologie du site actif, formant la première sphère

d'acides aminés autour de la diade catalytique (D400 et E438) et du stabilisateur de l'intermédiaire covalent (D511). Des études de « docking » du mutant D460M/H463Y/T464M/S512C avec le tetrasaccharide α -D-Glcp-(1→6)[α -D-Glcp-(1→3)] α -D-Glcp-(1→6)-D-Glcp suggèrent que ces 4 mutations, localisées au niveau du sous-site putatif +2, pourraient favoriser l'arrimage des chaînes linéaires liées en α (1→6) et ainsi faciliter la formation de ramifications en α (1→3). L'analyse encore plus fine des produits formés (localisation des liaisons α (1→3), longueur des branchements) et/ou la structure de complexes de glucane-saccharases-isomaltooligosaccharides ramifiés ou non en α (1→3) aurait été très utile pour améliorer nos modèles et nos interprétations. Par ailleurs, l'analyse structurale comparative des structures cristallographiques des enzymes GTF180, GTF-SI et celle du mutant Δ N₁₂₃-GBD-CD2, qui présentent toutes les trois des spécificités de liaison différentes devrait aussi être très informative. En l'absence de ces données, nos hypothèses doivent être considérées avec précaution. La caractérisation structurale d'autres glucane-saccharases, la comparaison de la topologie de leur site actif et l'analyse des interactions enzyme-ligand seront à l'avenir essentielles pour identifier de nouveaux déterminants de la spécificité des glucane-saccharases, guider la recherche de nouvelles cibles de mutagenèse et concevoir des glucane-saccharases artificielles capables de synthétiser des teneurs contrôlées en liaisons glucosidiques α (1→6), α (1→2), α (1→3) et α (1→4).

Caractérisation des propriétés physico-chimiques des nouveaux dextrans

Les polymères synthétisés par les sept mutants sélectionnés ont été analysés par différentes techniques afin de caractériser leurs propriétés physico-chimiques et mettre en relief leurs éventuelles différences. Ces travaux ont été réalisés au sein de l'unité Biopolymères Interactions Assemblages de l'INRA de Nantes.

L'analyse par AFFFF-MALLS a montré que les glucanes formés ont une masse molaire très élevée comprise entre 0,8 et 5,9.10⁸ g.mol⁻¹. Elle nous permet de suggérer que les α -glucanes les plus petits en taille ($\bar{M}_w \cong 1.10^8$ g.mol⁻¹) sont probablement les plus branchés, et qu'ils présentent les ramifications les plus longues. A l'inverse, les α -glucanes de plus grandes tailles sembleraient adopter une conformation plus linéaire.

L'analyse rhéologique des solutions de polymère concentrées à 5 % (m/v) a quant à elle révélé un comportement pseudo-plastique classique pour ce type de polymères, avec des valeurs de viscosités à faibles vitesses de cisaillement comprises entre 0,48 à 147 Pa.s. La détermination du comportement viscoélastique a permis de distinguer deux groupes de polysaccharides, sans qu'il soit possible de corréler ces observations à la teneur en liaisons α (1→3). En effet, les glucanes synthétisés par

l'enzyme parentale DSR-S vardel Δ 4N et les mutants F353W, S512C et D460M/H463Y/T464M/S512C forment des gels faibles, probablement dus à la formation de zones de jonction entre les chaînes. À l'inverse, les α -glucanes produits par les mutants F353T, H463R/T464D/S512T, H463R/T464V/S512T et D460A/H463S/T464L, présentent un comportement typique d'une solution de macromolécules enchevêtrées, sans réelles interactions.

Enfin, l'étude de filmabilité a révélé des caractéristiques très originales pour deux des polymères étudiés, synthétisés par les mutants F353T et H463R/T464V/S512T. En effet, ceux-ci présentent des valeurs de T_g de 19 et 30°C, respectivement, ce qui est significativement plus faible que les valeurs obtenues pour les autres α -glucanes, y compris le dextrane synthétisé par DSR-S vardel Δ 4N, pour lesquels les valeurs de T_g sont comprises entre 47 et 57°C. En accord avec ces valeurs de T_g , les films formés par les polymères synthétisés par les mutants F353T et H463R/T464V/S512T sont flexibles à température ambiante. Les valeurs de déformation et de module de Young, déterminées pour le polymère synthétisé par le mutant H463R/T464V/S512T, sont d'ailleurs tout à fait remarquables en comparaison avec d'autres biopolymères extraits de plantes ou produits par fermentation microbienne. Cette caractéristique peut être mise à profit pour l'obtention de nouveaux films biodégradables, et plus généralement pour la fabrication de matériaux bio-sourcés.

Ces travaux ont donc mis en évidence que la viscosité des polysaccharides étudiés, leur température de transition vitreuse et leur filmabilité ne sont pas corrélées au seul rapport entre liaisons $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow6)$, indiquant clairement que ces polysaccharides présentent des structures différentes. La caractérisation structurale approfondie de ces macromolécules, au niveau moléculaire et supra-moléculaire (notamment par spectroscopie RMN, dégradation chimique et enzymatique, méthylation et acétylation, mais aussi cryo-microscopie et diffusion des rayons X aux petits angles) est donc nécessaire pour déterminer plus finement les arrangements structuraux, notamment au niveau de l'organisation et de la longueur des ramifications. Cela permettrait d'établir avec une plus grande précision les relations entre structure et propriétés physico-chimiques de ces polysaccharides, de rationaliser la synthèse de ces molécules et d'être capable de sélectionner le catalyseur optimal pour la synthèse d'un α -glucane adapté à une application donnée.

En conclusion, ces travaux illustrent l'efficacité de l'ingénierie combinatoire couplée à une méthodologie de criblage haut-débit pour l'exploration des relations structure/spécificité des transglycosidases et la conception de nouvelles enzymes permettant de diversifier le panel de polysaccharides produits. Les résultats obtenus ouvrent de nouvelles perspectives d'applications de ces polymères en tant que bio-matériaux compatibles avec les exigences environnementales actuelles.

Agresti, J.J., Antipov, E., Abate, A.R., Ahn, K., Rowat, A.C., Baret, J.C., Marquez, M., Klivanov, A.M., Griffiths, A.D., Weitz, D.A., 2010. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *PNAS* **107**, 4004-4009.

Funane, K., Ishii, T., Terasawa, K., Yamamoto, T., Kobayashi, M., 2004. Construction of chimeric glucansucrases for analyzing substrate-binding regions that affect the structure of glucan products. *Biosci. Biotechnol. Biochem.* **68**, 1912-1920.

Kralj, S., Eeuwema, W., Eckhardt, T.H., Dijkhuizen, L., 2006. Role of asparagine 1134 in glucosidic bond and transglycosylation specificity of reuteransucrase from *Lactobacillus reuteri* 121. *Febs J.* **273**, 3735-3742.

Kralj, S., Van Leeuwen, S.S., Valk, V., Eeuwema, W., Kamerling, J.P., Dijkhuizen, L., 2008. Hybrid reuteransucrase enzymes reveal regions important for glucosidic linkage specificity and the transglucosylation/hydrolysis ratio. *FEBS J.* **275**, 6002.

Vujičić Žagar, A., 2007. Structural and functional investigations of *Lactobacillus reuteri* glucansucrase. University of Groningen, Groningen.

Zha, W., Zhu, T., Zhao, H., 2003. Family shuffling with single-stranded DNA. *Methods in Molecular Biology*, Volume 231, 1, 91-97, DOI: 10.1385/1-59259-395-X:91.

Tables des illustrations

FIGURES

Figure 1 : Réactions catalysées par les glucane-saccharases. +1 : formation du complexe β -D-glucosyl-enzyme ; -1 : échange isotopique ; 2 : réaction de polymérisation ; 3 : hydrolyse ; 4 : réaction d'accepteur ; 5 : formation d'isomères de saccharose. 28

Figure 2 : Représentation schématique des différentes structures des glucanes produits par les glucane-saccharases GH70. 1 : dextrane (résidus de la chaîne principale liés majoritairement en $\alpha(1\rightarrow6)$) ; 2 : reuterane (résidus de la chaîne principale liés en $\alpha(1\rightarrow4)$ et $\alpha(1\rightarrow6)$ avec des ramifications en $\alpha(1\rightarrow6)$) ; 3 : mutane (résidus de la chaîne principale liés majoritairement en $\alpha(1\rightarrow3)$) et 4 : alternane (résidus de la chaîne principale liés alternativement en $\alpha(1\rightarrow3)$ et $\alpha(1\rightarrow6)$). 29

Figure 3 : Structure tridimensionnelle de l'amylosaccharase de *N. polysaccharea* (code PDB : 1G5A). Légende des domaines: N (bleu sombre), A (bleu clair), B (jaune), B' (magenta) et C (rouge). Les résidus catalytiques Asp286 et Glu328, respectivement localisés à la fin des feuillets β -4 et β -5 dans le domaine A, sont également indiqués. 31

Figure 4 : Représentation des sous-sites de fixation du maltoheptaose au niveau du site catalytique de l'ASNp. A : Représentation selon la nomenclature proposée par Davies *et al.* (Davies and Henrissat, 1995). B : Représentation de l'ASNp en complexe avec le maltoheptaose dans le site actif (Champion, 2008). 32

Figure 5 : Mécanisme catalytique avec rétention de la configuration anomérique du substrat, identifié pour l'ASNp (Skov *et al.*, 2001). 34

Figure 6 : Structure globale des glucane-saccharases GTF180- Δ N et GTF-SI. 1. Structure cristalline de la GTF180- Δ N (code PDB : 3KLK). La sphère rouge représente l'ion calcium présent entre les domaines A et B ; 2. Structure cristalline de la GTF-SI (code PDB : 3AIE), 3. Représentation schématique de la conformation en U de la GTF180- Δ N. Adapté de (Ito *et al.*, 2011; Vujicic-Zagar *et al.*, 2010). 37

Figure 7 : Représentation schématique de la structure secondaire des glucane-saccharases des familles GH13 et GH70. Le domaine catalytique des amylosaccharases (GH13) est organisé en tonneau $(\beta/\alpha)_8$ correspondant à une succession alternée de huit feuillets β et hélices α . En ce qui concerne les glucane-saccharases de la famille GH70, le tonneau commence avec l'hélice α_3 et se termine avec le feuillet β_3 . De ce fait, le segment regroupant les éléments $\beta_1 \alpha_1 \beta_2 \alpha_2 \beta_3$ est situé à l'extrémité C-terminale du domaine catalytique. A : domaine A ; B et B' : domaines B et B' ; C : domaine C ; N-ter : Domaine N-terminal (comprenant les domaines IV et V chez les GH70) ; C-ter : Domaine C-terminal (comprenant les domaines IV et V chez les GH70). 38

Figure 8 : Motifs conservés entre les enzymes de la famille GH13 et GH70. En rouge : résidus strictement conservés correspondant à la diade catalytique et au stabilisateur de l'intermédiaire covalent. En noir et gras : autres résidus conservés entre les deux familles d'enzymes. En bleu : résidus formant le sous-site -1 de la GTF180- Δ N (Vujicic-Zagar *et al.*, 2010). 38

Figure 9 : Vues stéréo de la GTF180- Δ N en complexe avec du saccharose et du maltose. 1. Saccharose lié aux sous-sites -1 et +1 du mutant inactif D1025N de la GTF180- Δ N. Les résidus en gras

correspondent aux résidus strictement conservés entre les glucane-saccharases des familles GH13 et GH70. 2. Maltose M1 lié aux sous-sites +1 et +2 et Maltose M2 lié aux sous-sites +2 et +3. Les résidus du domaine A sont représentés en bleu. Les résidus du domaine B sont représentés en vert. La sphère magenta représente l'ion calcium. Les sphères rouges représentent les molécules d'eau. Les pointillés noirs représentent les liaisons hydrogène. (Codes PDB : 3HZ3 et 3KLL)..... 39

Figure 10 : Représentation schématique des interactions (A) acarbose-GTF-SI et (B) maltose-GTF-SI. Les résidus de la diade catalytique et le stabilisateur de l'intermédiaire covalent sont représentés en gras. Pointillés verts et gras : liaisons hydrogènes ; pointillés rouges et gras : interactions hydrophobes ; pointillés verts : interactions hydrophiles faibles. Tiré de (Ito *et al.*, 2011) 40

Figure 11 : Principe de l'évolution dirigée. 53

Figure 12 : Techniques de DNA et family shuffling (Cramer *et al.*, 1998; Stemmer, 1994a). 55

Figure 13 : Techniques de recombinaison non homologue ITCHY et SHIPREC (Ostermeier *et al.*, 1999; Sieber *et al.*, 2001). La première étape de la méthode ITCHY consiste à tronquer indépendamment les deux gènes parentaux en N- et C-terminal respectivement. Les fragments sont ensuite ligués pour aboutir à un mélange de séquences hybrides. La méthode SHIPREC commence par la fusion des deux gènes parentaux suivie d'une digestion contrôlée par la DNaseI. Les fragments sont ensuite circularisés et digérés au niveau du point de fusion originel aboutissant à une inversion de séquence générant la banque de mutants hybrides. Dans les deux méthodes, une dernière étape consiste à purifier les fragments dont la taille est identique à celle des gènes parents. 56

Figure 14 : Technique d'incorporation d'oligonucléotides dégénérés par réassemblage de gène ISOR (Herman and Tawfik, 2007). Dans un premier temps le gène codant pour la protéine d'intérêt est amplifié afin d'augmenter le nombre de copies. Cette étape est suivie d'une digestion par la DNaseI générant divers fragments couvrant l'ensemble du gène initial. Les fragments sont ensuite mélangés aux oligonucléotides dégénérés (la dégénérescence est portée au centre de l'oligonucléotide, flanquée de part et d'autre de séquences correspondant au gène parent). Pour finir, les oligonucléotides dégénérés sont incorporés lors de réassemblage des fragments grâce à une dernière étape de PCR sans amorces..... 58

Figure 15 : Stratégie générale de sélection par complémentation. (A) Sélection par complémentation génétique sur substrat non métabolisable. Dans la bactérie, l'enzyme d'intérêt convertit le substrat non métabolisable en substrat métabolisable lui fournissant ainsi l'énergie nécessaire pour croître. (B) Sélection par complémentation chimique. L'enzyme d'intérêt convertit le substrat en produit qui forme un complexe actif avec un facteur de transcription favorisant l'expression d'un gène essentiel pour la croissance bactérienne (gène de résistance par exemple) porté par un vecteur rapporteur. Adapté de (van Sin Fiet *et al.*, 2006)..... 61

Figure 16 : Stratégie de sélection par phage display adapté de (Fernandez-Gacio *et al.*, 2003). (A) Sélection indirecte. Le variant enzymatique présenté à la surface du phage est sélectionné par affinité pour un ligand spécifique attaché sur un support solide. (B) Sélection directe. Le substrat est lié au complexe enzyme-calmoduline-phage via un peptide de liaison à la calmoduline. Après catalyse, le produit se fixe par affinité à un ligand attaché sur support solide, d'après (Demartis *et al.*, 1999). (C) Sélection directe de métalloenzyme par phage-display. L'enzyme est dans un premier temps inactivée par extraction du co-facteur métallique Zn^{2+} et se lie avec le substrat attaché sur support solide. L'ajout du co-facteur entraîne la conversion du substrat en produit et conduit à l'élution du phage, d'après (Ponsard *et al.*, 2001).....59

Figure 17 : Schéma général de la stratégie de sélection par ribosome display adapté de (Boersma *et al.*, 2007). Dans un système *in vitro*, l'ADN double brin codant pour la protéine d'intérêt est transcrit en ARNm. L'ARNm étant dépourvu de codon stop, le relargage du variant enzymatique lors de la traduction est alors empêché ce qui aboutit à la formation du complexe ARNm-ribosome-protéine. Une sélection est ensuite réalisée par capture du complexe par affinité de la protéine pour un ligand. Après dissociation du complexe, l'ARNm peut être amplifié par RT-PCR isolant ainsi la séquence codante pour le variant sélectionné. 63

Figure 18 : Stratégies de sélection par compartimentation *in vitro*. (A) Couplage direct entre le gène et le produit de la réaction enzymatique utilisé pour la sélection de méthyltransférases (Cohen *et al.*, 2004). 1 : la banque de gènes et le mélange réactionnel permettant la transcription et traduction sont dispersés dans une émulsion eau dans huile (un gène par émulsion) ; 2 : chaque gène est transcrit et traduit ; 3 : l'enzyme active convertit le substrat en produit qui reste attaché au gène ; 4 : les émulsions sont cassées et les compartiments d'eau sont combinés ; 5 : les gènes codant pour des mutants enzymatiques actifs sont sélectionnés pour être caractérisés ou ; 6 : liés à un nouveau substrat pour d'autres tours de sélection (enrichissement de la banque de variants positifs). (B) Analyse par cytométrie de double émulsions employée pour la sélection de nouvelles β -galactosidases (Mastrobattista *et al.*, 2005). 1 : 1 : dispersion de la banque de gènes et le mélange réactionnel dans une émulsion eau dans huile ; 2 : transcription et traduction des gènes à l'intérieur de la microgouttelette ; 3 : transformation du substrat non fluorescent en produit fluorescent par les variants actifs et conversion de l'émulsion eau dans huile en double émulsion ; 4 : séparation des microgouttelettes fluorescentes des non-fluorescentes par FACS ; 5 : récupération des gènes codant pour une enzyme active et amplification par PCR ; 6 : enrichissement de la banque de variants positifs. 64

Figure 19 : Représentation schématique du (A) dextrane et (B) reuterane synthétisés par les glucane-saccharases GTF180 et GTF-A, respectivement. Adapté de (van Leeuwen *et al.*, 2008b; Van Leeuwen *et al.*, 2008c). 78

Figure 20 : Représentation schématique du dextrane synthétisé par la souche *Weissella confusa*. Les liaisons en rouge correspondent aux liaisons $\alpha(1\rightarrow3)$. Adapté de (Maina *et al.*, 2011). 78

Figure 21 : Modèle structural du dextrane (polymorphe haute température) suivant 2 plans de vue différents. Tiré de (Guizard, 1981) 79

Figure 22 : Principe de la séparation par AFFFF. Représentation schématique du système de séparation AFFFF. A l'intérieur de la cellule les échantillons sont sous l'influence de deux vecteurs de flux qui les transportent vers le détecteur (channel flow) et vers la paroi d'accumulation (crossflow). La séparation dans le temps est dépendante de la position de l'échantillon dans la cellule, elle-même déterminée par le coefficient de diffusion de chaque échantillon. Tiré de (Yohannes *et al.*, 2011) 81

Figure 23 : Elugramme (réponse de diffusion de lumière (LS) à l'angle 90° et réponse de l'indice de réfraction (RI)) et masses molaires en fonction du temps d'élution pour une solution d'un dextrane commercial de 150 kDa solubilisé dans une solution de nitrate de sodium à 0,1 M. L'analyse a été réalisée à 60°C . Les cercles représentent la réponse de l'indice de réfraction, les carrés la réponse de diffusion de lumière. Tiré de (Viebke and Williams, 2000). 83

Figure 24 : Courbes de R_G en fonction de M_w de divers biopolymères solubilisés dans une solution d'azide de sodium concentrée à $0,2 \text{ g}\cdot\text{L}^{-1}$. L'analyse a été réalisée à 40°C . (\square) Dextranes commerciaux

D4 (483.1 kDa) et D5 (2 906 kDa), (×) pullulane, (○) amylose et (Δ) glycogène de foie de lapin, huitre et phytoglycogène. Tiré de (Rolland-Sabaté *et al.*, 2008). 84

Figure 25 : Viscosité spécifique (η_{sp}) en fonction de la concentration réduite ($C[\eta]$) pour trois dextrans commerciaux de différentes masses molaires : $M_w = 2.10^6$ (■), (●) 2.10^5 et (▲) 5.10^5 g.mol⁻¹. Adapté de (Pinder *et al.*, 2006). 85

Figure 26 : Corrélation entre la viscosité et la concentration de différentes fractions de dextrane, à 25°C (Heinze *et al.*, 2006). 86

Figure 27 : (A) Evolution des modules élastique (G') et visqueux (G'') dans le cas (1) d'un gel, (2) d'une solution viscoélastique et (3) d'une solution diluée. (B) Mesures en mode dynamique dans le domaine linéaire (oscillation entre 0 à 10 Pa) pour des solutions de dextrans synthétisés par Thio-DSR-S vardel Δ4N-His, variant tronqué de DSR-S (voir partie I, paragraphe I.3.5), à partir d'extraits cellulaires (○ (i) et ● (ii); $G' < G''$ comportement typique d'une solution macromoléculaire; solution concentrée à 5 %) et à partir de l'enzyme pure (□ (iv) and ■ (iii); $G' > G''$ comportement d'un gel; solution concentrée à 0,4 %) (Moulis *et al.*, 2006a). 87

Figure 28 : Thermogramme obtenu par analyse calorifique différentielle illustrant la transition vitreuse pour un échantillon de dextrane de masse molaire de 6.10^3 g.mol⁻¹ dont la teneur en eau est de 8,3 % (g d'eau/100g de masse sèche). Adapté de (Gloria Hernández *et al.*, 2011). 88

Figure 29 : Effet de la masse molaire et de la teneur en eau sur la T_g . Tiré de (Icoz *et al.*, 2005). 89

Figure 30 : Outline of NMR-based screening for carbohydrate active enzyme specificity. The method was set-up with *E. coli* cells encoding empty plasmid (negative control) and parental glucansucrases with known specificities. Cells were cultured in 96-deepwells format for growth and gene expression. After lysis, recombinant enzymes were incubated with 292 mM sucrose and 50 mM iso-maltooligosaccharides to produce oligosaccharides. (trimethylsilyl)propionic acid-2,2,3,3-d4 acid (37 mM), used as internal standard, was added to each samples before 1D ¹H NMR analysis carried out on the crude reaction media. The method was validated by screening a library of glucansucrase variants. Mutants were first selected on their abilities to cleave sucrose. Then, for each of the 4,032 selected active glucansucrase variants, the synthesized oligosaccharide structure was screened at a flow rate of one sample per 3 minutes. 126

Figure 31 : A. "Structural reporter group" used to screen the carbohydrate structures synthesized by the glucansucrases DSR-S vardel Δ4N, DSR-E, ASR C-del, and by the DSR-S vardel Δ4N H463R/T464D/S512T variant by 1D ¹H NMR. Spectra are centered at the anomeric region (4.94 to 5.48 ppm). Peaks were referenced to internal TSP-d4 (1H = 0 ppm). (a) α-D-Glcp unit of the sucrose, (b) α-D-Glcp reducing unit, (b') free D-Glcp unit resulting from sucrose hydrolysis reaction and (c) α(1→5)-D-Glcp unit of the leucrose resulting from sucrose isomer synthesis. B. Examples of glucooligosaccharides synthesized from (1) sucrose and (2) isomaltooligosaccharides by (3) DSR-S vardel Δ4N, (4) DSR-E, (5) ASR C-del and (6) DSR-S vardel Δ4N H463R/T464D/S512T variant (structural representations show the osidic linkage composition of the products), n corresponds to the number of α-D-Glcp unit ranging from 3 to 5. 125

Figure 32 : Identification of glucansucrase mutants with modified specificity by NMR-based screening. a) Ratio of α(1→3) to α(1→6) linked D-Glcp units in the glucooligosaccharides produced by the 303 isolated variants compared to DSR-S vardel Δ4N (black dotted line). These results are also presented

in details in supplementary table 3. b) Comparison of synthesized oligosaccharides primary structures determined after screening using flow injection technology (sc) and large-scale characterization (ch) for DSR-S vardel $\Delta 4N$, c) for mutant F353W, d) for mutant S512C and e) for mutant H463R/T464D/S512T.

..... 130

Figure 33 : Superimposition of 1D 1H NMR spectra of the oligosaccharides synthesized by DSR-S vardel $\Delta 4N$ (black) and the three mutants S512C (blue), F353W (orange) and H463R/T464D/S512T (gray). These mutants show an alteration in their capacity to synthesize $\alpha(1\rightarrow 3)$ glucosidic linkages, visible as disturbance of the relative intensities of $\alpha(1\rightarrow 3)$ linked D-Glcp units anomeric proton signals at 5.33 ppm. Spectra were calibrated according to the internal standard TSP-d4 present in each sample at the same concentration ($^1H = 0$ ppm, 37 mM). (a) α -D-Glcp reducing unit, (b) free D-Glcp unit and (c) α -D-Glcp (1 \rightarrow 5)-D-frup (leucrose). 131

Figure 34 : HPAEC-PAD profiles of the reaction products obtained from sucrose and iso-maltooligosaccharides by using parental DSR-S vardel $\Delta 4N$ and variants S512C, F353W and H463R/T464D/S512T. IMOS: Isomalto-oligosaccharides; DP: Polymerisation degree (referenced to IMOS used as acceptors). 132

Figure 35 : Superimposition of 1D 1H NMR spectra of the α -glucans synthesized by DSR-S vardel $\Delta 4N$ (black) and the mutants S512C (blue), F353W (orange) and H463R/T464D/S512T (gray) that show an altered specificity. Spectra were calibrated according to the internal standard TSP-d4 present in each sample at the same concentration ($^1H = 0$ ppm, 37 mM). (a) α -D-Glcp reducing unit, (b) α -D-Glcp (1 \rightarrow 5)-D-frup. 132

Figure 36 : Alignment of GH70 amino acid sequences in the regions chosen for the combinatorial site directed mutagenesis of DSR-S vardel $\Delta 4N$. Sequences used are clustered according to the specificity of the characterized glucansucrases. Dextransucrase sequences are ordered from the higher to lower ratio between the $\alpha(1\rightarrow 6)$ and the $\alpha(1\rightarrow 3)$ linkages contained in the α -glucans produced by these enzymes. Mutansucrase sequences are ordered from the lower ratio between the $\alpha(1\rightarrow 3)$ and the $\alpha(1\rightarrow 6)$ linkages contained in the produced α -glucans. Reuteransucrase sequences are ordered from the higher to lower ratio between the $\alpha(1\rightarrow 4)$ and the $\alpha(1\rightarrow 6)$ linkages. aa: strictly conserved amino acids; ★: catalytic amino acids (D400 and E438, DSR-S vardel $\Delta 4N$ numbering) and transition state stabilizer (D511, DSR-S vardel $\Delta 4N$ numbering); •: amino acids chosen for mutagenesis. 151

Figure 37 : HPSEC chromatogram of products synthesized by *Thio*-DSR-S vardel $\Delta 4N$ -His and the 7 selected variants, from 292 mM sucrose. HMW: High Molecular Weight polymers; IMW: Intermediary Molecular Weight products; DP2: disaccharides; DP1: monosaccharides. Red dotted lines indicate the retention time of standard dextrans. 154

Figure 38 : (A) 3D-model of DSR-S vardel $\Delta 4N$ catalytic domain. (B) Catalytic domain of *Lactobacillus reuteri* N-terminally truncated glucansucrase GTF180 (PDB accession code: 3KLLK). 156

Figure 39 : (A) Molecular docking of sucrose in the active site of DSR-S vardel $\Delta 4N$. (B) Overview of mutation sites affecting DSR-S vardel $\Delta 4N$ linkage specificity. Docked sucrose molecule is shown for reference. Targeted residues are showed in orange and green for LibA and LibB respectively. 157

Figure 40 : Molecular docking of α -D-Glcp-(1 \rightarrow 6)[α -D-Glcp-(1 \rightarrow 3)] α -D-Glcp-(1 \rightarrow 6)-D-Glcp tetrasaccharide in the active site of D460M/H463Y/T464M/S512C mutant. 158

SI Figure 1 : (a) Western blot on unpurified enzymes using anti-thioredoxin antibody. (b) Western blot on unpurified enzymes using anti-6 x his. (c) Zymograms on unpurified enzymes. (d) SDS-PAGE gel and colloidal blue coloration on purified enzymes. (e) Western blot on purified enzymes using anti-6 x his antibody. 1, *Thio-DSR-S vardel Δ4N-His*; 2, *Thio-F353W-His*; 3, *Thio-F353T-His*; 4, *Thio-D306G/F353L/N404G/W440T-His*; 5, *Thio-S512C-His*; 6, *Thio-H463R/T464D/S512T-His*; 7, *Thio-D460M/H463Y/T464N/S512C-His*; 8, *Thio-H463R/T464V/S512T-His* and 9, *Thio-D460A/H463S/T464L-His*..... 170

Figure 41 : (A) ¹H NMR spectra of the α-glucans synthesized by the seven selected variants. (B) Osidic linkage content of the α-glucans produced by the parental enzyme DSR-S vardel Δ4N and the seven selected mutants..... 186

Figure 42 : (A) and (B) Elugrams (Light Scattering at 90° and DRI responses) of α-glucans versus the elution volume. (C), (D) and (E) Elugrams (Relative Light Scattering at 90° responses) of α-glucans and their molar masses versus the elution volume. (F) Elugrams (Relative Light Scattering at 90° and DRI responses) of Gcn4 and Gcn5 versus the elution volume. The red arrows indicate the VLMM (very low molecular mass), LMM (low molecular mass) and HMM (high molecular mass) observed for Gcn4. .. 188

Figure 43 : Logarithm of the radii of gyration (R_{Gi}) versus the logarithm of the molar masses (M_i) of the various α-glucans. 189

Figure 44 : (A) Apparent viscosity as a function of the shear rate of the 5% (w/v) polymer solutions. (B), (C) and (D) Storage modulus (G') and loss modulus (G'') as a function of angular frequency of the 5% (w/v) polymer solutions. 192

Figure 45 : Differential scanning calorimetry thermograms for polymers (a) Gcn_DvΔ4N, (b) Gcn4, (c) Gcn3 (d) Gcn2 (e) Gcn6, (f) Gcn5 and (g) Gcn1..... 194

Figure 46 : (A) Films obtained after casting and conditioning at a_w 0.57. (B) Tensile stress-elongation curve of Gcn1 films. 195

TABLEAUX

Tableau 1 : Liste des glucane-saccharases caractérisées en janvier 2011.....	30
Tableau 2 : Ingénierie de la spécificité des glucane-saccharases de la famille GH70 : mutations et effets.	46
Tableau 3 : Exemples de méthodes de criblage utilisées et taille des banques analysées pour l'évolution dirigée d'enzymes.	67
Tableau 4 : Attribution des déplacements chimiques par RMN ¹ H pour différents glucanes solubilisés dans du D ₂ O à 27°C (Van Leeuwen <i>et al.</i> , 2008d).	77
Tableau 5 : Masse molaire de certains α-glucanes trouvée dans la littérature.....	80
Tableau 6 : Relation entre la masse molaire et le rayon de giration pour des dextrans commerciaux.	84
Table 7 : ¹ H NMR chemical shifts (in ppm) of the synthesized products.....	128
Table 8 : Parameters for sample train preparation and automated direct injection of crude reaction media to the NMR spectrometer.	135
Table 9 : NMR measurement parameters.	135
Tableau 10 : Similarité de séquences entre les gènes codant pour les glucane-saccharases DSR-S, DSR-E, GTF-I et ASR.	145
Table 11 : Sequence analysis of the LibA and LibB variants and evolution of the mutation distribution at each step of the screening strategy.	153
Table 12 : Products synthesized by the parental <i>Thio</i> -DSR-S varde1 Δ4N- <i>His</i> enzyme and the seven selected variants, from 292 mM sucrose, as determined by HPSEC and 1D ¹ H NMR spectroscopy... ..	154
Table 13 : Linkage specificity and kinetic properties of the 8 selected variants, in comparison with those of the WT <i>Thio</i> -DSR-S varde1 Δ4N- <i>His</i>	155
SI Table 1 : Primers and degenerated oligonucleotides used for combinatorial site-directed mutagenesis of <i>dsr-s</i> varde1 Δ4N. Underlined nucleotides correspond to the introduced restriction sites.....	167
SI Table 2 : Mutants from LibA and LibB with altered specificity, identified by NMR-based screening.	168
Table 14 : AFFFF-MALLS analysis of α-glucans produced by the parental enzyme DSR-S varde1 Δ4N and its mutants. (\bar{M}_w) weight average Molar Masses, (\bar{R}_G) z-average Radii of Gyration, apparent density (d_{Gapp}) and the slope of the log-log plot of the radius of gyration versus the molar mass (v_G).	187
Table 15 : Water content, glass transition temperature (T_g) and changes in heat capacity at T_g	193
Table 16 : Mechanical properties of Gcn1 in comparison with those of other polymers referred in the literature.	196