

THÈSE

présentée devant

L'INSTITUT NATIONAL DES SCIENCES APPLIQUÉES DE TOULOUSE

en vue de l'obtention du

DOCTORAT

N° d'ordre : 931

Spécialité: Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries

Filière : Microbiologie et Biocatalyse industrielle

par **Elise CHAMPION**

Ingénieur de l'Ecole Nationale Supérieure de Chimie de Montpellier

GLUCANE-SACCHARASES OPTIMISÉES

POUR LA GLUCOSYLATION DE MOLÉCULES

INTERVENANT DANS LA COMPOSITION DE VACCINS

Soutenue le 1^{er} Avril 2008 devant la commission d'examen :

| | |
|----------------------|---|
| Bruno CANARD | Directeur de Recherche CNRS, Marseille |
| Philippe MINARD | Professeur, PARIS-SUD 11 |
| Antoni PLANAS | Professeur, Institut Quimic de Sarria, Barcelone, Espagne |
| <u>Pierre MONSAN</u> | Professeur, INSA Toulouse |
| Magali REMAUD-SIMÉON | Professeur, INSA Toulouse |
| Isabelle ANDRÉ | Chargée de Recherche CNRS, INSA Toulouse |
| Laurence MULARD | Directrice de Recherche, Institut PASTEUR, Paris |

Cette thèse a été préparée au Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés de l'INSA de Toulouse (UMR CNRS 5504, UMR INRA 792), dans le cadre de l'Ecole Doctorale SEVAB.

NOM : CHAMPION

Prénom : Elise

Titre : Glucane-saccharases optimisées pour la glucosylation de molécules intervenant dans la composition de vaccins

Spécialité: Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries, Filière : Microbiologie et Biocatalyse industrielle

Année : 2008

Lieu : INSA de Toulouse

RESUME :

La synthèse d'oligosides complexes et glycoconjugués reste difficilement réalisable par voie chimique, freinant ainsi l'exploitation de ces molécules à grande échelle. Le recours aux catalyseurs enzymatiques apparaît comme une alternative prometteuse mais les enzymes naturelles ne présentent pas toujours les propriétés adéquates, en particulier sur le plan de leurs spécificités, pour être intégrées dans des voies de synthèse chimio-enzymatiques. Le défi de cette thèse a consisté à mettre à profit les techniques d'ingénierie enzymatique pour concevoir des glyco-enzymes déployant de nouvelles spécificités, produire à façon des intermédiaires de synthèse chimique et ouvrir la voie à de nouveaux procédés d'obtention d'oligosides d'intérêt thérapeutique. Il s'agissait de développer de nouvelles transglucosidases capables de glucosyler des dérivés de la *N*-acétyl-D-glucosamine et du L-rhamnose pour produire des disaccharides n'existant pas dans la nature mais pouvant entrer dans une voie de synthèse chimique destinées à synthétiser des haptènes oligosaccharidiques mimant les motifs antigéniques des lipopolysaccharides des sérotypes 1b et 3a de *Shigella flexneri*. A terme, l'objectif est de développer un vaccin multivalent de troisième génération contre ces pathogènes, responsables du décès de milliers de personnes chaque année dans le monde.

Pour relever ce défi, nous avons tout d'abord évalué le potentiel de quatre glucane-saccharases sauvages, spécifiques de la synthèse de liaisons osidiques distinctes, à glucosyler les accepteurs cibles. Les résultats obtenus ont montré qu'aucune des enzymes testées ne catalysait efficacement la réaction attendue. Cette étude préliminaire a cependant permis de sélectionner l'amylosaccharase de *Neisseria polysaccharea*, comme enzyme candidate au développement de biocatalyseurs à stéréo- et régio-spécificités contrôlées. Le remodelage du catalyseur a été guidé par l'analyse des modèles des complexes enzyme:accepteur. Sept acides aminés situés dans le site de fixation de l'accepteur (sous-site +1) ont été identifiés et individuellement remplacés par les 19 acides aminés possibles par mutagenèse dirigée. Une banque de 133 mutants a été créée qui renfermait 43 mutants d'intérêt pour les réactions de glucosylation considérées. Notamment, le mutant I228Y a révélé une toute nouvelle régio-spécificité vis-à-vis de l' α -methyl L-rhamnopyranose. Il s'agit du premier exemple de création de nouvelle spécificité pour les enzymes de cette famille et plus largement pour des α -transglycosidases. Le mutant F290K, également isolé dans cette première librairie, s'est montré capable de glucosyler l' α -allyl *N*-acétyl-D-glucoamine avec une efficacité catalytique 130 fois plus élevée que celle observée avec l'enzyme sauvage. Ces deux nouvelles enzymes ont été employées pour la synthèse des disaccharides ciblés et la validation de la voie chimio-enzymatique proposée pour la synthèse des haptènes saccharidiques.

Cette stratégie a été poursuivie par la construction de banques de double-mutants permettant de générer une diversité plus importante, focalisée sur ces positions, et rechercher des effets synergiques entre les différentes mutations pour améliorer la réaction de glucosylation des accepteurs cibles. Vingt double-mutants ont pu être ainsi identifiés.

MOTS CLES :

Glucane-saccharase, amylosaccharase, glucosylation, réaction d'accepteur, ingénierie semi-rationnelle, *Shigella*, vaccins synthétiques, *N*-acétyl-D-glucosamine et L-rhamnose.

Cette thèse a été préparée au Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés de l'INSA de Toulouse (UMR CNRS 5504, UMR INRA 792), et soutenue le 1^{er} Avril 2008 à l'INSA de Toulouse.

SUMMARY:

The difficult access to complex carbohydrates and glycoconjugates by chemical synthesis impairs their development on a large scale. Therefore, the use of biocatalysts appears as an appealing alternative, yet poorly exploited in spite of the fast-growing development of engineering technologies allowing the search and the construction of better performing enzymes, displaying novel specificities.

The objective of this thesis aimed to apply enzyme engineering techniques to conceive novel glyco-enzymes well-adapted for the glucosylation of naturally non- or poorly recognized molecules (*N*-acetyl-D-glucosamine and L-rhamnose derivatives fulfilling the subsequent chemical step requisites). The glucosylated motives obtained *via* the enzymatic route, and corresponding to the serotype-specific branched α -D-glucopyranosyl linkages of *Shigella flexneri* 2a, 1b and 3a, are subsequently elongated through a chemical synthetic pathway to synthesize the antigenic oligosaccharides. This chemo-enzymatic road could open the way to oligosaccharide haptens at low cost, that could in the future be used in the development of a third-generation vaccine against *S. flexneri* 1b and 3a.

In a first study, we have evaluated the potential of four wild-type glucansucrases, specific for the synthesis of distinct osidic linkages, to glucosylate both types of targeted acceptors. Our preliminary results demonstrated that none of the tested glucansucrases allowed an efficient glucosylation of the target acceptors with the desired linkage specificity.

On the basis of these results, amylosucrase from *Neisseria polysaccharea* was chosen as a candidate enzyme for the development of biocatalysts with controlled stereo- and regio-specificities that could enable the glucosylation of both types of acceptors of interest. A semi-rational approach based on the engineering of the acceptor binding site (subsite +1) was undertaken to evolve this enzyme. By individually exploring 7 positions of the active site using site-directed mutagenesis, the strategy led to the isolation of 47 mutants of interest for the considered glucosylation reactions (over the 133 generated mono-mutants) and to the identification of 2 key positions (228 and 290). Noteworthy, I228Y mutant displayed an entirely novel regio-specificity towards the α -methyl L-rhamnopyranose and in the presence of allyl *N*-acetyl-D-glucosamine, F290K mutant permitted a 130 fold enhancement of catalytic efficiency compared to wild-type amylosucrase.

This strategy was further pursued with the construction of double-mutant libraries enabling the generation of an enlarged diversity, focused on the identified positions, to investigate the synergistic effects between the different mutations in order to improve the glucosylation reactions of the target acceptors. Twenty double-mutants were thus identified.

KEY WORDS :

Glucansucrase, amylosucrase, glucosylation, acceptor reaction, semi-rational engineering, *Shigella*, synthetic vaccines, *N*-acetyl-D-glucosamine and L-rhamnose

JURY :

| | |
|----------------------|---|
| Bruno CANARD | Directeur de Recherche CNRS, Marseille/ Rapporteur |
| Philippe MINARD | Professeur, PARIS-SUD 11/ Rapporteur |
| Antoni PLANAS | Professeur, Institut Quimic de Sarria, Barcelone, Espagne/ Rapporteur |
| Pierre MONSAN | Professeur, INSA Toulouse/ Directeur de thèse |
| Magali REMAUD-SIMÉON | Professeur, INSA Toulouse/ Présidente du jury |
| Isabelle ANDRÉ | Chargée de Recherche CNRS, INSA Toulouse/ Invitée |
| Laurence MULARD | Directrice de Recherche, Institut PASTEUR, Paris/ Invitée |

Au cours de ces presque quatre années de thèse, mes deux plus belles découvertes ont été mon engouement pour la recherche appliquée et ma passion pour la montagne pyrénéenne. Recherche, plutôt la semaine et montagne, plutôt le week-end. Aussi, je ne peux résister à l'idée de les associer symboliquement dans un p'tit récit de thèse et dans les remerciements qui l'accompagnent....

Alors que je m'appête à partir sur la HRP qui traverse les Pyrénées de l'Atlantique à la Méditerranée, je me rends compte que ce voyage-là, je l'ai déjà un peu vécu :

Nous sommes parties, mon enzyme et moi, un jour d'octobre 2004, depuis un milieu aqueux et riche en NaCl certes, mais drôlement agité : l'Océan Atlantique. Direction : la mer Méditerranée, une destination de rêve pour que mon enzyme puisse *Rhamer* et *NAGer* un peu plus à sa guise. Un pari dingue, cette transpyrénéenne de l'AS !

Je me suis, donc mise en route, avec mon enzyme, pleine d'enthousiasme pour cette aventure qui s'annonçait passionnante. Une foule d'itinéraires possibles germaient déjà dans ma tête. Itinéraires qui se sont modifiés, affinés, remodifiés, grâce notamment aux précieux conseils que j'ai pu recevoir au fil du parcours. Parfois, nous nous sommes engagées dans une voie sans issue et nous avons du renoncer pour mieux changer de route. D'autres fois, ce sont les curiosités des lieux qui nous ont fait dévier du chemin. Au cours de cette traversée, il y a eu des moments éprouvants, de belles tempêtes, que j'ai pu essuyer en me réfugiant dans les Kfêts qui jalonnaient le parcours. Un lieu de ravitaillement inédit, avec des gâteaux au goût indescriptible. Il y eu aussi ces moments de doute, qui s'effaçaient lors de rencontres avec d'autres « accompagnateurs d'enzyme ». Certains faisaient route en sens inverse, parce que leur enzyme avait envie de surfer. Et puis il y a eu ces moments intenses, lorsque de magnifiques pics chromatographiques se déployaient devant mes yeux. On pouvait apercevoir, au loin, la mer Méditerranée!!!

Ces quatre années ont été une succession d'étapes riches en expériences et en émotions, une alternance de passage entre vallées et sommets. Une période formidablement enrichissante, aussi bien scientifiquement que personnellement. Aussi je voulais remercier toutes les personnes qui m'ont aidé à réaliser cette belle traversée, à tracer mon itinéraire, à franchir cols et sommets pour arriver à destination avec l'AS.

Je tiens tout d'abord à remercier Bruno Canard, Philippe Minard et Antoni Planas pour avoir accepté d'être rapporteurs de cette thèse. Merci pour votre enthousiasme pour le manuscrit. Un merci tout particulier à Antoni pour m'avoir fait découvrir le monde des enzymes espagnoles, pour m'avoir présenté à mon futur directeur de thèse et pour tes nombreux conseils lors de ton séjour à Toulouse.

Merci à Pierre, mon directeur de thèse, pour m'avoir accueillie dans l'équipe « Enzymo », pour m'avoir proposé ce sujet de thèse et avoir été présent dans les étapes importantes. Merci, Pierre, pour votre confiance, votre dynamisme et votre bonne humeur quotidienne.

Merci à Magali, Isabelle et Sandrine pour m'avoir encadré pendant cette thèse et pour avoir contribué, chacune à votre manière à l'avancée de cette thèse. Merci d'avoir partagé ces beaux moments, et d'avoir été compréhensives quand le ciel était voilé.

Merci Magali, pour ton enthousiasme scientifique contagieux et pour y introduire tout plein d'humanité. Merci de toujours y croire quand, moi, j'ai plus vraiment envie d'y croire...

Merci Isa, pour m'avoir montré le début de la route semi-rationnelle. Sans toi, j'en serais encore à essayer de modéliser ces sucres autour du Mont Perdu. Merci pour ton précieux soutien, pour ta disponibilité de tous les instants et pour tout ce que tu m'as appris tout au long du chemin. Merci 1000 fois !

Merci Sandrine, pour ton immense aide pour l'attribution de tous ces pics qui formaient les massifs de **Ramougn**, **Maladeta** et **Néouvielle**. Merci pour ta patience.

Merci à toute l'équipe « Enzymo », qui ont fait du LISBP un endroit où j'ai adoré venir travailler (quasi ☺) tous les jours pendant ma thèse. Merci pour la chaleur humaine qui irrigue cette équipe. Merci pour tous ces instants Kfêts où nous avons refait 1000 fois le labo ...et autant de fois le monde!

Un merci tout particulier à Claire, pour avoir amené l'AS jusqu'en mer rouge, une mer encore plus paisible. Merci d'avoir partagé cette dernière année-paillasse avec moi. C'était tellement mieux d'être à deux. Merci, c'est super agréable de travailler avec toi.

Merci Marie-Sophie et Albert, de m'avoir accompagné sur un petit bout de chemin. Merci pour votre aide, merci pour avoir ouvert d'autres sentiers.

Merci Sandra, pour avoir pris le temps de me montrer tout plein de choses pendant ma thèse et toujours avec le sourire. Merci pour ton écoute et ta bonne humeur qui font du bien !

Merci Stef pour tes conseils éclairés sur les secrets cachés de la Biomol et sur tout ce qui est gogglisable.

Merci Régis et Vincent pour toutes vos petites attentions (...et les chips dans les tiroirs ☺)

Et merci, à mon plus fidèle compagnon de travail, l'HTC, pour n'avoir jamais défailli devant la cadence infernale que tu as du subir !

Ce travail n'aurait certainement pas pris cette dimension sans cette collaboration avec l'Institut Pasteur et le CERMAV et sans toutes ces réunions entre Grenoble, Paris et Toulouse... Merci Laurence, Julien, Karine, Hugues, Carole, Sami, Sylvie, Armelle et Farida pour toutes nos discussions autour du rétro.

Merci, Laurence, pour être à l'initiative de ce très beau projet, qui a donné tout son sens à cette traversée. Merci d'avoir coordonné OPTIGLUC avec autant d'enthousiasme. Merci pour tous les échanges fructueux.

Merci Julien et Karine d'avoir purifié les immondes mélanges de disaccharides qu'on vous a envoyés.

Je voudrais également remercier à tous mes adorables amis pour ces moments privilégiés qu'on a partagés ensemble à Toulouse, en montagne ou aux quatre coins du monde. Merci pour votre fantaisie, votre fraîcheur et votre simplicité. Merci pour nos délires, nos fous-rires, nos moments remonte-moral et notre complicité.

Merci Alice, Cécile, Constantin, Céline, Eme, Estelle, Hélène K., Hélène V., Isa, Jess, Lolotte, Laurette, Nico, Olive, Rémi, Vincent ...Merci pour cette amitié très précieuse.

Merci aussi à toutes les personnes du club Rando de l'UPS avec qui j'ai partagé une, deux... cinquante randos. Merci pour ces fabuleux bols d'air, les belles siestes au soleil, ces moments où tout a plus de saveur !

Merci à mes proches : Davidou, Mandou, Fredou et Binou. Merci, les cousins, pour vos chaleureux et nombreux accueils à Paris.

Et enfin, merci à mes plus proches : Maman, Papa, Ben et Clem. Merci d'être toujours, toujours là à mes côtés.

BREVETS

1- Mutants of Amylosucrase and uses thereof for synthesizing complex oligosaccharides and disaccharide intermediates. Série 3a.

Champion E., André I., Moulis C., Morel S., Boutet J., Monsan P., Remaud-Siméon M., Mulard L.

Demande de brevet en cours- Pasteur DI2007-51.

2- Mutants of Amylosucrase and uses thereof for synthesizing complex oligosaccharides and disaccharide intermediates. Série 1b.

Champion E., André I., Moulis C., Morel S., Descroix K., Monsan P., Remaud-Siméon M., Mulard L.

Demande de brevet en cours- Pasteur DI2007-52.

PUBLICATIONS

1- Enzymatic synthesis of disaccharide motives entering in the composition of bacterial polysaccharides by use of glucansucrases

Champion E., Morel S., Mulard L., Monsan P., Remaud-Siméon M., André I.

Prêt à être soumis.

2- Tailored transglucosidases for innovative chemo-enzymatic synthesis of antigenic oligosaccharides

Champion E., André I., Moulis C., Morel S., Boutet J., Descroix K., Monsan P., Mulard L., Remaud-Siméon M.

Prêt à être soumis.

3- Exploration of amylosucrase sequence-function space for non-natural acceptor glucosylation using a novel screening method

Moulis C., **Champion E.**, André I., Monsan P. and Remaud-Siméon M.

COMMUNICATION ORALE

1- Alteration of transglucosidases acceptor specificities by semi-rational engineering

Champion E., André I., Bounaix M.S., Pizzut-Serin S., Skov L., Morel S., Monsan P., Remaud-Siméon M.

21^{ème} Colloque du Club de Bioconversions en Synthèse Organique, Ax-les-Thermes, France, 29 Mai-1^{er} Juin 2006.

COMMUNICATIONS PAR AFFICHE

1- Characterization of a sucrose hydrolase from *Xanthomonas campestris*

Champion E., Pizzut-Serin S., Monchois V., Monsan P., Remaud-Siméon M.

6th Carbohydrate Bioengineering Meeting, Barcelona, 3-6 April 2005

2- Modelage de la saccharose-hydrolase de *Xanthomonas Campestris* adaptée à la synthèse de glucoconjugués

Champion E., André I., Pizzut-Serin S., Skov L., Morel S., Monsan P., Remaud-Siméon M.

21^{ème} Journées du Groupe Français des Glucides, Le Croisic, France, 15-18 Mai, 2006

3- Alteration of transglucosidases acceptor specificities by semi-rational engineering

Champion E., André I., Bounaix M.S., Pizzut-Serin S., Skov L., Morel S., Monsan P., Remaud-Siméon M.

21ème Colloque du Club de Bioconversions en Synthèse Organique, Ax-les-Thermes (France), 29 Mai-1^{er} Juin 2006

4- Transglucosylation reactions catalysed by amylosucrase from *Neisseria Polysaccharea*

Bounaix M.S., **Champion E.**, Morel S., André I., Monsan P., Remaud-Siméon M.

21ème Colloque du Club de Bioconversions en Synthèse Organique, Ax-les-Thermes (France), 29 Mai-1^{er} Juin 2006

5- Screening of unnatural acceptors for the specificity study of a transglucosidase, amylosucrase from *Neisseria polysaccharea*.

Morel S., **Champion E.**, André I., Monsan P., Remaud-Siméon M.

7th Carbohydrate Bioengineering Meeting, Braunschweig, Allemagne, 22-25 Avril 2007.

6- Transglucosidase specificity optimization by semi-rational engineering.

Champion E., André I., Pizzut-Serin S., Skov L., Morel S., Monsan P., Remaud-Siméon M.

7th Carbohydrate Bioengineering Meeting, Braunschweig, Allemagne, 22-25 Avril 2007

7- Modifying glucansucrase regioselectivity to synthesize new oligosaccharides and glycoconjugates.

Champion E., Emond S., Bertrand A., André I., Pizzut-Serin S., Skov L., Morel S., Monsan P., Remaud-Siméon M.

Benzon Symposium 54 – Glycosylation : Opportunities in drug development, Copenhagen, Danemark, 11-14 Juin 2007.

8- Molecular engineering of glucansucrases for novel applications

Champion E., Emond S., Bertrand A., Bounaix M.S., Morel S., Albenne C., Potocki-Véronèse G., André I., Monsan P., Remaud-Siméon M.

IX^{ème} Colloque de la section 30 – De la recherche à la découverte, Nice, France, 8-9 Janvier 2008

EMBL Workshop EMBL (European Molecular Biology Laboratory) –Directed Evolution Approaches in Structural Biology, Grenoble, France, 30-31 Janvier 2008

Sommaire

| | |
|--|-----------|
| SOMMAIRE..... | 13 |
| INTRODUCTION..... | 19 |
| CHAPITRE I: ÉTUDE BIBLIOGRAPHIQUE..... | 25 |
| Partie I : Les Vaccins basés sur les Antigènes Saccharidiques..... | 27 |
| A. Des vaccins traditionnels aux nouvelles générations de vaccins | 27 |
| A.1. Vaccins traditionnels | 27 |
| A.2. Nouvelles générations de vaccins | 29 |
| B. Vaccins basés sur les antigènes saccharidiques | 33 |
| B.1. Antigène saccharidique bactérien et potentiel antigénique | 33 |
| B.2. Les Vaccins Composés de PC Purifiés : 1 ^{ère} Génération de vaccins polysidiques..... | 36 |
| B.3. Les Vaccins Conjugués PS-Protéine : 2 ^{nde} génération de vaccins polysidiques..... | 37 |
| B.4. Les Vaccins Semi-Synthétiques Conjugués PS-Protéine : 3 ^{ème} génération de vaccins polysidiques . | 43 |
| Partie II : <i>Shigella</i>, notre modèle d'étude | 48 |
| A. <i>Shigella</i>, agent étiologique de la shigellose | 48 |
| A.1. L'organisme pathogène | 48 |
| A.2. La Maladie..... | 48 |
| A.3. Historique et épidémies causées par la shigellose | 49 |
| B. Mécanisme de l'infection | 51 |
| C. Réponse immunitaire induite au cours de la shigellose | 52 |
| C.1. Réponse innée | 52 |
| C.2. Réponse spécifique | 53 |
| C.2.a. Réponse cellulaire | 53 |
| C.2.b. Réponse humorale..... | 53 |
| D. Avancées des stratégies vaccinales (état de l'art) | 54 |
| D.1. Les souches vivantes de virulence atténuée..... | 55 |
| D.2. Les vaccins de 2 ^{ème} génération..... | 56 |
| D.3. Le recours aux Vaccins Semi-Synthétiques Conjugués PS-Protéine:..... | 57 |
| Une nouvelle approche vaccinale prometteuse..... | 57 |
| Partie III : Les Glyco-enzymes..... | 60 |
| A. Choix des glyco-enzymes | 60 |
| A.1. Les glyco-enzymes impliquées dans la glycosylation du LPS de <i>S. flexneri</i> ?..... | 60 |
| A.2. Vers d'autres glycosyltransférases ?..... | 61 |
| A.3. Une préférence pour les glycoside-hydrolases ... plus particulièrement les glucane-saccharases | 62 |
| B. Les outils enzymatiques : les glucane-saccharases | 63 |
| B.1. Présentation des glucane-saccharases | 64 |
| B.1.a. Biodiversité des glucane-saccharases..... | 64 |
| B.1.b Données structurales | 67 |
| B.1.c Mécanisme catalytique..... | 68 |
| B.1.d Réactions catalysées par les GS | 70 |
| B.2. Les amylosaccharases | 75 |
| B.2.a. Organismes producteurs | 75 |
| B.2.b. Séquences consensus de la famille GH13.4 | 77 |
| B.2.c. Caractéristiques biochimiques de l'ASNP..... | 80 |
| B.2.d. Structure de l'amylosaccharase de <i>Neisseria polysaccharea</i> | 81 |
| C. Techniques d'Ingénierie et Applications..... | 86 |
| C.1. Ingénierie des protéines : Pourquoi ? | 86 |
| C.2. Ingénierie des protéines : Comment ? | 87 |
| C.3. L'ingénierie des glycoside-hydrolases en vue de la modification de leur activité | 91 |
| C.3.a. Les glycosynthases, les thioglycoligases et les thioglycosynthases | 91 |
| C.3.b. Des approches alternatives aux glycosynthases | 96 |

| | |
|--|------------|
| Positionnement du sujet et Strategies | 100 |
| A. Stratégie de synthèse des haptènes 3a et 1b | 100 |
| B. Stratégie de synthèse des disaccharides par voie enzymatique | 101 |
| B.1. Choix des couples glucane-saccharase/accepteur | 101 |
| B.1.a. Choix des glucane-saccharases | 101 |
| B.1.b Choix des accepteurs | 102 |
| B.2. Stratégie employée pour le développement de la voie enzymatique | 103 |
| Références bibliographiques | 104 |

CHAPITRE II : SYNTHÈSE ENZYMATIQUE DE DISACCHARIDES INTERVENANT DANS LA COMPOSITION DE POLYSACCHARIDES BACTÉRIENS125

| | |
|--|------------|
| Enzymatic synthesis of disaccharide motives entering in the composition of bacterial polysaccharides by use of glucansucrases | 128 |
| A. Summary | 128 |
| B. Introduction | 129 |
| C. Results and Discussion | 130 |
| C.1. Glucosylation of <i>N</i> -acetyl-D-glucosamine | 131 |
| C.2. Glucosylation of L-rhamnose | 135 |
| C.3 Glucosylation of methyl α -L-Rhamnopyranoside | 138 |
| D. Conclusion | 140 |
| E. Experimental Section | 140 |
| E.1. Materials | 140 |
| E.1.a. Choice of recombinant glucansucrases | 140 |
| E.1.b. Chemical material | 141 |
| E.2. Acceptor reaction assay | 141 |
| E.3. Glucosyl acceptor production | 141 |
| E.4. Analytical methods | 142 |
| E.4.a. High Performance Liquid Chromatography (HPLC) | 142 |
| E.4.b. High Resolution Mass Spectrometry (HRMS) and Nuclear Magnetic Resonance (NMR) | 142 |
| F. References | 144 |

CHAPITRE III : CONCEPTION A FAÇON DE TRANSGLUCOSIDASES POUR LA SYNTHÈSE CHIMIO-ENZYMATIQUE D'ANTIGÈNES OSIDIQUES147

| | |
|--|------------|
| Tailored transglucosidases for innovative chemo-enzymatic synthesis of antigenic oligosaccharides | 151 |
| A. Summary | 151 |
| B. Introduction | 152 |
| C. Results | 154 |
| C.1. Chemo-enzymatic pathway | 154 |
| C.2. Choice of acceptor substrates | 155 |
| C.3. Screening for native glucansucrases able to synthesize the starting building blocks | 157 |
| C.4. Engineering of amylosucrases able to glucosylate the target acceptors | 157 |
| C.4.a. Structural analysis of AS: α -D-Glcp-(1 \rightarrow 3)- L-RhapOMe [EA'] complex | 158 |
| C.4.b. Structural analysis of AS: α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc [ED] complex | 159 |
| C.5. Detection of mutants able to glucosylate the target acceptors | 160 |
| C.5.a. Target building block α -D-Glcp-(1 \rightarrow 3)- L-RhapOMe [EA'] | 160 |
| C.5.b. Target building block α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc [ED] | 160 |
| C.6. Characterization of I228Y and F290K mutants | 161 |
| C.7. Structural insight on mutation effects | 163 |
| C.8. EA' as starting material for the synthesis of <i>S. flexneri</i> building blocks | 164 |
| C.9. ED' as starting material for the synthesis of <i>S. flexneri</i> building blocks | 164 |
| D. Discussion | 165 |
| E. Methods | 168 |

| | |
|--|------------|
| E.1. Bacterial strains, plasmids and chemicals | 168 |
| E.2. Molecular modelling | 168 |
| E.3. Construction of mutant library | 169 |
| E.4. Expression of mutant library | 171 |
| E.5. Screening of mutant library | 171 |
| E.6. Characterization of the Selected Variants: I228Y and F290K | 172 |
| E.6.a. Production and Purification of I228Y and F290K | 172 |
| E.7. Preparative synthesis of the acceptor reaction product. | 173 |
| F. Supplementary Results..... | 174 |
| F.1. Supplementary Results 1 : Scale-up of the enzymatic processes | 174 |
| F.2. Supplementary Results 2 | 175 |
| F.2.a. Enzymatic glycosylation of L-Rhap and α -L-RhapOAll derivative | 175 |
| F.2.b. Enzymatic glycosylation of D-GlcpNAc derivatives | 175 |
| G. Supplementary Methods | 176 |
| H. Supplementary Figures | 179 |
| I. References | 182 |

CHAPITRE IV : EXPLORATION DU SOUS-SITE (+1) DE L'AMYLOSACCHARASE POUR LA GLUCOSYLATION D'ACCEPTEURS NON-NATURELS A L'AIDE D'UNE NOUVELLE METHODE DE CRIBLAGE.....187

| | |
|--|------------|
| Exploration of amylosucrase sequence-function space for non-natural acceptor glycosylation using a novel screening method | 190 |
| A. Summary..... | 190 |
| B. Introduction | 191 |
| C. Results | 192 |
| C.1 Development of a colorimetric assay for detecting sucrose-utilizing variants | 192 |
| C.2 Library construction and pre-screening of sucrose-utilizing variants | 194 |
| C.3. Isolation of improved Mutants for EA' and ED synthesis | 196 |
| D. Discussion..... | 199 |
| E. Methods | 201 |
| E.1. Plasmids, bacterial strains and chemicals..... | 201 |
| E.2. Construction of libraries containing two vicinal mutations (libraries 1-3)..... | 201 |
| E.3. Construction of library containing two distant mutations (library 4) | 202 |
| E.4. Development of the BBT test in liquid medium | 202 |
| E.5. Development of the BBT test at colony level | 203 |
| E.6. HPLC screening of variants able to glycosylate α -L-RhapOMe and D-GlcpNAc | 203 |
| F. References | 205 |

CONCLUSION207

ANNEXE215

TABLE DES ILLUSTRATIONS219

INTRODUCTION

Le sujet de cette thèse s'inscrit dans un programme de recherche, dont la finalité est le **développement de vaccins anti-bactériens multivalents, destinés à prévenir la shigellose**. Cette maladie, plus connue sous le nom de dysenterie bacillaire, est causée par *Shigella*, une entérobactérie Gram négatif, et est à l'origine d'environ un million de décès chaque année. La mise au point d'un vaccin contre la shigellose doit concilier **les critères de sécurité** définis par les autorités de la santé et les **contraintes humaines** (populations concernées : pays en voie de développement et enfants en bas âge) **et économiques** (coût de développement) liées à son application. Dans ce contexte, les constructions cibles retenues dans ce programme de recherche sont des **vaccins polyosidiques-conjugués chimiquement définis**. De manière générale, les composantes osidiques (haptènes) des vaccins glycoconjugués sont des oligosides synthétiques mimant les déterminants antigéniques portés par les polyosides (PS) de surface des bactéries. Ces antigènes polyosidiques sont les cibles majeures de la réponse immune protectrice. Le concept de ces vaccins dits "**de troisième génération**" a été récemment validé, en 2004, par la commercialisation du premier vaccin de cette catégorie (anti-Hib)¹. Concernant *Shigella*, la synthèse par voie chimique d'haptènes osidiques a permis d'élaborer un glycoconjugué à potentiel vaccinal prometteur contre les infections de *Shigella flexneri* sérotype 2a². Cette démarche, si elle est appliquée aux sérotypes les plus virulents de *Shigella*, permettrait d'élaborer à terme un vaccin multivalent à large spectre.

C'est dans ce cadre que s'inscrit le **projet "OptiGluc"**, soutenu en 2005 par l'Agence Nationale de la Recherche et dans lequel s'intègre la présente thèse. Ce projet multidisciplinaire coordonné par l'Institut Pasteur a pour objectif d'étendre la démarche développée pour le sérotype 2a de *Shigella flexneri* **aux sérotypes 1b et 3a**, espèces également prévalentes et responsables de la forme endémique de la shigellose. Il propose plus particulièrement **d'ouvrir un nouvel accès aux oligosides complexes** en mettant **en jeu des voies de synthèse chimio-enzymatiques originales** des haptènes cibles (1b et 3a). Le principal obstacle du procédé de synthèse de ce type d'antigènes se situe au niveau des réactions de glycosylation, difficilement réalisables par voie chimique. L'alternative proposée pour lever ce verrou consiste à réaliser cette réaction *via* des glyco-enzymes, en l'occurrence des glucane-saccharases natives ou construites « à façon » de manière à satisfaire les exigences de la voie de synthèse proposée par les chimistes. Nos travaux ont donc été consacrés au développement de cette étape enzymatique et sont présentés dans ce manuscrit.

L'ingénierie des glucane-saccharases des familles 13 et 70 des glycoside-hydrolases constitue une des thématiques majeures du groupe de Catalyse et Ingénierie Moléculaire EnzymatiqueS (CIMES) du Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés de l'INSA de

Toulouse. Ces transglucosidases possèdent des spécificités réactionnelles variées par le type de liaisons synthétisées, la taille des produits formés ou encore le spectre d'accepteurs reconnus et renferment des potentialités remarquables pour la synthèse de glycoconjugués de structure définie. Le défi majeur du projet de cette thèse consiste à mettre en œuvre les technologies d'ingénierie enzymatique disponibles pour créer des outils enzymatiques déployant de **nouvelles spécificités autorisant la glucosylation de molécules naturellement peu ou non reconnues, et produire à façon des synthons intervenant dans la voie de synthèse chimio-enzymatique des haptènes de série 1b et 3a**. Dans notre cas, il s'agira de glucosyler des dérivés de la *N*-acétyl-D-glucosamine et du L-rhamnose, porteurs de groupements protecteurs, qui permettront de poursuivre par voie chimique l'élongation des oligosaccharides antigéniques ciblés.

Dans l'étude bibliographique (**CHAPITRE I**), les deux premières parties ont pour objectif de familiariser le lecteur aux **thématiques abordées par ce projet de recherche** alors que la troisième partie est focalisée sur les **éléments bibliographiques se rapportant au développement de l'étape enzymatique**.

Aussi, la première partie exposera les progrès accomplis dans la mise au point de vaccins, plus particulièrement de ceux basés sur les antigènes saccharidiques qui ont récemment abouti à des vaccins "de troisième génération". La deuxième partie, portant sur notre modèle d'étude, *Shigella*, fera l'état des connaissances acquises sur cette bactérie qui ont permis d'orienter les recherches vers des stratégies vaccinales plus adaptées. Enfin, la troisième partie sera consacrée à la présentation des glucane-saccharases et des techniques d'ingénierie envisageables pour remodeler ces enzymes.

A la suite de ce chapitre, nous détaillerons **la stratégie envisagée pour l'ingénierie de catalyseurs enzymatiques** intervenant dans la synthèse des haptènes 1b et 3a.

Nous avons choisi de présenter les principales avancées réalisées au cours de cette thèse sous la forme de publications rédigées en anglais, dans un esprit de concision et afin d'en faciliter la divulgation. Le **CHAPITRE II** décrit les travaux initiaux portant sur **la glucosylation des accepteurs cibles** par un panel de glucane-saccharases de spécificités différentes, produites sous forme recombinante. Cette étude préliminaire nous a amenés à retenir l'amylosaccharase de *Neisseria polysaccharea*, comme enzyme modèle candidate à la mise en œuvre d'une stratégie raisonnée d'ingénierie pour l'optimisation de la réaction de glucosylation des deux accepteurs ciblés. Le **CHAPITRE III** fait la **démonstration du potentiel de la stratégie** d'ingénierie semi-rationnelle adoptée. Fondée sur la modélisation de complexes de l'amylosaccharase avec les produits de glucosylation attendus, cette stratégie a permis de guider avec succès la construction de banques

de variants adaptés à la voie de synthèse chimio-enzymatique envisagée. Enfin, le dernier chapitre est consacré à la **poursuite de la recherche de variants améliorés** par des techniques de recombinaison nécessitant le développement de nouvelles méthodes de criblage.

CHAPITRE I: Étude Bibliographique

PARTIE I : LES VACCINS BASÉS SUR LES ANTIGÈNES SACCHARIDIQUES

Le principe de la vaccination consiste à **administrer à l'organisme un principe actif qui déclenche une réponse et une mémoire immunitaire** contre un agent pathogène. En cas d'infection réelle, le **système immunitaire mobilisera plus rapidement ses défenses spécifiques** contre ce même pathogène et protégera efficacement l'organisme. Depuis le lancement en 1974 du **Programme Elargi de Vaccination (PEV)**, les campagnes de vaccination entreprises dans le monde entier ont permis de sauver des millions d'individus. Cependant, les maladies infectieuses restent une des principales causes de mortalité à l'échelon mondial et sont encore responsables d'environ 15 millions de décès annuels sur un total de 57 millions. Nombre de maladies infectieuses subsistent, en partie parce que l'efficacité du vaccin existant a été surestimée, s'est estompée au fil du temps (l'agent infectieux s'étant adapté par mutation à son environnement) ou tout simplement parce qu'aucun vaccin contre l'agent infectieux (comme le parasitisme du paludisme) n'a encore été trouvé. A cela, s'ajoute l'apparition de maladies émergentes (comme le SIDA, le SRAS, et la grippe aviaire). Pour éradiquer ces fléaux, la découverte de nouveaux vaccins et l'élaboration de nouvelles stratégies vaccinales s'avèrent primordiales. Cette première partie a pour objectif **d'exposer sommairement les avancées accomplies dans la conception de vaccins** et de présenter les **approches vaccinales disponibles ou envisageables** à ce jour. Elle s'étendra davantage sur les vaccins basés sur les antigènes saccharidiques, puisqu' à terme, c'est dans cette finalité que s'inscrit le sujet de cette thèse.

A. Des vaccins traditionnels aux nouvelles générations de vaccins

A.1. Vaccins traditionnels

L'histoire de la vaccination a commencé à la fin du XVIII^{ème} siècle avec les travaux du britannique Edward Jenner. Jenner constata que les fermiers contaminés par un virus bovin appelé vaccine (variole de la vache) développaient une réaction dermatologique mais ne contractaient jamais la variole humaine. Il imagina une pratique systématique de la vaccination. En 1796, il inocula le contenu d'une pustule prélevée sur la main d'une servante atteinte de la vaccine à un garçon de 8 ans et vérifia ainsi le pouvoir protecteur de la vaccine contre la variole.

Après la découverte de Jenner qui était basée sur un fait exceptionnel (l'immunisation envers une maladie humaine par un virus responsable d'une maladie animale), les travaux de Louis Pasteur **marquèrent le début de l'ère des vaccins modernes**. Avec ses collaborateurs, Louis Pasteur

découvrit que certaines bactéries, responsables du choléra des poules perdaient leur virulence après plusieurs semaines, ou exposition à des conditions défavorables. Il observa que des animaux inoculés avec ces bactéries ne développaient pas la maladie et se trouvaient même protégés contre une infection future avec l'agent pathogène virulent. C'est ainsi que **le concept de vaccination, basé sur l'emploi de germes pathogènes atténués prit naissance**. Appliquant ce principe à d'autres maladies animales puis à la rage, il sera à l'origine de la mise au point du premier vaccin antirabique en 1885.

De nombreux vaccins reposent encore **de nos jours sur l'utilisation de micro-organismes vivants atténués** (rougeole, rubéole, varicelle et fièvre jaune). L'atténuation des souches virulentes est obtenue par voie chimique, par passage prolongé du pathogène virulent dans divers hôtes, ou en faisant croître l'agent pathogène dans des conditions de culture anormales. Ce dernier procédé permet de sélectionner les mutants qui sont les mieux équipés pour une croissance dans des conditions différentes de celles rencontrées dans l'hôte naturel et par conséquent, sont moins capables de croître au sein de celui-ci. Ces vaccins sont principalement dirigés contre des virus et non contre des bactéries. En effet, il n'existe qu'un seul vaccin anti-bactérien de ce type largement utilisé chez l'homme, le vaccin contre la tuberculose ou B.C.G. (Bacille de Calmette-Guérin) dont l'efficacité est d'ailleurs controversée.

L'autre approche classique employée pour la production de vaccins consiste à utiliser **des micro-organismes tués (ou inactivés)**. L'agent pathogène, inactivé par la chaleur ou plus communément par des traitements chimiques comme le formaldéhyde (employé pour les vaccins contre la poliomyélite (Salk) et la coqueluche) ou divers agents alkylants, est alors incapable de se répliquer chez l'hôte. Tout comme les vaccins atténués, les vaccins inactivés doivent être capables de simuler une infection naturelle et déclencher une réponse immunitaire. Il est donc extrêmement important, au cours de l'inactivation, de maintenir la structure des antigènes de surface qui seront reconnus par le système immunitaire. Plusieurs vaccins viraux (contre la grippe, l'hépatite A et B, l'encéphalite à tiques et japonaise, et la rage) et bactériens (contre la leptospirose et le choléra) de ce type sont commercialisés actuellement.

Les vaccins, utilisant des organismes pathogènes entiers, présentent néanmoins **des inconvénients en termes de difficultés de mise au point, d'efficacité, de sécurité et de coût de production**. Ne nécessitant généralement qu'une seule et faible dose, les vaccins vivants atténués sont plus efficaces que les vaccins tués pour lesquels des rappels réguliers contenant une masse importante de pathogènes et d'adjuvants sont nécessaires pour maintenir l'état immun de l'hôte. Ces rappels expliquent leur coût de production plus élevé. D'autre part, atténuer ou éliminer la virulence d'un

agent infectieux en conservant ses capacités immunogènes peut s'avérer extrêmement long, voire impossible. A titre d'exemple, 13 ans ont été nécessaires pour qu'une souche de *Mycobacterium bovis* se soit adaptée à croître sur un milieu contenant de la bile en concentrations croissantes et soit alors suffisamment atténuée pour être utilisable comme vaccin contre la tuberculose humaine (seul vaccin disponible actuellement en France). Utilisé depuis 85 ans, ce vaccin, dont l'efficacité (estimée à 50% chez l'adulte) a été vivement critiquée ces dernières années, ne peut enrayer la transmission de la maladie à l'échelle mondiale. Cette année, l'administration de ce vaccin jusque-là obligatoire en France a été suspendue. Enfin, il se peut aussi qu'un virus atténué redevienne pathogène ou soit contaminé par d'autres virus présents lors de la production du vaccin, (cas du vaccin Sabin contre la poliomyélite) ou enfin, qu'un vaccin inactivé contienne des pathogènes qui n'ont pas été tués lors du traitement (cas du vaccin Salk). Du fait d'une composition difficilement contrôlable, ces vaccins présentent toujours des risques d'effets secondaires, pas toujours prévisibles et tolérables par l'organisme.

A.2. Nouvelles générations de vaccins

Ces 20 dernières années, les progrès scientifiques acquis **dans les domaines de l'immunologie, de la microbiologie, de l'analyse structurale et de la biologie moléculaire** ont permis de mieux comprendre les mécanismes de multiplication des agents pathogènes et du déterminisme de leur virulence, d'identifier les facteurs responsables de leur caractère antigénique et finalement d'analyser plus finement la réponse immunitaire anti-infectieuse. Ainsi, de nouvelles stratégies vaccinales sont proposées **en complément plutôt qu'en remplacement des approches traditionnelles**³.

- La première voie, la moins utilisée, consiste à créer de nouvelles souches mutantes de microorganismes pathogènes totalement inoffensives. Il s'agit d'inactiver ou d'éliminer les gènes impliqués dans la virulence du pathogène. On parle alors de **vaccins vivants atténués par génie génétique**.

- La seconde, plus courante, ne retient de l'agent pathogène **qu'une fraction immunogène de celui-ci, préalablement identifiée comme majeure** pour l'utiliser comme vaccin.

Habituellement, les immunogènes viraux majeurs correspondent à des (glyco)protéines de structure, de la capsid (composant interne des virus) ou de l'enveloppe. Les immunogènes bactériens, étudiés depuis plus longtemps, sont de nature saccharidique ou protéique. Pour les immunogènes de type saccharidique, on distingue principalement les polysaccharides capsulaires (PC) (**antigène K ou Ag-K**), les lipopolysaccharides (LPS) (**antigène somatique ou antigène O ou Ag-O**) ou les lipooligosaccharides (LOS), constituants des parois bactériennes. Les

immunogènes protéiques sont les protéines de pili et de flagelles (**antigène H ou Ag-H**) ainsi que les exotoxines bactériennes. L'identification de ces structures antigéniques combinée à l'utilisation des nouvelles technologies a ouvert de nouvelles portes aux stratégies vaccinales qui se distinguent selon le mode de présentation de ces immunogènes au système immunitaire. On distingue :

- **les vaccins sous unités (vaccins inertes) :** Les antigènes sont obtenus soit en fractionnant le pathogène et en conservant uniquement les parties immunogènes, soit en le produisant *in vitro* par des techniques de génie génétique (antigène recombinant). Les sous fractions peuvent également être synthétisées par synthèse chimique, par exemple. Lorsque les fragments antigéniques sont de courte taille (dénommés alors **haptènes**) ou sont formés par des sucres, il est très souvent nécessaire de les coupler chimiquement à une molécule porteuse (anatoxine tétanique ou diphtérique) pour les rendre immunogènes. Il s'agit alors de **vaccins conjugués**.
- **les vaccins utilisant des vecteurs recombinants (vaccins vivants) :** Les gènes codant pour les principaux antigènes du pathogène sont introduits dans d'autres microorganismes "vecteurs" (généralement dans des virus atténués ou bactéries) capables de se répliquer au sein de l'hôte et d'exprimer le produit du gène du pathogène à leur surface ou de les sécréter dans le milieu extérieur. Ces "microbes présentoirs" sont des vaccins mixtes puisqu'ils permettent de vacciner à la fois contre le vecteur et contre le virus ou la bactérie dont ils présentent les antigènes au système immunitaire.
- **les vaccins de délivrance de l'antigène par des vecteurs "non-réplicatifs" ou sans vecteur (vaccins intermédiaires) :** Les gènes codant pour l'antigène étranger au système immunitaire sont cette fois-ci délivrés par l'intermédiaire de vecteurs viraux incapables de se multiplier ("enveloppes vides"), ou sous forme d'ADN nu par injection intra-musculaire ou par "bombardement" de particules sous la peau. Ces vaccins sont inertes au sens où ils ne se multiplient pas dans l'organisme mais ils présentent l'antigène au système immunitaire sous sa forme native.

La découverte de nouveaux antigènes cibles utilisables à des fins vaccinales est aujourd'hui largement facilitée par les outils de la génomique⁴ qui permettent de déterminer la séquence génomique complète d'un agent pathogène en quelques semaines voire quelques jours. A ce titre, la **vaccinologie inverse** ou "**reverse vaccinology**"⁵ peut remédier aux lacunes de la vaccinologie classique qui traite uniquement des pathogènes cultivables *in vivo* et n'identifie que les antigènes les plus abondants. Cette approche vaccinale débute, *in silico*, en utilisant les séquences de génomes microbiens plutôt que le pathogène lui-même. Il s'agit d'identifier et de caractériser, les gènes codant pour tous les antigènes potentiels (de l'ordre de la centaine), d'analyser leur

localisation et le niveau d'expression des protéines correspondantes. Les candidats antigènes les plus prometteurs sont, par exemple, exprimés à haut-débit chez *E. coli* et leur immunogénicité est évaluée. Selon les estimations, deux ans suffiraient alors pour développer un vaccin alors que la vaccinologie actuelle nécessite un délai de temps très long, de 5 à 15 ans (**FIGURE 1**). Une des limites de cette approche est son incapacité à **identifier des antigènes autres que protéiques** tels que les composés saccharidiques.

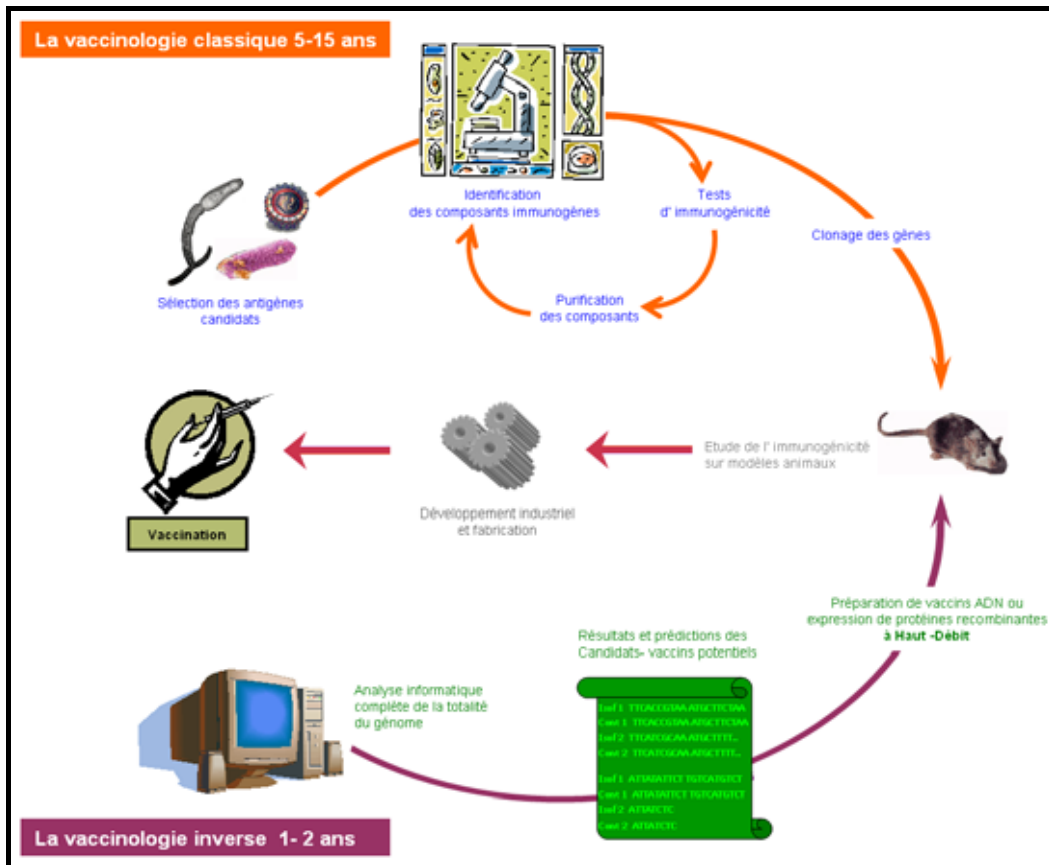


FIGURE 1| Représentation schématique des étapes essentielles pour la mise au point d'un vaccin par la vaccinologie classique et de la vaccinologie inverse. (D'après Rapuoli et coll. ⁵)

Ces nouveaux vaccins offrent des perspectives très prometteuses puisque nombre d'entre eux ont déjà fait leur preuve, atteignant le stade de la commercialisation (vaccins recombinants, sous-unités et conjugués). Quant aux autres, ils sont soumis à des essais cliniques sur des modèles animaux puis chez l'homme. Un article⁶ paru en 2004 dans *Nature Médecine*, montre que les catégories de vaccins actuellement mis sur le marché se répartissent équitablement entre vaccins tués, vaccins vivants atténués et vaccins sous-unités. En 2020, l'auteur prévoit que ces vaccins soient en partie supplantés par les stratégies nouvelles faisant intervenir les sous-unités obtenues par vaccinologie inverse, des vecteurs non réplicatifs, des vaccins ADN ou muqueux (**FIGURE 2**). Il faut cependant rester prudent sur le temps nécessaire pour introduire une nouvelle stratégie et

généraliser son utilisation. Vingt cinq ans après le premier vaccin recombinant (contre l'hépatite B), peu de vaccins de cette catégorie ont été mis sur le marché.

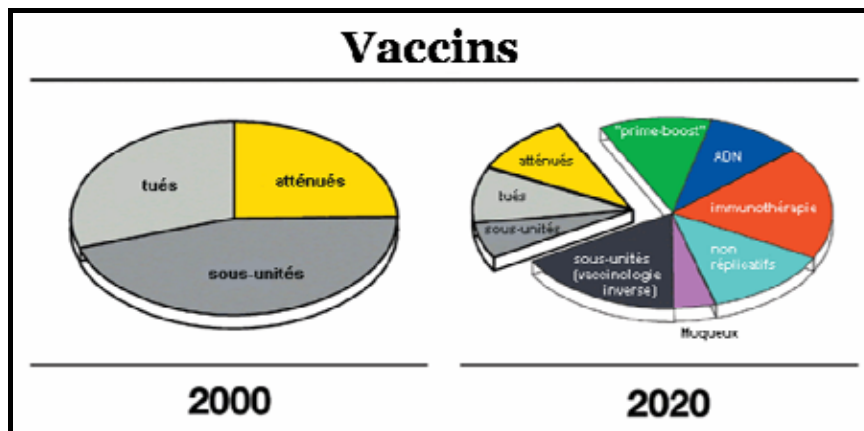


FIGURE 2| Type de vaccins enregistrés en 2000 et prédiction des types de vaccins disponibles en 2020 (D'après Rapuoli et coll.⁶)

La vaccinologie reste un domaine en pleine évolution tant au niveau de la découverte de nouveaux vaccins qu'au niveau de l'amélioration de la qualité et de l'innocuité des vaccins disponibles et futurs. Cependant, le développement d'un nouveau vaccin est une entreprise très longue (au moins 5-15 ans), coûteuse (au moins \$500 Millions) et de plus en plus risquée. La demande croissante en nouveaux vaccins n'est généralement pas traitée en fonction des besoins urgents des populations mais du marché potentiel pour ce vaccin. Or, les pays en voie de développement, où surviennent plus de 90% des maladies infectieuses, sont généralement lents pour adopter les nouveaux vaccins et les marges bénéficiaires y sont faibles. **Les vaccins destinés principalement à ces pays ne sont donc pas souvent développés en priorité.** De tels vaccins, que l'OMS⁷ qualifie de « négligés » sont ceux qui proposent une prévention contre la shigellose -projet de l'étude- et le choléra (maladies bactériennes), la dengue et l'encéphalite japonaise (maladies virales), la leishmaniose et la schistosomiase (maladies parasitaires). Une des priorités que s'est fixée l'Alliance mondiale pour les vaccins et la vaccination (GAVI) est **d'accélérer la recherche et le développement de ces vaccins « négligés »**. En plus de participer financièrement et activement à la R&D de ces vaccins, cette alliance, qui réunit les principaux acteurs (privés et publics) du secteur de la vaccination (OMS, UNICEF, Groupe de la Banque Mondiale...) aide les pays les plus pauvres à augmenter leur couverture vaccinale en vaccins existants, à développer des technologies qui facilitent la distribution et l'administration des vaccins localement et à introduire des vaccins sous utilisés (tels que l'hépatite B, Hib et la fièvre jaune). La recherche et le développement de vaccins se heurtent également à un autre obstacle. Les nouveaux vaccins mis au point contre des maladies rencontrées dans les pays industrialisés **ne sont souvent pas adaptés aux pays en voie de développement.** Jusqu'à récemment, les essais cliniques étaient effectués

tardivement sur les populations touchées. Maintenant, ils sont réalisés dès les premiers stades de la conception du vaccin pour pouvoir réorienter plus rapidement les recherches si nécessaire.

B. Vaccins basés sur les antigènes saccharidiques

Après avoir donné une vision générale des avancées récentes accomplies dans la conception de nouveaux vaccins, nous nous consacrerons plus en détail à l'une d'entre elles : les vaccins sous-unités basés sur les antigènes saccharidiques. En premier lieu, nous nous focaliserons sur les antigènes saccharidiques d'origine bactérienne dont le potentiel vaccinal a été validé pour trois générations de vaccins⁸⁻¹⁰. Ce succès a ouvert de nouvelles perspectives d'application. Ainsi, cette stratégie vaccinale est aujourd'hui extrapolée à d'autres pathologies comme les infections fongiques, parasitaires et aussi certains cancers¹⁰⁻¹³.

B.1. Antigène saccharidique bactérien et potentiel antigénique

Les sucres exposés à la surface de la plupart des bactéries sont les cibles majeures de la réponse immune protectrice de l'hôte envers l'agent infectieux. On peut les trouver sous forme de **polysaccharides capsulaires (PC)** chez les bactéries Gram positif et Gram négatif, et sous forme de **lipopolysaccharides (LPS)** (et plus rarement LOS) comme composants de la membrane externe des bactéries Gram négatif (**FIGURE 3**).

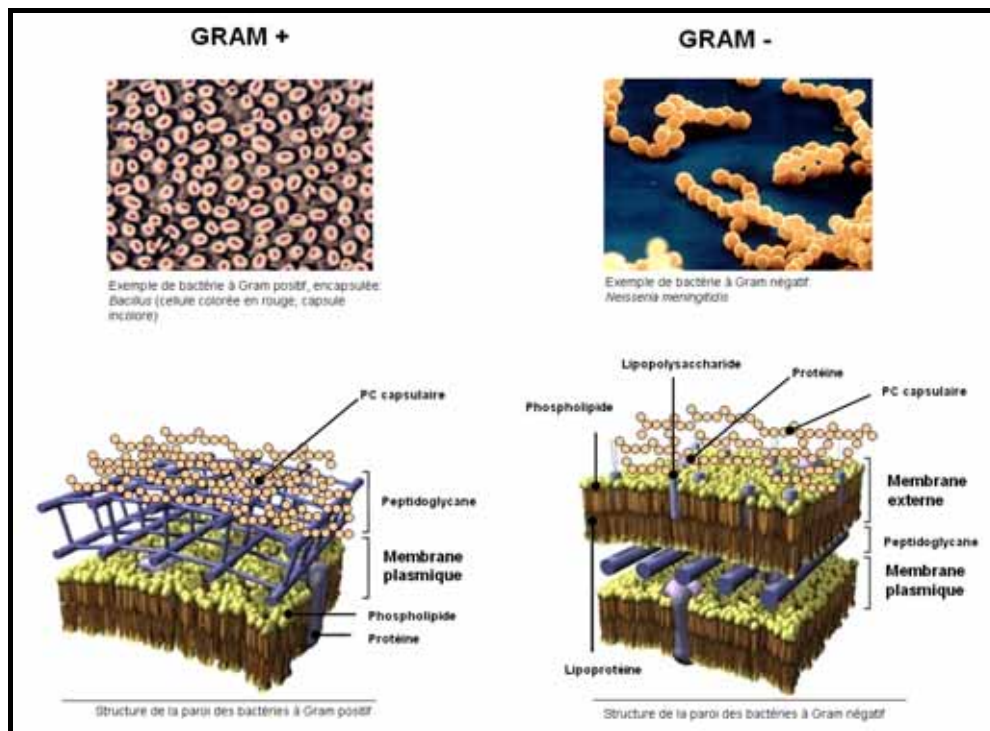


FIGURE 3| Structure de la paroi des bactéries à Gram positif et à Gram négatif.

(Source images: <http://www.bact.wisc.edu/themicrobialworld/N.meningitidis.jpg>;
http://stl_bjb.ac-dijon.fr/microbio/image/gramplus.jpg et [grammoins.jpg](http://stl_bjb.ac-dijon.fr/microbio/image/grammoins.jpg))

Les LPS sont composés de trois éléments : le lipide A, ancré dans la membrane grâce à des interactions hydrophobes et responsable de la toxicité du LPS; l'oligoside central (appelé "core") dont la structure diffère peu d'une souche bactérienne à l'autre, au sein d'une même espèce. Il assure le lien entre le lipide A et le polysaccharide spécifique (**O-PS**) exposé à la surface et présentant la plus grande diversité de composition et de structure (**FIGURE 4**).

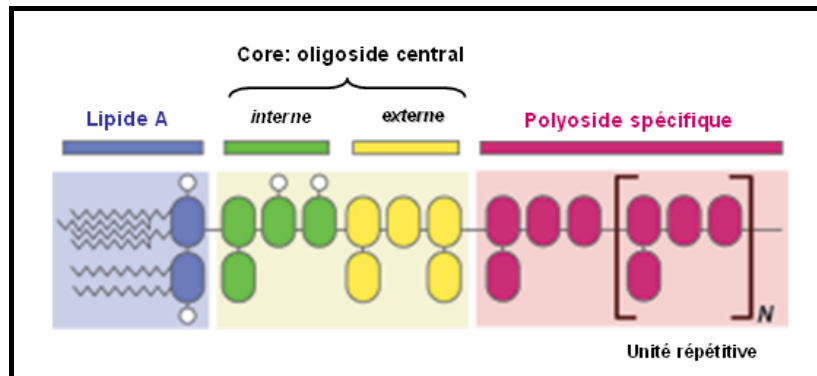


FIGURE 4 | Structure des LPS

Source Image: http://www.uoguelph.ca/~cwhitfie/lps_core.html

La majorité des bactéries possède des PC et/ou O-PS. Les PC sont des polymères de haut poids moléculaire fortement hydratés (plus de 95 % d'eau), constitués de quelques centaines d'unités répétitives, linéaires ou ramifiées, le plus souvent de nature anionique. Les O-PS, généralement neutres, sont de plus petite taille, puisque composés de quelques dizaines d'unités répétitives. Ces motifs répétitifs sont formés d'au plus 8 monosaccharides constitutifs¹⁴, classiques (tels que le D-glucose et le D-galactose) ou moins conventionnels¹⁵ (tels que le tyvélose, le paratose, l'abéquose et le colitose). Outre la nature diverse de ces résidus, les variations rencontrées au niveau des liaisons glycosidiques, les configurations des carbones anomériques (forme α ou β), et les isomères de forme (forme D ou L) ou de conformation (forme pyranose ou furanose) contribuent à l'étonnante diversité structurale de ces polysaccharides bactériens. De surcroît, ces unités répétitives peuvent aussi contenir des substituants non saccharidiques, tels que des groupes O-acétyl, glycérol phosphate ou pyruvate. Les différentes empreintes saccharidiques induisent des réponses immunitaires distinctes qui, pour une même espèce, permettront de classer les bactéries par sérotype et/ou sérotype (**TABLEAU 1**). Cette diversité antigénique est très bien illustrée par les 90 et 2200 sérotypes identifiés respectivement pour *Streptococcus pneumoniae* et *Salmonella enterica*. Il faut, de plus, souligner que de très faibles variations de la structure polysaccharidique peuvent modifier la réponse immunitaire. Par exemple, les PC des sérotypes 1a et 1b de *Streptococcus agalactiae* ne diffèrent que par la liaison osidique entre le D-galactose et le N-acetyl-D-glucosamine ($\beta(1-3)$ vs $\beta(1-4)$) au niveau de leur unité répétitive¹⁶. Du fait de la variabilité antigénique, il est donc difficile de concevoir un vaccin **offrant des protections croisées** contre tous les sérotypes d'une même espèce.

TABLEAU 1| Exemples d'unités répétitives de PC et O-PS des bactéries à Gram positif et à Gram négatif

| Bactérie (espèce) | Gram | Type d'antigène polysaccharidique | Sérogroupe/ Sérotype | Unités répétitives | Références |
|---------------------------------|---------|-----------------------------------|----------------------|---|------------|
| <i>Streptococcus pneumoniae</i> | positif | capsule polysaccharidique (PC) | Type 2 | →4)-β-D-Glcp-(1→3)-[α-D-GlcpA-(1→3)-α-D-Glcp-(1→2)]-α-L-Rhap-(1→3)-α-L-Rhap-(1→3)-β-L-Rhap-(1→ | 17 |
| | | | Type 6b | →2)-α-D-Galp-(1→3)-α-D-Glcp-(1→3)-α-L-Rhap-(1→4)-D-Rib-ol-(5→P→ | 17 |
| | | | Type 18c | →4)-β-D-Glcp-(1→4)-[α-D-Glcp(6OAc)-(1→2)]-[Gro-(1→P→3)-β-D-Galp-(1→4)-α-D-Glcp-(1→3)-β-L-Rhap-(1→P→ | 17 |
| | | | Type 19a | →4)-β-D-ManpNAc-(1→4)-α-D-Glcp-(1→3)-α-L-Rhap-(1→P→ | 17 |
| <i>Neisseria meningitidis</i> | négatif | capsule polysaccharidique (PC) | Groupe A | →6)-α-D-ManpNAc(3OAc)-(1→OPO ₃ → | 17 |
| | | | Groupe C | →9)-α-D-Neu5Ac(7/8OAc)-(2→ | 17 |
| | | | Groupe W 135 | →6)-α-D-Galp-(1→4)-α-D-Neu5Ac(9OAc)-(2→ | 17 |
| | | | Groupe Y | →6)-α-D-Glcp-(1→4)-α-D-Neu5Ac(9OAc)-(2→ | 17 |
| <i>Escherichia coli</i> | négatif | Lipolysaccharide (LPS) | O111 | →4)-[α-Col-(1→6)]-[α-Col-(1→4)]-α-D-Glcp-(1→4)-α-D-Galp(1→3)β-D-GlcpNAc(1→ | 18, 19 |
| | | | O157 | →2)-α-D-Rhap4NAc-(1→3)-α-L-Fuc(1→4)-β-D-Glcp(1→3)-α-D-GalpNAc(1→ | 20 |

Le **recours aux vaccins multivalents** visant les souches les plus virulentes est alors nécessaire pour assurer une protection efficace.

D'autre part, il a été démontré que les PC sont impliqués dans la virulence des bactéries²¹. En l'absence de leur capsule, la pathogénicité des bactéries Gram⁺ (telles que les streptocoques des groupes A²² et B²³) est affectée autant que celle des Gram⁻ (telles que *E. coli* K1²⁴ et *H. influenzae* type b²⁵). En effet, les capsules confèrent généralement une capacité antiphagocytaire aux bactéries. Le rôle du LPS, comme facteur de virulence, est bien moins défini et est variable suivant la bactérie. Il pourrait cependant, dans certains cas, jouer aussi le rôle de « bouclier », protégeant la bactérie du système immunitaire. Concernant *Shigella flexneri*, les souches dépourvues de LPS sont incapables, *in vivo*, de résister aux facteurs anti-bactériens. De plus, le LPS participe aussi au caractère invasif des shigelles : ces bactéries ont développé, au cours de leur évolution, la capacité de modifier, par une simple activité enzymatique, leurs molécules protectrices (LPS) pour les raccourcir et faciliter ainsi l'invasion des bactéries pour franchir la barrière épithéliale²⁶. Chez certaines espèces telles que *N. meningitidis*, *N. gonorrhoeae*, *H. influenzae*, *Campylobacter jejuni* et *Helicobacter pylori* les structures polysaccharidiques du LPS miment la structure de certaines molécules présentes chez l'hôte et ce camouflage permettrait à ces bactéries d'échapper aux réponses immunitaires.

C'est au début du XX^{ème} siècle, que les structures saccharidiques de type PC ont pour la première fois été identifiées, chez *S. pneumoniae*, et leur caractère antigénique mis en évidence. En 1931, ces observations aboutirent à la conception du premier vaccin polyosidique purifié qui permit d'enrayer une épidémie de pneumonie au Massachusetts (USA). Supplanté par l'efficacité et le faible coût des antibiotiques, ce vaccin fut rapidement retiré du marché. Lorsque la résistance à la pénicilline apparut à partir de 1947, les recherches reprirent mais il fallut progresser dans les techniques de purification et d'analyse pour exploiter le potentiel vaccinal des PC purifiés et aboutir dans les années 1970-80 à la première génération de vaccins polysaccharidiques.

B.2. Les Vaccins Composés de PC Purifiés : 1^{ère} Génération de vaccins polyosidiques

Ces vaccins utilisent les **polysaccharides capsulaires comme immunogènes**. Les polysaccharides capsulaires purifiés^{27, 28} sont obtenus par extraction à partir de cultures bactériennes. Ils doivent avoir conservé leurs caractéristiques chimiques et antigéniques et être dépourvus de tout contaminant biologique pour être dotés d'**une excellente innocuité**. Notons que sur ce plan, l'emploi des PC purifiés représente un avantage énorme par rapport à celui de bactéries entières inactivées. Administrés à faible dose (quelques dizaines de µg), ils stimulent la production d'anticorps bactéricides et, par suite, la phagocytose immune. Cependant, ces vaccins

polysaccharidiques ne sont **pas ou peu immunogènes chez les enfants de moins de deux ans**, chez certaines personnes âgées ou individus présentant des déficiences immunitaires. En effet, les polysaccharides sont reconnus par les lymphocytes B mais n'activent les lymphocytes T ; ils sont dits T-indépendants (TI)²⁹. Or, chez l'humain, les cellules B des nouveaux-nés ne répondent pas à la plupart des antigènes polysaccharidiques. En effet, ce type de réponse ne se développe que progressivement dans les premières années de la vie. Le caractère antigénique T-indépendant des vaccins polysaccharidiques a comme autre conséquence de ne conférer qu'une **protection de courte durée** (~ 5 ans). En effet, la réponse T-indépendante n'induit ni cellule mémoire, ni mutation somatique, ni maturation d'affinité, produisant essentiellement des immunoglobulines d'isotype M (IgM). Des rappels sont alors nécessaires. La faible immunogénicité des vaccins polysaccharidiques est donc un **inconvenient majeur**. De plus, certains PC peuvent leurrer le système immunitaire et donc se révéler totalement inefficaces. C'est le cas du PC séro groupe B de *N. meningitidis*, bactérie responsable d'environ la moitié des cas de maladies dues aux méningocoques, dont le motif composé d'un homopolymère $\alpha(2-8)$ d'acide *N*-acétylneuraminique est retrouvé dans les tissus humains. Dans ces cas particuliers, le recours à la vaccinologie inverse (décrite précédemment) prend tout son intérêt³⁰.

Trois vaccins à base de PC purifiés sont encore largement utilisés aujourd'hui chez l'adulte et l'enfant de plus de 2 ans. Il s'agit tout d'abord des vaccins protégeant contre la méningite. Disponibles depuis 1976 sous forme bivalente (séro groupe A et C), ces vaccins ciblés contre les méningites provoquées par *N. meningitidis* sont aussi commercialisés sous forme tétravalente (séro groupes A, C, Y et W135)³¹ depuis 1982. L'année suivante, un vaccin composé des PC de 23 sérotypes de *S. pneumoniae* est mis sur le marché en complément du vaccin 14-valent produit depuis 1977. Il assure théoriquement une protection contre 85 à 90% des pneumocoques responsables d'infections chez l'adulte. Enfin, le dernier vaccin dans cette catégorie, commercialisé depuis 1989, est le vaccin contre la fièvre typhoïde. Composé du PC purifié de *S. enterica* sérotype Typhi (antigène Vi), il supplanta l'ancien vaccin préparé à partir de bactéries entières inactivées et utilisé depuis 1896.

B.3. Les Vaccins Conjugués PS-Protéine : 2^{nde} génération de vaccins polysaccharidiques

Les vaccins polysaccharidiques de deuxième génération, qui apparaissent dans les années 90, sont des **conjugués polysaccharide-protéine**. Contrairement aux polysaccharides, les glycoprotéines sont des antigènes T-dépendants (TD), induisant ainsi une réponse immunitaire beaucoup plus efficace pour le même antigène. Cette propriété fut mise en évidence dès 1931 par Avery et Goebel³² qui convertirent le PC d'un pneumocoque, un antigène peu immunogène ou « haptène »

en un antigène très immunogène par couplage covalent à une protéine. Elle fut ensuite employée pour apporter une solution au problème majeur des vaccins de première génération, à savoir **leur utilisation chez les enfants de moins de deux ans**. Purifiés à partir de cultures bactériennes puis liés à une protéine porteuse ou *carrier*, ces PC induisent une immunité T-dépendante, qui aboutira, grâce à une coopération efficace entre cellules immunitaires B et T, à **l'expression d'immunoglobulines IgG** spécifiques du polysaccharide et à **la production de cellules B mémoire** (FIGURE 5).

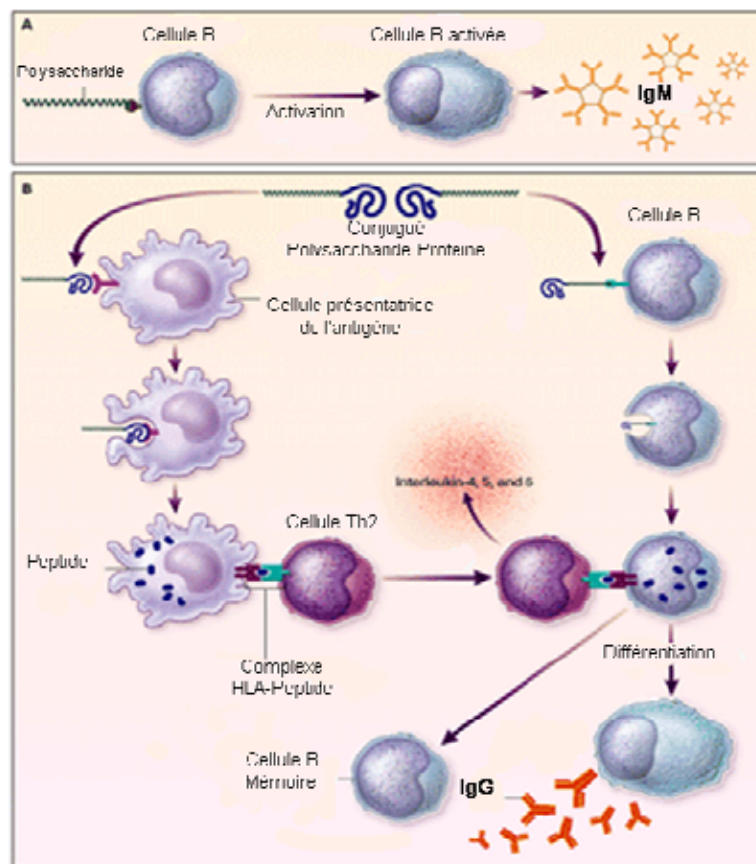


FIGURE 5 | Production d'anticorps stimulés par des antigènes polysaccharidiques conjugués ou non à des protéines porteuses.
(D'après Ada³³)

En (A), un antigène polysaccharidique se lie à un récepteur de la surface d'une cellule B (lymphocyte B). Cette dernière sera activée et produira des IgM.

En (B), des conjuguats polysaccharide-protéines sont pris en charge par des cellules dendritiques, lesquelles présentent ensuite des peptides provenant de la fraction protéique de l'antigène aux cellules Th2. D'autres conjuguats se lient aux cellules B qui ont des récepteurs spécifiques IgM spécifiques pour la partie polysaccharidique de l'antigène; ces cellules vont exprimer à leur surface des peptides de la partie protéique en association avec le HLA de classe II. L'ensemble va être reconnu par les Th2 activés qui secréteront IL-4, IL-5 et IL-6. Ceci conduira à la différenciation des cellules B et à l'expression de IgG spécifiques des polysaccharides, puis à leur maturation en plasmocytes.

Le développement des vaccins glycoconjugués est considéré comme un succès majeur dans les sciences biomédicales. Les premiers vaccins monovalents de cette catégorie (ProHibit[®], Hib-

Titer[®], PedVax Hib[®], Act-Hib[®]), enregistrés entre 1986 et 1993, préviennent les méningites à *Haemophilus influenzae* de sérotype b (Hib) touchant principalement les enfants âgés de 4 à 18 mois. Après leur introduction dans les programmes de vaccination infantile, les pathologies à Hib ont considérablement diminué dans les pays développés (Europe occidentale, Canada, Etats-Unis, Australie et Nouvelle-Zélande). Une utilisation généralisée à l'échelle mondiale pourrait conduire à l'éradication complète de la maladie. Ce premier succès s'est poursuivi par la commercialisation de vaccins conjugués visant les infections à *N. meningitidis* Groupe C (Meningitec[®], Menjugate[®] et NeisVacC[®]) et celles à *S. pneumoniae* (vaccin heptavalent : Prevnar[®]). Des vaccins conjugués ciblant les pathologies induites par *Streptococcus* Groupe A et B, *S. typhi*, *Staphylococcus aureus*, et *Vibrio cholerae* sont en phase de développement. Le sont également des glycoconjugués destinés à améliorer les vaccins préexistants contre *S. pneumoniae* (de valence 7, 9 ou 11) et *N. meningitidis* Groupe A, C, W135 et Y.

La mise au point de tels vaccins fait appel à une **stratégie de construction modulable**, basée sur :

- (a) le choix de la protéine porteuse
- (b) la taille des polysides utilisés : polyside natif ou fragmenté
- (c) les modes de couplage entre le polyside et la protéine porteuse
- (d) l'utilisation ou non d'un bras espaceur (*spacer*)
- (e) la méthode chimique de couplage du polyside sur la protéine
- (f) le taux de greffage

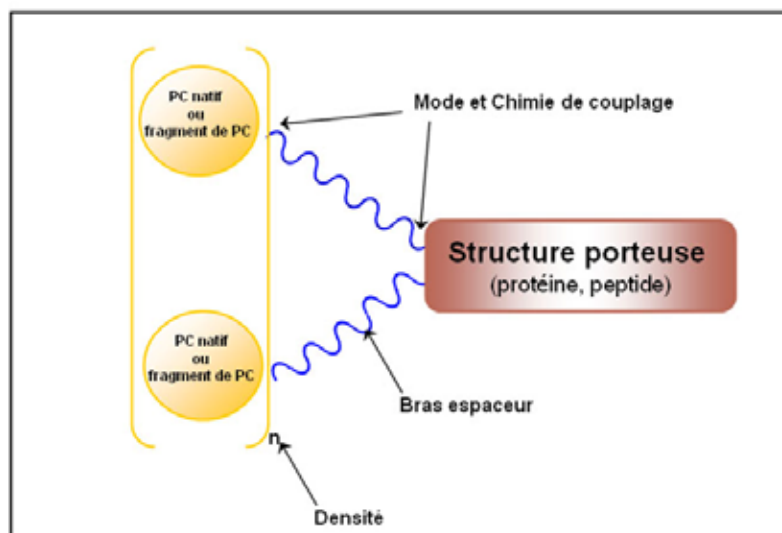


FIGURE 6| Représentation schématique de la stratégie de construction des vaccins glyco-conjugués.

Ces multiples paramètres révèlent la **complexité** de l'approche **mais aussi la flexibilité d'adaptation de la méthode pour l'optimisation à façon du vaccin ciblé** (TABLEAU 2 / FIGURE 6).

- **La protéine porteuse.** Son choix doit répondre à des critères d'innocuité, de faible coût et de disponibilité, de pureté, de solubilité, de stabilité, d'accessibilité au niveau de ses sites réactifs et de potentiel immunogène. Aujourd'hui, quatre protéines porteuses sont autorisées chez l'homme : l'anatoxine tétanique (TT), l'anatoxine diphtérique (DT) et un de ses mutants non toxique (CRM197)³⁴ et un complexe protéique de la membrane externe de *N. meningitidis* Groupe B³⁵ (OMPC), ce qui est un nombre très insuffisant face au développement croissant des vaccins conjugués. La **diversification** a pour but d'éviter les phénomènes de **suppression épitopique** induits notamment par l'existence d'anticorps formés contre la protéine porteuse lors d'une précédente vaccination. Ces phénomènes sont également observés lors de formulations multivalentes (combinaison de vaccins conjugués utilisant une même protéine porteuse), pour lesquelles se produit un effet de **compétition antigénique entre les cellules B spécifiques des PS et celles spécifiques de la protéine**³⁶. Pour limiter cet effet, il est recommandé d'utiliser des protéines porteuses différentes pour chaque conjugué de la composition vaccinale^{37, 38}.

De nouvelles protéines bactériennes, cibles de l'immunité protectrice chez l'homme sont donc proposées. L'étude la plus avancée porte sur un mutant recombinant non toxique de l'exoprotéine A de *Pseudomonas aeruginosa* (**rEPA**) qui a déjà fait l'objet d'essais cliniques pour de multiples conjugués³⁹⁻⁴¹. Le **recours à des peptides contenant des épitopes T auxiliaires, mais dépourvus d'épitopes B** est une alternative encore plus prometteuse. De taille réduite, ces peptides seraient plus performants que les protéines de grande taille⁴². Grâce à une meilleure reproductibilité des réactions de conjugaison, la production des conjugués à grande échelle serait de plus facilitée. D'origine virale ou bactérienne, les peptides sont sélectionnés pour leur capacité à se fixer aux molécules HLA-DR (*antigènes leucocytaires humains de classe II*) et par conséquent, à induire une réponse immunitaire T dépendante chez l'humain. De ce fait, ils sont dits universels. Le concept a été validé en assemblant, par ingénierie génétique, les gènes codant dix-neuf épitopes issus de la combinaison de dix épitopes T universels différents, pour produire la protéine recombinante N19. Conjuguée à des PS anti-Hib⁴³ ou anti-MenC^{44, 45}, cette protéine s'avère autant voire plus efficace qu'une protéine porteuse conventionnelle.

Les évolutions les plus récentes s'orientent vers l'utilisation de **peptides T auxiliaires universels non naturels** conçus par ingénierie moléculaire pour être spécifiquement immunogène chez l'humain⁴⁶, comme le peptide le PADRE[®] (*Pan HLA-DR-binding Epitope*) qui est composé de 13 acides aminés.

- **La taille des polysides utilisés.** La taille de l'antigène saccharidique est un élément crucial de la conception du glycoconjugués. Considérations théoriques, données expérimentales et contraintes vaccinales de reproductibilité incitent à l'utilisation de **fragments antigéniques**

spécifiques de taille courte et contrôlée, obtenus par dépolymérisation sélective⁴⁷ plutôt qu'à celle du PC natif, difficilement soluble et de répartition hétérogène (polydisperse)⁴⁸⁻⁵⁰. Le degré de polymérisation de la composante saccharidique correspond à celui de l'épitope complémentaire du site de reconnaissance des anticorps protecteurs. Il est généralement compris entre 2 et 8 résidus mais pourra être plus élevé si les anticorps reconnaissent un épitope conformationnel.

- **Les modes de couplage entre le polyside et la protéine porteuse.** Trois modes de greffage du polyside sur la protéine sont envisageables. Le premier, préféré pour le greffage de chaîne courte, fait intervenir **un seul point d'ancrage** entre la protéine et le polyside, généralement à son extrémité réductrice et conduira à l'obtention d'une néoglycoprotéine. Si le polyside est aléatoirement activé le long de sa chaîne, il pourra se lier à plusieurs protéines voire à d'autres polysides formant ainsi des réseaux réticulés. Ce mode de greffage à **points d'ancrage multiple** est recommandé pour accroître l'immunogénicité des polysides de haut poids moléculaires. Pour ces deux procédés, il est possible de coupler directement le polyside sur la protéine porteuse ou d'introduire un bras espaceur. L'introduction d'un espaceur de longueur variable présente deux avantages. Elle facilite l'étape de conjugaison, évitant toute gêne stérique entre la protéine et le polyside et favorise l'exposition des épitopes osidiques limitant ainsi un éventuel masquage par la protéine. La troisième approche, moins utilisée, consiste à conjuguer le polysaccharide à des protéines présentes sur la membrane externe de liposomes (OMPC). Protéines et polysaccharides doivent, dans ce cas, être au préalable activés pour être associés et former une vésicule vaccinale (« vesicle vaccine »⁵¹).

- **La méthode chimique de couplage du polyside sur la protéine.** Le couplage du polyside sur la protéine requiert l'activation chimique du polysaccharide et quelquefois aussi de la protéine. L'activation du polyside dépend étroitement **de la nature et de la structure** de celui-ci. Elle tiendra compte des fonctions réactives de type aldéhyde, carboxyle ou amine, présentes dans le polyside. Elle devra également préserver l'intégrité des épitopes impliqués dans la réponse protectrice. Côté protéine, les fonctions réactives les plus utilisées sont les fonctions amine.

- **Le taux de greffage.** Le rapport protéine/polyside est aussi un facteur-clé pour l'induction optimale de la réponse immunitaire. La densité trop élevée d'osides masquant les épitopes protéiques, un compromis doit être trouvé entre accessibilité aux épitopes protéiques et saccharidiques.

TABEAU 2| Caractéristiques des vaccins polysidiques conjugués commercialisés contre les méningites à *Haemophilus influenzae* de sérotype b (Hib).

| Vaccins | Autorisation (AMM) | Nombre d'injections | Paramètres de construction | | | | |
|---|--------------------|---------------------|---|-----------------------|----------------------------|--------------------|-----------------------|
| | | | Protéine porteuse | Composante osidique | Site de couplage | Bras espaceur | Chimie de couplage |
| PRP-D (ProHibit [®]) | >18 mois (1987) | | Anatoxine diphtérique | PC natif | points d'attache multiples | bras à 6 carbones | activation au CNBr |
| HbOC (Hib-Titer [®]) | >2 mois (1990) | 3 | Mutant de l'anatoxine diphtérique (CRM197) | fragments du PC natif | point d'attache unique | / | amination réductrice |
| PRP-OMP (PedVax Hib [®]) | >2 mois (1990) | 2 | Protéine de la membrane externe de <i>N.Meningitidis</i> groupe B | PC natif | points d'attache multiples | bras à 14 carbones | fonctionnalit é thiol |
| PRP-T (Act-Hib [®] ou Hiberix [®]) | >2 mois (1992) | 3 | Anatoxine tétanique | PC natif | points d'attache multiples | bras à 6 carbones | activation au CNBr |

La flexibilité dans la préparation du glycoconjugué entraîne néanmoins **des difficultés de mise au point**. L'influence des différents paramètres est mal maîtrisée et difficile à établir en raison de leur **interdépendance**. De plus, la méthode employée dépendra aussi des propriétés intrinsèques du polysaccharide. Taille, structure et fonctions réactives du polyside orienteront la stratégie de construction du conjugué. Le procédé élaboré **ne sera pas forcément extrapolable à un autre polysaccharide**. Il en résulte des études au cas par cas. En particulier, la conception de vaccins à n valences est très contraignante. Elle implique de construire n vaccins monovalents pour lesquels il sera préférable d'utiliser des protéines porteuses. Cependant, l'emploi du peptide PADRE pourra potentiellement lever cette double contrainte : ne nécessitant qu'un seul couplage chimique, différents mélanges de PC pourraient être conjugués en une seule étape au peptide, avec lequel aucun phénomène de suppression épitopique n'a été observé. Un tel procédé simplifierait largement l'approche.

Ces procédés de préparation de vaccins de 2^{ème} génération peuvent également être appliqués au couplage d'haptènes isolés **à partir du LPS de bactéries pathogènes** comme *S. flexneri*, *V. cholerae* ou *S. typhi*. Dans ce cas, le polysaccharide spécifique (O-PS) du LPS, est assimilable à un haptène et doit donc être couplé à une protéine porteuse pour être immunogène. Au préalable, le lipide A (endotoxine) présent dans le LPS doit être impérativement éliminé^{52, 53} ou masqué⁵⁴ en raison de sa très forte toxicité. Cette étape reste très délicate et difficile à optimiser. Aucun vaccin

de 2^{ème} génération utilisant le LPS ou le LOS comme haptène n'est encore commercialisé. Toutefois, différentes équipes s'intéressent à cette approche pour lutter contre une dizaine d'espèces pathogènes¹⁷. Les travaux les plus avancés concernent les vaccins de 2^{ème} génération prévenant les infections par *Shigella*, actuellement en phase II de développement clinique (cf. partie II.D.2, page 44).

L'engouement pour les vaccins de 2^{ème} génération est manifeste. Utilisés sans adjuvant, ces vaccins ne requièrent qu'une faible quantité de polysaccharides (~10µg par dose vaccinale). Cependant, leur conception est plus complexe que celle des vaccins de 1^{ère} génération et leur coût environ 10 fois plus élevé (de l'ordre de \$1 à \$50 par dose suivant la complexité et le marché potentiel offert par le vaccin)¹⁷. L'optimisation des voies d'obtention doit être poursuivie pour accroître leur potentiel vaccinal et commercial.

B.4. Les Vaccins Semi-Synthétiques Conjugués PS-Protéine : 3^{ème} génération de vaccins polyosidiques

Les vaccins de 3^{ème} génération ou semi-synthétiques dérivent des vaccins de 2^{ème} génération. Ils reposent sur l'utilisation des propriétés immunogènes d'une glycoprotéine obtenue par couplage d'une protéine porteuse (non synthétique) et d'un haptène osidique **synthétique**, mimant généralement un motif antigénique naturel. L'utilisation **d'haptènes chimiquement définis** permet de contrôler la taille et la structure du polyoside et offre la possibilité d'explorer la relation entre **la taille des antigènes saccharidiques et leur performance vaccinale**.

Ces deux dernières décennies, les voies d'obtention des dérivés osidiques susceptibles d'être employées pour la synthèse d'haptènes oligosaccharidiques se sont considérablement développées et diversifiées. Elles sont au nombre de quatre :

- (a) Synthèse chimique multi étapes
- (b) Synthèse enzymatique
- (c) Synthèse chimio-enzymatique : combinaison de (a) et (b)
- (d) Synthèse et extraction à partir d'une usine cellulaire

- **La synthèse chimique multi-étapes**

La synthèse chimique d'oligosaccharides, entreprise depuis plus d'un siècle, est longtemps restée fastidieuse^{55, 56}. En effet, elle requiert **de multiples séquences d'activation/protection/**

glycosylation/déprotection/purification de façon à assembler régiosélectivement les intermédiaires réactionnels qui conduiront à l'oligosaccharide cible. Dans les années 90, les progrès accomplis dans les méthodes de glycosylation et de combinaison de groupes protecteurs ont rendu possible l'accès à des oligosaccharides synthétiques de 15 résidus⁵⁵. Malgré tout, ces synthèses sont longues à mettre en œuvre et restent peu adaptées à des productions à grande échelle. Deux avancées majeures réalisées dans le domaine de la glycochimie et déjà exploitées par des industriels devraient faciliter à l'avenir l'accès à ce type de composés. La première fait appel à des synthèses automatisables basées sur l'assemblage de synthons fonctionnalisés sur phase solide^{57, 58} (**FIGURE 7A**). Créée en 2002, *Ancora pharmaceuticals* (MA, USA) bénéficie d'une plate-forme mettant à profit cette technologie. Cette société réalise **des synthèses rapides** (en moins d'un jour) **et extrapolables à l'échelle industrielle** (de 10g à 1kg) d'oligosaccharides complexes⁵⁹, disposant ainsi d'une cadence de production 500 fois plus élevée que celle des méthodes classiques de glycosylation en solution⁶⁰. La deuxième méthode implique des synthèses "One-Pot" en solution basée sur les réactivités différentielles de donneurs thioglycosides^{61, 62}. A partir d'une base de données rassemblant les différents synthons et leurs réactivités relatives, le **programme "Optimer" conçoit la synthèse optimale d'un oligosaccharide (sélection et ordre d'assemblage des synthons) en une seule étape (FIGURE 7B)**. Développée par *Optimer Pharmaceutical* (CA, USA), la technologie **OPopS™ (Optimer Programmed One-Pot Synthesis)** offre les mêmes avantages (cadence et quantité) que la synthèse automatisée d'oligosaccharides. Ces nouveaux outils de glycosylation ont fait leur preuve mais restent encore marginaux.

Utilisée pour la construction de conjugués ciblant les sérotypes 3 et 6B de *S. pneumoniae* ainsi que le sérotype 1b de *H. influenzae*, la synthèse chimique a permis de générer une gamme d'oligosaccharides de taille croissante mimant des fragments (ou sous-fragments) de leur PC et d'évaluer leur potentiel immunogène⁶³⁻⁶⁵. Immunogènes chez la souris, des structures de taille minimale et réduite par rapport à celles intervenant dans les vaccins de seconde génération ont été identifiées.

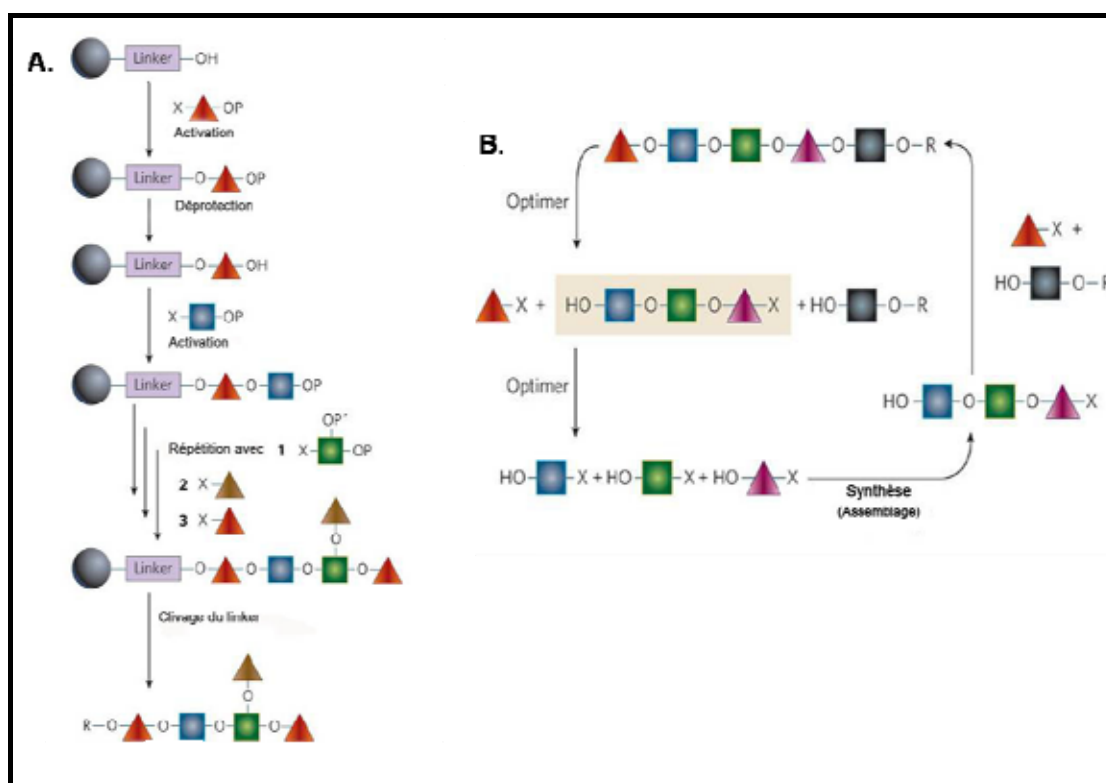


FIGURE 7| Synthèses chimiques automatisées
(D'après Seeberger et coll.¹²)

A- Assemblage automatisé en phase solide de 5 synthons monomériques sur une résine polymérique (cercle noir) par l'intermédiaire d'un linker. Le cycle, qui se compose d'étapes d'activation et de déprotection, est répété 5 fois pour l'assemblage du pentasaccharide. A la fin, le linker est clivé de manière à libérer le produit désiré.

B- Synthèse automatisée en solution utilisant la technologie OPOps® pour la synthèse d'un pentasaccharide. Un programme informatique sélectionne les synthons monomériques appropriés selon leur réactivité relative pour avoir le meilleur rendement lors de l'assemblage conduisant à l'oligosaccharide final. Les triangles et carrés colorés représentent les différents synthons. P et P' sont les groupes protecteurs temporaires. R est un résidu hydrocarboné qui sera fonctionnalisé. X est un groupe partant.

• La synthèse enzymatique

Parallèlement aux méthodes chimiques, la synthèse enzymatique est reconnue comme étant une alternative performante pour la synthèse d'oligosaccharides^{62, 66}. **Naturellement stéréo- et régiosélective**, la voie enzymatique a recours à **des schémas de synthèse simplifiés**, évitant toutes les étapes complexes de protection/déprotection de la chimie traditionnelle. Impliquées dans la biosynthèse naturelle des glucides ou glycoconjugués (glycoprotéines et glycolipides), les glycosyltransférases et les glycosidases se révèlent être les deux classes d'enzymes les plus adaptées au développement d'une telle approche. Cependant, les premières catalysent la synthèse de glucides ou glyco-conjugués à partir de sucres nucléotidiques, substrats rares et coûteux. De plus, très peu sont commercialisées. Les deuxièmes utilisent des substrats moins coûteux. Elles sont aussi plus faciles d'accès mais parfois peu efficaces en réaction de synthèse.

D'autre part, la spécificité des catalyseurs a ses revers. En effet, elle limite les réactions faisant intervenir des substrats non naturels. Ces dernières années, le **remodelage d'enzymes basé sur**

des techniques d'ingénierie aussi bien rationnelle qu'aléatoire voire semi-rationnelle, s'est montré extrêmement prometteur pour conférer de nouvelles spécificités aux enzymes⁶⁷. Ces approches et leurs avancées seront plus largement décrites en *partie III.C*.

Les étapes enzymatiques, une fois sélectionnées, peuvent être facilement intégrées dans des procédés de synthèses automatisées. Aucune synthèse purement enzymatique d'oligosaccharides mimant des haptènes anti-bactériens n'a été décrite à ce jour. En effet, la synthèse enzymatique d'un oligosaccharide composé de n résidus différents implique la formation de $(n-1)$ liaisons osidiques et par conséquent, l'action de $(n-1)$ enzymes distinctes. Du fait de la complexité des motifs oligosaccharidiques bactériens, ces enzymes sont rarement disponibles. C'est pourquoi la voie chimio-enzymatique est l'alternative préférée.

- **La synthèse chimio-enzymatique**

Elle **combine la flexibilité de la synthèse chimique et la haute sélectivité de la synthèse enzymatique en intégrant une ou plusieurs étapes enzymatiques dans une voie de synthèse chimique**. Le plus souvent cette étape est intégrée en fin de synthèse. Cette méthodologie a été mise en œuvre pour la préparation d'une série de fragments représentatifs de l'unité répétitive du PC chez *S. pneumoniae* sérotype 14. L'étape enzymatique met en œuvre une β -(1-4) galactosyltransférase qui catalyse le greffage d'une unité galactose sur des intermédiaires oligosaccharidiques linéaires synthétisés chimiquement^{68, 69}. La combinaison chimio-enzymatique a été également proposée pour la synthèse d'un pentasaccharide, composant du LOS de *N. meningitidis*. Le pentamère est obtenu en faisant agir successivement 3 enzymes sur du lactose préalablement dérivatisé *via* 4 étapes chimiques⁷⁰. Dans ce cas, les étapes enzymatiques prédominent. La simplicité et le rendement global (80%) de cette synthèse illustrent l'efficacité de l'approche.

- **La synthèse et extraction à partir d'une usine cellulaire**

L'utilisation de microorganismes génétiquement modifiés, nommés "usines cellulaires" s'ils sont utilisés en croissance, apparaît comme une méthode très efficace pour la production d'oligosaccharides à grande échelle⁷¹. Il s'agit de modifier le microorganisme (généralement *Escherichia coli*) de façon **à orienter certaines de ses voies métaboliques vers la synthèse de l'oligosaccharide d'intérêt**. Les gènes codant pour les glycosyltransférases impliquées dans la biosynthèse de l'oligosaccharide sont introduits dans la bactérie dite alors recombinante et devront être exprimés sous une forme active. Si nécessaire, les gènes requis pour la synthèse de donneurs de type nucléotide-sucres sont aussi intégrés dans le génome ou dans un vecteur approprié. Le substrat de départ ou "accepteur" doit pouvoir être internalisé efficacement dans les cellules sans avoir subi de modification (hydrolyse) ou être produit par la cellule si elle est pourvue des

enzymes adéquates. Ce système offre l'avantage d'un accès direct aux nucléotide-sucres sans avoir recours à l'isolement d'enzymes. Des niveaux de production de trisaccharide de l'ordre de 200g/L⁷² ont été décrits. Néanmoins, cette voie a été expérimentée principalement pour des structures osidiques simples (di- ou trisaccharidique) et reste encore peu décrite pour la synthèse d'oligosaccharides plus complexes.

Conclusion. De la première à la troisième génération, les vaccins polysaccharidiques ont gagné en qualité et en efficacité. Ils répondent de mieux en mieux aux critères de recevabilité des autorités de la santé. De plus, les domaines de la glycochimie et de la glycobiologie bénéficient aujourd'hui de nouveaux outils de synthèse dédiés à la conception d'oligosaccharides. Tous ces progrès technologiques sont en faveur du développement des vaccins de 3^{ème} génération et devraient être à l'origine de la quatrième génération de vaccins entièrement synthétiques. Cependant, **seul un type de vaccin de 3^{ème} génération** a été commercialisé à ce jour, il s'agit d'un vaccin anti-Hib¹ synthétisé chimiquement. Composé de 8 unités répétitives en moyenne, l'oligosaccharide est obtenu en une unique étape (au lieu de 16) de polycondensation du monomère préfonctionnalisé (**FIGURE 8**). Ce glycoconjugué, disponible à Cuba sous le nom de Quimi Hib[®] depuis 2004, est déjà utilisé dans un programme de vaccination de masse chez l'enfant.

Par ailleurs, l'équipe de L. Mulard de l'Institut Pasteur concentre ses efforts de recherche sur la production de vaccins de 3^{ème} génération prévenant la shigellose. C'est dans ce cadre que s'inscrivent nos travaux qui ont pour objectif d'évaluer le potentiel de nouvelles voies de synthèse chimio-enzymatique d'haptènes oligosaccharidiques.

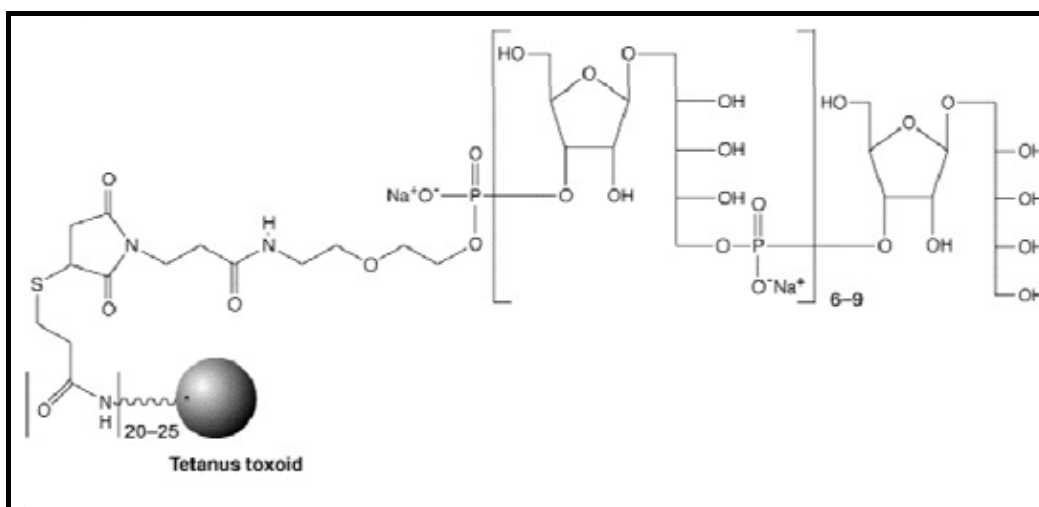


FIGURE 8 | Structure du premier vaccin commercialisé semi synthétique anti-Hib.

(D'après Roy et coll.¹³)

L'oligomère, comprenant de 6 à 9 unités, est doté d'un bras espaceur permettant un couplage sur l'anatoxine tétanique par construction sélective d'une liaison thioéther

PARTIE II : *SHIGELLA*, NOTRE MODÈLE D'ÉTUDE

La shigellose ou dysenterie bacillaire tue chaque année des millions de jeunes enfants dans les pays en voie de développement. Après plus de 60 ans de recherche, **aucun vaccin n'est encore disponible** pour lutter contre cette maladie, causée par une bactérie très complexe aux nombreux sous-types : *Shigella*. Cette deuxième partie propose de faire le bilan des connaissances acquises sur la maladie qui ont permis d'orienter les recherches vers des stratégies vaccinales adaptées et destinées à une couverture multi-sérotype⁷³⁻⁷⁵.

A. *Shigella*, agent étiologique de la shigellose

A.1. L'organisme pathogène

Shigella est un genre bactérien appartenant à **la famille des entérobactéries** (*Enterobacteriaceae*) qui sont définies classiquement par les sept critères suivants : bacilles à Gram négatif, aéro-anaérobies, mobiles ou immobiles, facilement cultivables, fermentant le glucose, réduisant les nitrates en nitrites et dépourvus d'oxydase. Présentes en tant qu'hôtes normaux ou pathologiques du tube digestif de l'homme et des animaux, ces bactéries sont communément appelées bactéries entériques (du grec *enterikos*, appartenant à l'intestin). Cette famille regroupe plus de 40 genres bactériens, dont le plus connu est *Escherichia* et son espèce prédominante : *Escherichia coli*. Les données récentes de la taxonomie moléculaire (basées sur l'hybridation de l'ADN) et du séquençage de gènes multiples indiquent que toutes les espèces et tous les sérotypes de *Shigella* (excepté le sérotype 13 de *S. boydii*) constituent avec, *E. coli*, une même espèce génomique. **Les *Shigella* sont génétiquement des variétés invasives de *E. coli***, qui sont non mobiles, anaérogènes, auxotrophes, porteuses de plasmides invasifs, métaboliquement moins actives et strictement **adaptées au parasitisme du colon de l'homme**⁷⁶. La classification en deux genres distincts est conservée en référence aux travaux antérieurs.

A.2. La Maladie

La shigellose est une **infection rectocolique aigüe invasive**. Ce n'est pas la plus fréquente des maladies diarrhéiques mais, dans sa forme typique dysentérique, elle est sans aucun doute la plus sévère: la shigellose, dont le nombre de cas est estimée chaque année à 150 millions, **est responsable d'un million de décès annuels**. 60% de ces décès surviennent chez les enfants de moins de 5 ans⁷⁷.

Les shigelles sont pathogènes uniquement pour l'homme et les autres primates. **L'homme est leur seul réservoir naturel**. Le plus souvent, **la transmission est directe**, du malade à son

entourage. Elle est parfois indirecte par l'intermédiaire de l'eau non traitée, d'aliments souillés par les déjections d'un malade ou par les mouches qui peuvent transporter les shigelles de manière passive. **Extrêmement infectieux**, 10 à 100 bacilles suffisent à induire la maladie⁷⁸. Répandue dans le monde entier, **la shigellose sévit surtout dans les régions aux conditions d'hygiène précaire** (insuffisance de l'assainissement et absence d'accès à l'eau salubre) **et à forte promiscuité** (guerres, camps de réfugiés, collectivités d'enfants). Dans les pays tropicaux en voie de développement, la dysenterie bacillaire associée à la malnutrition est responsable d'une forte morbidité et mortalité.

A.3. Historique et épidémies causées par la shigellose

A la fin du XIX^{ème} siècle, de violentes épidémies de dysenterie bacillaire accompagnées d'un taux de mortalité élevé (20%) touchent le Japon. Un jeune microbiologiste japonais, Kiyoshi Shiga isole alors une bactérie provenant des selles de malades et décrit, **en 1898, la découverte du premier membre du genre *Shigella***, connu aujourd'hui sous le nom de *Shigella dysenteriae* 1. Au cours des 40 ans qui suivirent cette découverte, les 3 autres espèces du groupe sont identifiées. Le terme *Shigella* est proposé en 1930. Les noms des quatre espèces connues aujourd'hui : *Shigella dysenteriae*, *S. flexneri*, *S. boydii* et *S. sonnei* ont été attribués en référence aux chercheurs impliqués dans leur découverte : Shiga, Flexner, Boyd et Sonne. Ces 4 espèces (ou groupes) sont elles-mêmes divisées en sous-espèces sur la base de caractères biochimiques et antigéniques (de type Ag-O)⁷⁹ (TABLEAU 4).

TABLEAU 3 | Espèces et sérotypes de *Shigella* (D'après Levine et coll.⁷⁵)
(* Il y a 14 sérotypes et sous-sérotypes de *Shigella flexneri*)

| Espèces <i>Shigella</i> (groupe) | Nombre de sérotypes et sous-sérotypes | Niches épidémiologiques les plus importantes | Sérotypes qui nécessitent un vaccin |
|--|---------------------------------------|---|-------------------------------------|
| <i>S. dysenteriae</i> (groupe A) | 15 | épidémies et pandémies sévères de <i>S. dysenteriae</i> 1 | <i>S. dysenteriae</i> 1 |
| <i>S. flexneri</i> (groupe B) | 14* | sous forme endémique dans les pays en voie de développement | les 14 sérotypes |
| <i>S. boydii</i> (groupe C) | 20 | foyers endémiques dispersés dans les pays en voie de développement | Aucun |
| <i>S. sonnei</i> (groupe D) | 1 | diarrhée des voyageurs/ sous forme endémique dans les pays développés | <i>S. sonnei</i> |

De toutes les espèces de *Shigella*, *S. dysenteriae* 1 (***Sd1***) est particulièrement **menaçante en raison de la gravité de la maladie et de son potentiel épidémique**. Les pandémies de *Sd1* ont tendance à se déclarer à des intervalles d'une dizaine d'années. Survenue à la fin des années 60, l'épidémie envahissant tous les pays d'Amérique centrale localisés à proximité du foyer situé au Guatemala fut particulièrement dévastatrice (112 000 cas/ 13 000 décès). Durant les 3 décennies

qui suivirent, d'autres pandémies ont été recensées en Asie du sud (années 70), en Afrique centrale (années 80) et en Afrique de l'est (années 90) (FIGURE 9).

S. flexneri, est responsable des formes endémiques de la maladie touchant l'Afrique centrale, le sud-est asiatique et l'Inde. *S. sonnei*, moins virulente, est prévalente dans les pays industrialisés. Enfin, la shigellose induite par *S. boydii* et les sérotypes de *S. dysenteriae* autres que le 1 est peu fréquente (bien qu'une augmentation de cas due à *S. boydii* ait été rapportée dans certains foyers).

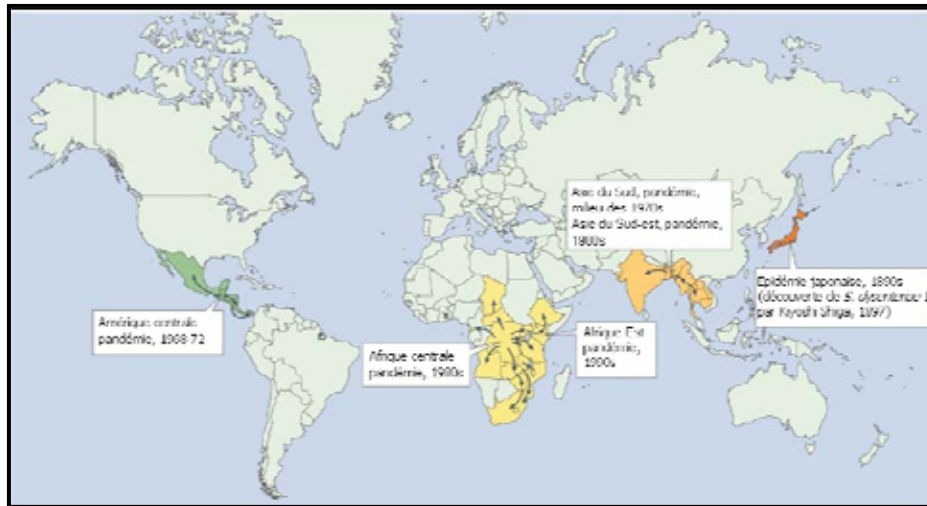


FIGURE 9] Pandémies de dysenteries de Shiga.
(D'après Levine et coll.⁷⁵)

Le premier membre du genre *Shigella* identifié, *Shigella dysenteriae* 1, fut isolé par Kiyoshi Shiga durant les épidémies affectant le Japon à la fin des années 1890. Débutant en Amérique centrale à la fin des années 1960, les pandémies dues à *S. dysenteriae* 1 ont décimé la population affectant plus 112 000 personnes et causant plus de 13 000 décès. Les pandémies suivantes touchèrent l'Asie et l'Afrique durant les 3 décennies qui se succèdent.

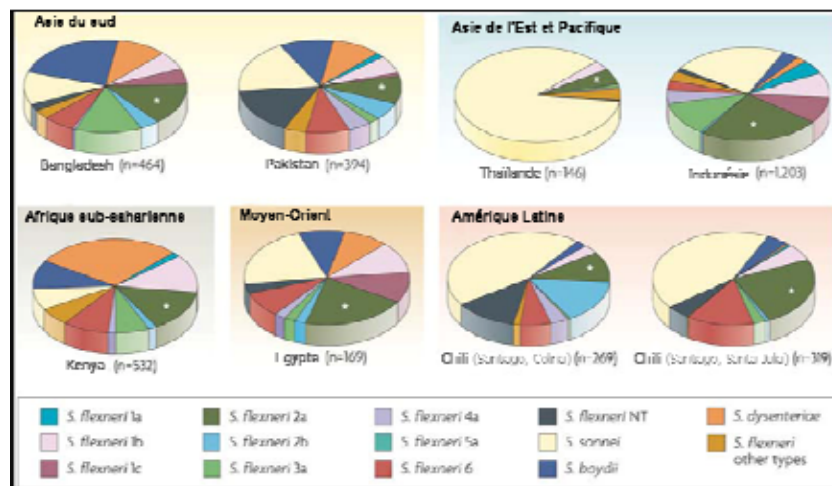


FIGURE 10] Distribution des espèces de Shigella provenant d'études menées en Amérique Latine, Asie et Afrique.
(D'après Levine et coll.⁷⁵)

La surveillance des shigelloses est réalisée durant des périodes d'au moins 2 ans en effectuant l'identification et le sérotypage des souches de *Shigella*. Dans tous les pays, *S. flexneri* 2a (marquée d'un asterisque) est le sous-sérotipe de *S. flexneri* le plus représenté. Les pays les plus développés (Chili et Thaïlande) montrent une proportion significative de cas dus à *S. sonnei*. Dans chaque pays, 3-4 sérotypes prédominent à 75%. Cependant, les sérotypes prévalents de *S. flexneri* autres que 2a diffèrent de pays à pays. Un vaccin qui conférerait une protection élevée contre *S. dysenteriae* 1, *S. sonnei* et tous les sérotypes de *S. flexneri* constituerait un vaccin à large spectre. NT, non typable.

B. Mécanisme de l'infection

En « disséquant » les mécanismes moléculaires et cellulaires mis en jeu lors du processus d'invasion de la bactérie, de nombreux travaux ont permis durant ces 2 dernières décennies d'établir le schéma physiopathologique de cette maladie et de pointer ainsi des cibles thérapeutiques.

Les études physiopathologiques récentes se sont portées essentiellement sur *S. flexneri*. Le mécanisme invasif de *Shigella* est fondé sur la propriété des bactéries à pénétrer et à se disséminer dans l'épithélium et sur leur capacité à induire une réaction inflammatoire intense marquée par une sévère destruction tissulaire. Ainsi le processus d'infection se décompose en 3 étapes:

1- Invasion de l'épithélium par internalisation des bactéries⁸⁰

Les déterminants de l'entrée (et de la dissémination) des bactéries dans les cellules épithéliales sont codés par un plasmide de virulence de 213kb codant, en particulier, pour une région de 30 kb qui spécifie un appareil de sécrétion de type III (l'appareil Mxi-Spa). Cet appareil, très répandu parmi les bactéries Gram négatif, permet l'injection à l'intérieur des cellules, de protéines IpaA, IpaB, IpaC et IpaD appelées invasines (ou Invasion plasmid antigen)^{81, 82}. Ces sécrétions déclenchent alors une réorganisation massive du cytosquelette ainsi que la formation de filaments d'actine. De longues projections membranaires issues de ces filaments vont entourer la bactérie conduisant finalement à son internalisation au sein d'une vacuole de phagocytose (FIGURE 11).

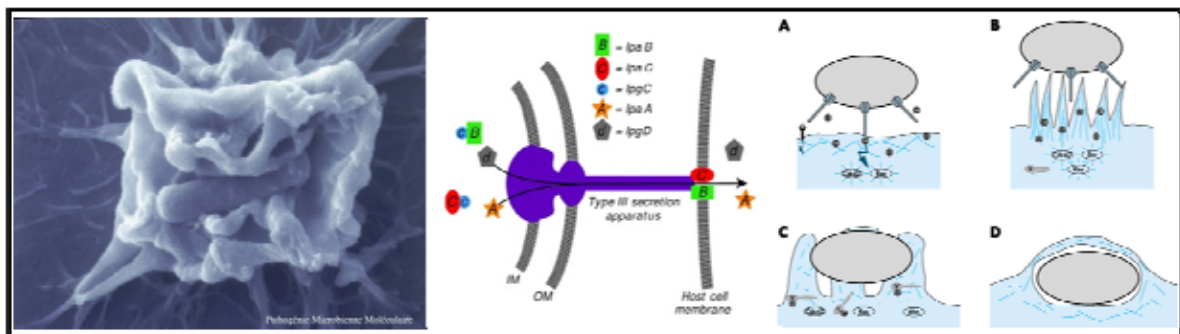


FIGURE 11| Invasion de l'épithélium par les shigelles

1- Image de microscopie électronique montrant l'entrée de *Shigella* dans une cellule *Hela*.

2- Schéma de l'appareil de sécrétion de type III de *Shigella*

3- Internalisation induite par les bactéries du genre *Shigella*. (A) Au contact de la cellule, *Shigella* sécrète des protéines (IpaA, IpaB, IpaC cercles gris) par son appareil de sécrétion de type III. Certaines Ipa, sécrétées dans le milieu extracellulaire lient des récepteurs membranaires et activent des voies de transduction du signal. La protéine IpaC sécrétée dans le cytoplasme active les protéines Rac, Cdc42 et Rho, entraînant une forte polymérisation et stabilisation du cytosquelette d'actine (B). La protéine IpaA lie et active la vinculine (V), conduisant à une dépolymérisation de l'actine. Les protéines Rho, Rac et Cdc42 sont inactivées, et des feuillettes membranaires entourent la bactérie (C). Après fusion des membranes, *Shigella* se retrouve dans une vacuole de phagocytose (D).

Source Images: http://www.pasteur.fr/recherche/unites/Pmm/contents/fr/04_banque-images/index.html#
http://www.xray.chem.rug.nl/Home_r.htm

2- Apoptose des macrophages : Déstabilisation de l'épithélium

Une fois internalisées, les shigelles lysent la vacuole d'endocytose entraînant **l'apoptose des macrophages infectés**, une réaction inflammatoire au niveau des sites d'invasion et le recrutement de leucocytes polymorphonucléaires (PMN) en réponse à l'inflammation. Ces leucocytes vont infiltrer le site infecté et faciliter l'entrée de nouvelles bactéries en ouvrant les espaces paracellulaires et de ce fait, **déstabiliseront l'épithélium. L'intense destruction tissulaire observée lors de la shigellose résulte probablement plus de l'activité des PMN activés que des bactéries elles-mêmes**, ces dernières agissant comme stimulus et facteur d'entretien de la réaction inflammatoire.

3-Motilité des bactéries : colonisation de l'épithélium⁸³

Les shigelles, libres dans le cytoplasme vont pouvoir se multiplier. Immobiles en culture, elles peuvent se déplacer dans le cytoplasme. En effet, la motilité intracellulaire de *Shigella* est médiée par la protéine IcsA (120 kDa) codée par le plasmide de virulence; cette protéine de membrane externe localisée à un pôle de la bactérie induit la polymérisation de l'actine, l'assemblage et la croissance polarisés de ces filaments. La bactérie est alors propulsée dans le cytoplasme cellulaire à la vitesse de polymérisation de l'actine (6 à 60µm/min). Ce mouvement permet le passage de la bactérie de cellule à cellule au sein d'une monocouche épithéliale. **Ce processus aboutit à une colonisation rapide et efficace de l'épithélium caractéristique de l'infection par *Shigella*.**

Outre le caractère invasif de *Shigella*, certaines souches sécrètent aussi **des toxines**, renforçant encore leur virulence. La toxine dysentérique ou Shiga toxine, codée par le gène *stx* est exclusivement produite par *S. dysenteriae* 1. Sa présence semble provoquer la destruction des cellules endothéliales et pourrait expliquer le syndrome hémolytique-urémique (SHU) associé à l'infection par *S. dysenteriae* 1.

C. Réponse immunitaire induite au cours de la shigellose

C.1. Réponse innée

La réponse innée correspond à la première ligne de défense contre l'infection bactérienne. Elle permet de localiser l'infection et joue également un rôle dans l'induction de la réponse spécifique. Dans le cas de *Shigella*, cette réponse de type inflammatoire comporte deux étapes. Au cours de la première phase, le processus inflammatoire est localisé au niveau des follicules lymphoïdes, qui représentent les sites d'entrée des bactéries dans l'épithélium intestinal. Son déclenchement est principalement lié **à la libération massive des cytokines pro-inflammatoires, IL-1β et IL-18** par les macrophages pré-apoptotiques⁸⁴.

Dans une seconde phase, ce processus s'étend à l'épithélium villositaire, conduisant à une destruction tissulaire intense. En effet, l'invasion des cellules épithéliales par les bactéries active **une voie de signalisation nommée NF-kB dont le facteur protéique activera les gènes pro-inflammatoires** des cellules de l'hôte. Il en résulte une production de cytokines pro-inflammatoires, en particulier l'IL-8⁸⁵. Toutefois, les shigelles parviennent à **moduler cette réponse inflammatoire** en bloquant la voie NF-kB à plusieurs niveaux^{86,87}.

A un stade ultérieur, les bactéries sont finalement tuées par les PMN qui jouent donc un double rôle dans le processus physiopathologique: d'abord, en faveur des bactéries en permettant leur entrée après effraction de la barrière intestinale, puis en faveur de l'hôte en les éliminant.

C.2. Réponse spécifique

En réponse à l'infection par *Shigella*, l'hôte va développer des réponses immunitaires cellulaires et humorales. Les cibles et les effecteurs de la réponse humorale, aussi bien que leur rôle dans la protection, ont été largement étudiés alors que peu de données sont disponibles en ce qui concerne la réponse cellulaire induite par ce pathogène.

C.2.a. Réponse cellulaire

Certaines études mettent en évidence une production élevée d'interférons IFN- γ et d'interleukine 10, suggérant la participation de cellules T auxiliaires de type 1 en réponse à une infection par *Shigella*^{88, 89}. Cependant, bien que l'activation de ces cellules T semble apparaître localement, la réponse cytotoxique et son rôle éventuel dans la protection restent mal connus et controversés⁹⁰.

C.2.b. Réponse humorale

Lors d'une infection induite par *Shigella*, l'hôte met en place deux types de réponses immunitaires humorales : une réponse humorale systémique médiée par la production d'anticorps sériques, principalement des immunoglobulines d'isotype G (IgG) et une réponse humorale locale caractérisée par la production d'immunoglobulines d'isotype A (IgA) sécrétées locales. Ces anticorps sont dirigés **essentiellement contre la partie polysidique du LPS (Ag-O)** et mineurairement contre des antigènes protéiques (codés par le plasmide de virulence : Ipa, VirG et IcsA). Si leur contribution individuelle n'est pas bien établie, il est clair en revanche que les anticorps anti-Ag-O jouent un rôle clé dans la protection de l'hôte contre une ré-infection. La plupart des données épidémiologiques⁹¹ et expérimentales⁹² indiquent que **l'immunité protectrice est spécifique de l'espèce, et pour une espèce donnée, spécifique du sérotype. La protection est donc liée à la structure de la partie antigène O du LPS, structure qui est propre à chaque souche prévalente (FIGURE 12)**. Du fait de la diversité des structures

antigéniques, seul un vaccin multivalent pourra prétendre à l'éradication de la maladie (**FIGURE 10**). Ce point est un obstacle majeur au développement d'un vaccin efficace anti-*Shigella*. La fonction protectrice des anticorps protéiques, quant à elle, n'est pas encore clairement définie.

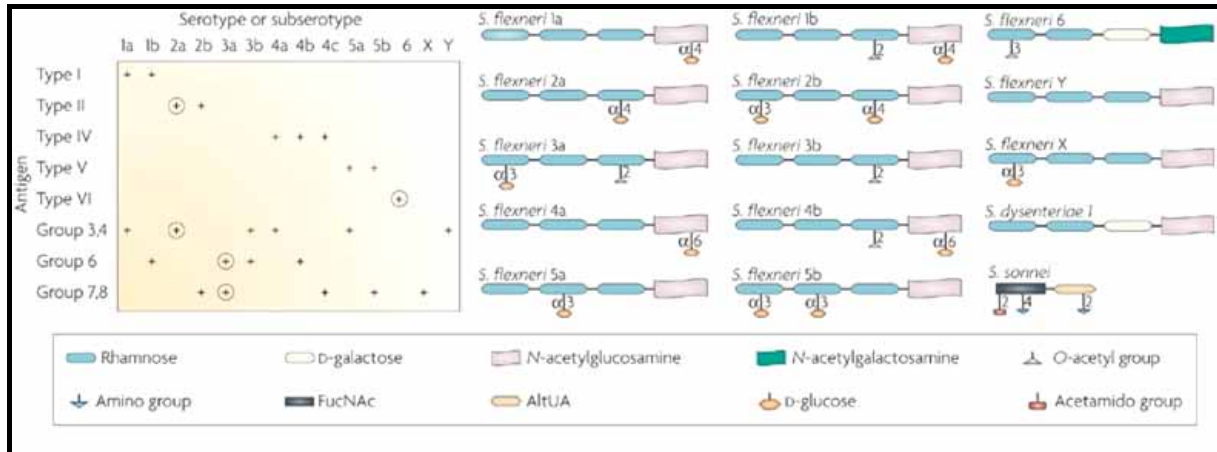


FIGURE 12| Antigènes Ag-O de *S. flexneri*, *S. dysenteriae 1* et *S. sonnei*
(D'après Levine et coll.⁷⁵)

Figure de gauche: Les antigènes Ag-O permettent de définir les espèces et sous-espèces (sérotypes) de *Shigella*. Le sérotypage de *S. flexneri* est complexe. La taxonomie actuelle reconnaît 14 sérotypes et sous sérotypes de *S. flexneri* qui sont définis par des antigènes de type (I, II, IV–VI) et de groupes (3, 4, 6, 7, 8)

Figure de droite: 13 des 14 sérotypes ou sous-sérotypes de *S. flexneri* (tous à l'exception de *S. flexneri* 6) partagent une structure commune: un tétrasaccharide linéaire composé de trois résidus rhamnose, liés en 2 ou 3 et d'un D-glucose N-acétylé lié en position 2. Ils se distinguent par le greffage d'un résidu D-glucose et/ou d'un groupe O-Acétyl à des positions variables sur le tétrasaccharide. L'Ag-O de *S. flexneri* 6 présente un D-galactose en tant que 3^{ème} résidu du tétrasaccharide et un D-Galactose N-acétylé en position terminale. La structure de *S. flexneri* 4c n'a pas encore été rapportée. Sont également présentées les structures de l'Ag-O de *S. dysenteriae 1* et *S. Sonnei*, qui devraient être inclus dans un vaccin multivalent.

D. Avancées des stratégies vaccinales (état de l'art)

Pour lutter contre la shigellose ou d'autres maladies infectieuses, les approches envisagées sont généralement de **type préventif et/ou thérapeutique**. A la différence des autres maladies diarrhéiques, la shigellose ne peut être traitée par la seule réhydratation. L'approche thérapeutique, basée sur **l'emploi d'antibiotiques dits de première ligne**, s'est révélée efficace par le passé mais se heurte de plus en plus à **l'augmentation de souches pharmacorésistantes**. Ainsi, dans les années 1940, *Shigella* a acquis une résistance aux sulfamides ; dans les années 1950 à la tétracycline et au chloramphénicol ; dans les années 70 à l'ampicilline et à la fin des années 80, au triméthoprime-sulfaméthoxazole⁹³. Aujourd'hui, elle oblige l'usage d'antibiotiques plus rares et chers (fluoroquinolones et céphalosporines de 3^{ème} génération) pour lesquels certaines souches semblent avoir développé une résistance. **L'approche préventive réserve aussi certaines limites**. Les faibles ressources des pays pauvres rendent peu vraisemblable l'amélioration des conditions sanitaires des zones endémiques. Les difficultés structurelles que rencontre le continent africain pour assurer son développement économique et la crise asiatique

sont autant de facteurs négatifs, mais **le poids de l'augmentation de la population reste certainement le facteur dominant.**

Les premières tentatives vaccinales datant des années 1940-1950 ont été basées sur l'utilisation de vaccins inactivés administrés par voie parentérale⁹⁴. Leur échec a contribué à entretenir la vision selon laquelle il n'y aurait jamais de vaccin contre la shigellose. Ce concept a prévalu jusqu'à récemment, en dépit des résultats prometteurs obtenus dans les années 1960, après vaccination orale avec des souches atténuées⁹¹.

En 1996, le **développement d'un vaccin anti-*Shigella* a été considéré comme une priorité par l'OMS**⁹³ qui a défini des critères pour leur conception incluant : l'immunogénicité suite à l'administration d'une seule dose vaccinante, la facilité d'administration, la stabilité de la préparation vaccinale dans des conditions variées de climat et l'absence d'effets secondaires. La mise au point d'un vaccin multivalent et unique est aussi fortement encouragée. Suite à ces recommandations, différentes stratégies vaccinales ont été investiguées qui sont porteuses de résultats prometteurs.

D.1. Les souches vivantes de virulence atténuée

De nombreuses souches de virulence atténuée ont été construites ces dernières années et certaines d'entre elles ont atteint un stade avancé dans les essais cliniques. Les stratégies d'atténuation, basées sur l'inactivation rationnelle de gènes précis, ont pour but de concevoir des souches **au pouvoir pathogène réduit**, mais qui restent **néanmoins capables de stimuler efficacement une réponse immunitaire protectrice**. Or, pour induire une telle réponse, les souches doivent conserver leur caractère invasif. D'autre part, il existe une étroite relation entre le processus d'invasion des cellules épithéliales par *Shigella* et la réaction inflammatoire entraînant une destruction tissulaire massive. Trouver un **bon équilibre entre tolérance et efficacité** se révèle donc être **la principale difficulté** à surmonter dans le développement de ce type de vaccin.

Deux types d'atténuation ont été ciblés dans ce but. Elles visent soit des gènes impliqués dans une voie métabolique essentielle **à la survie de la bactérie dans les tissus**, soit **des gènes de virulence**.

En 2006, l'OMS a dressé un bilan des vaccins candidats les plus prometteurs qui utilisent cette technique⁹⁵. Dans le cas de *S. flexneri* 2a, la souche SC602⁹⁶ et la gamme **CVD**⁹⁷⁻⁹⁹ développée par l'université de Maryland (University of Maryland Center for Vaccine Development, USA) font l'objet d'essais cliniques avancés. Ces essais consistent à vérifier l'innocuité et l'immunogénicité d'une dose croissante unique de la souche candidate et à définir, ainsi, une dose

tolérée et protectrice qui serait utilisée dans la composition vaccinale finale. L'ajustement de cette dose s'avère particulièrement délicate¹⁰⁰.

Des souches atténuées contre *S. dysenteriae* (SC599) et contre *S. sonnei* (WRSS1)^{101, 102} sont également disponibles. Ces différentes souches pourraient être à la base d'un vaccin oral trivalent contre la shigellose.

Les vaccins vivants de virulence atténués ont l'avantage de pouvoir être administrés par voie orale. De cette manière, cette approche permet, d'une part, de **délivrer une grande quantité et diversité d'antigènes**, et d'autre part de les **présenter directement aux structures lymphoïdes** associées à la muqueuse intestinale qui sont les sites inducteurs des réponses locales et systémiques. L'administration d'une seule dose peut être suffisante, ce qui représente un avantage supplémentaire. Cependant, il existe des obstacles majeurs au développement de cette approche. **L'influence de la délétion** introduite sur l'induction de la réponse immune spécifique **ne peut être prédite**. Il faut alors construire et tester de nombreux mutants de façon à trouver le meilleur équilibre entre atténuation et immunogénicité. Par ailleurs, la **réversion possible vers la virulence** nécessite la délétion de plusieurs gènes, alourdissant d'autant plus la construction de ces souches. La possibilité d'effets secondaires chez certains individus, notamment les personnes immunodéprimées est également problématique. Enfin, **le maintien de la chaîne du froid** requis pour ce type de vaccin reste peu adapté à leur utilisation dans les pays pauvres où l'infrastructure nécessaire est souvent inadéquate voire inexistante. D'autres approches s'avèrent nécessaires.

D.2. Les vaccins de 2^{ème} génération

Puisque le polysaccharide O-PS du LPS de *Shigella* est indispensable à la virulence de la bactérie et aux phénomènes immunitaires, le développement de vaccins sous-unités de type glyco-conjugués est la deuxième approche prometteuse. Des vaccins candidats sous-unités sont disponibles, en particulier un vaccin parentéral anti-*S. sonnei* composé de LPS détoxifié conjugué à un toxoïde. Les premières données cliniques sont encourageantes dans la mesure où ce vaccin a démontré une efficacité protectrice de 74% chez l'adulte lors d'une épidémie de *S. sonnei* dans un camp d'entraînement de l'armée israélienne⁴⁰. D'autres études menées sur *S. sonnei* et *S. flexneri* 2a utilisent de nouvelles protéines porteuses (rEPA et CRM9) dans le but d'accroître l'immunogénicité de la construction finale¹⁰³. Efficace chez l'adulte, cette stratégie a été par la suite validée chez les jeunes enfants¹⁰⁴.

D'autres approches particulières moins avancées existent, préparées à partir de protéosomes¹⁰⁵ (mélange de LPS et de protéines bactériennes de membrane externe) ou de ribosomes^{106, 107} (complexes non covalents de polysaccharide O et de particules ribosomales de *Shigella*).

L'utilisation de vaccins sous-unités permet d'éviter les difficultés rencontrées lors de la mise au point de vaccins vivants de virulence atténuée. Pourtant, d'autres limitations apparaissent : la **purification** des différents composés polysaccharidiques, le **manque de reproductibilité** (ou standardisation du vaccin) et le coût de production élevé sont autant de points à prendre en compte pour améliorer ce type de vaccins. Dans ce contexte, l'approche semi-synthétique, récemment validée pour *H. influenzae* (cf partie I.B.4, page 31), peut lever les limites des vaccins sous-unités.

D.3. Le recours aux Vaccins Semi-Synthétiques Conjugués PS-Protéine: Une nouvelle approche vaccinale prometteuse

Cette approche a été validée dans le cas de *S. dysenteriae* 1 et plus récemment de *S. flexneri* 2a. Concernant *S. dysenteriae* 1, l'unité répétitive de l'Ag-O est un tétrasaccharide¹⁰⁸ (**FIGURE 13**).

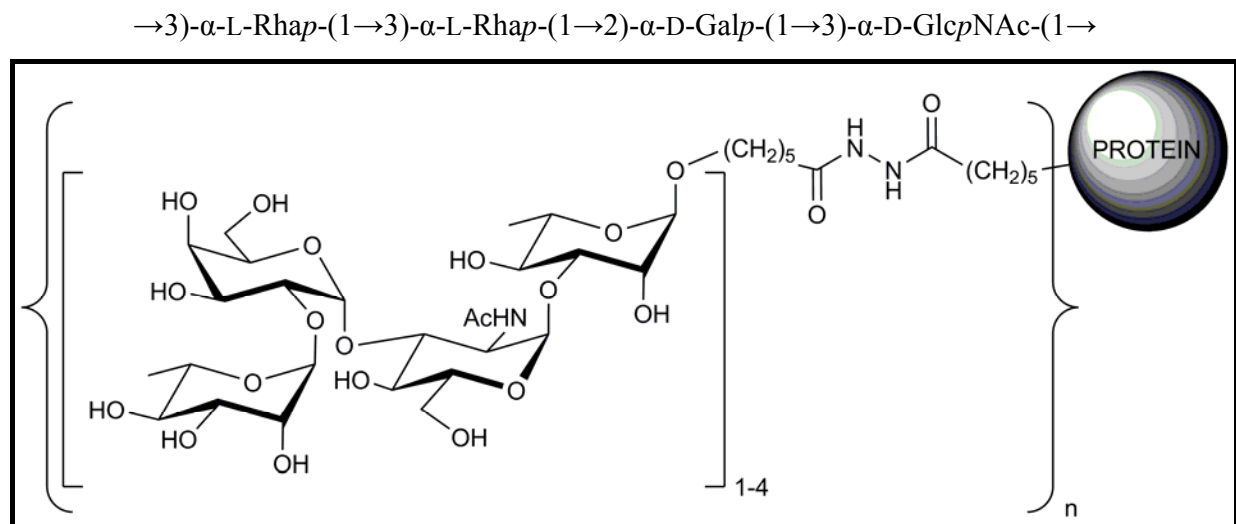


FIGURE 13| Glycoconjugués semi-synthétiques à potentiel vaccinal contre *S. dysenteriae* 1
(D'après Pozgay et coll.¹⁰⁹)

Des fragments tétra-, octa-, dodéca et hexadécasaccharidiques de l'Ag-O munis d'un bras espaceur sélectivement activable à leur extrémité réductrice ont été synthétisés¹⁰⁹ et couplés à la HSA (**H**uman **S**erum **A**lbumin) par la technique du point d'ancrage unique. Une famille de conjugués différant par la taille de la composante saccharidique et le taux de greffage a ainsi été construite. A l'exception du conjugué tétrasaccharidique, tous les conjugués se sont révélés être de meilleurs immunogènes chez la souris qu'un conjugué de deuxième génération utilisé comme référence. D'autre part, il a été observé une corrélation entre la taille de l'haptène saccharidique et leur densité sur la protéine⁵⁰. L'immunogénicité maximale des conjugués pourrait donc être optimisée en ajustant ces 2 paramètres et en permettant ainsi l'accessibilité aux épitopes aussi bien protéiques que saccharidiques. Dernièrement, l'importance du résidu terminal et sa contribution à

l'effet immunogène de l'haptène saccharidique ont été mises en évidence. Ainsi, à taille et taux de greffage similaires, les glyco-conjugués comportant un résidu *N*-acetyl-D-glucosamine ou D-galactose à leur extrémité non réductrice induisent une sécrétion d'IgG beaucoup plus élevée que ceux pourvus d'un résidu L-rhamnose¹¹⁰.

Concernant *S. flexneri* 2a, l'unité répétitive (UR) de l'Ag-O est un pentasaccharide composé d'un tétrasaccharide linéaire ramifié par un résidu D-glucose (**FIGURE 14**).

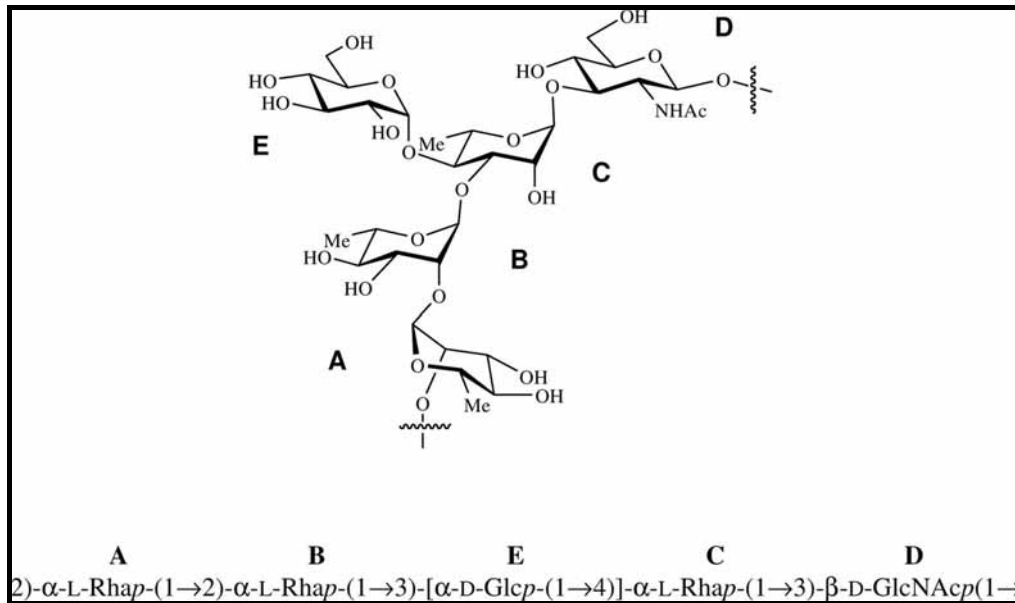


FIGURE 14 | Unité répétitive de *S. flexneri* 2a
(D'après Phalipon et coll.²)

Une large gamme (23) d'oligosaccharides synthétiques allant du mono au pentasaccharide dont les séquences sont obtenues par permutation circulaire des différents résidus de l'unité répétitive de l'Ag-O a été évaluée selon les techniques conventionnelles utilisées en ELISA d'inhibition. Le motif ECD s'avère être le motif déterminant pour induire une réponse protectrice. D'autre part, le tétrasaccharide ramifié B(E)CD apparaît comme l'épitope immunodominant. Enfin, l'élongation de la chaîne (du tétra au décasaccharide) améliore significativement la reconnaissance pour 4 des 5 anticorps testés². Les motifs ECD, B(E)CD, AB(E)CD¹¹¹ (=UR), [AB(E)CD]₂¹¹² (=2UR), et [AB(E)CD]₃¹¹² (=3UR), ont donc été synthétisés chimiquement et couplés à l'anatoxine tétanique (TT). Seuls, les conjugués comportant de une à trois unités répétitives induisent une production d'anticorps chez la souris. Cette réponse est croissante avec le nombre d'UR. Ainsi, le **pentadécasaccharide, dont l'immunogénicité est supérieure à celle du glycoconjugué de seconde génération** utilisé comme référence, s'avère être un bon candidat pour le développement d'un futur vaccin de troisième génération contre *S. flexneri* 2a.

Cette démarche, appliquée aux sérotypes les plus virulents de *Shigella*, permettrait d'élaborer à terme un vaccin multivalent répondant aux différents critères de pureté, de qualité, d'efficacité et

de reproductibilité. **C'est dans ce cadre que s'inscrit notre projet qui a pour objectif le développement d'une voie de synthèse chimio-enzymatique pour les haptènes 1b et 3a de *Shigella flexneri*** qui sont également des espèces prévalentes, responsables de la forme endémique de la shigellose (**FIGURE 10**).

La synthèse de ces haptènes oligosaccharidiques de *S. flexneri* 1b et 3a implique des réactions d' **α -glucosylation de la *N*-acétyl-D-glucosamine et du L-rhamnose**, difficilement réalisables par voie chimique. En revanche, ces réactions pourraient être catalysées plus aisément *via* des glyco-enzymes. Le projet ANR « Optigluc » dans lequel s'intègrent nos travaux propose de développer une stratégie de synthèse chimio-enzymatique employant des transférases natives ou construites « à façon » pour satisfaire les exigences de la voie de synthèse proposée par les chimistes. Le défi majeur du projet est de démontrer qu'il est possible de mettre à profit les technologies d'ingénierie enzymatique pour modifier la spécificité de ces enzymes et les adapter à la glucosylation de molécules naturellement mal reconnues et destinées à intégrer la voie de synthèse chimique. Dans notre cas, il s'agira de glucosyler des dérivés de la *N*-acétyl-D-glucosamine et du L-rhamnose, porteurs de substituants qui permettront de poursuivre par voie chimique la synthèse des oligosaccharides antigéniques. La glucosylation de molécules peut être envisagée par l'emploi de différents types d'enzymes qu'il convient maintenant de présenter.

PARTIE III : LES GLYCO-ENZYMES

Les glucides et les glycoconjugués sont des composants essentiels des systèmes vivants, qu'ils soient microbiens, végétaux ou animaux. Ils y sont présents sous des formes extrêmement variées et interviennent dans le stockage d'énergie, la structure des parois cellulaires, les phénomènes d'interactions ou de signalisation cellule-cellule, hôte-pathogène, et dans la glycosylation des protéines. Cette étonnante diversité repose sur l'action d'une multiplicité d'enzymes capables d'agir sur des molécules glucidiques et plus particulièrement d'intervenir dans la synthèse de motifs extrêmement variés.

Dans la classification proposée par B. Henrissat¹¹³, les enzymes responsables de la synthèse des sucres sont répertoriées dans deux classes distinctes : les **Glycoside-Hydrolases (GHs)** et les **Glycosyltransférases (GTs)**. Cette classification, dénommée classification CAZy ("Carbohydrate-Active enZymes") a d'abord été proposée en 1991 pour les glycoside-hydrolases¹¹³⁻¹¹⁶. Elle a ensuite été élargie aux autres classes d'enzymes actives sur les sucres (glycosyltransférases, carbohydrate-esterases et polysaccharide-lyases)¹¹⁷, ainsi qu'à leurs modules non catalytiques¹¹⁸. L'ensemble des données est disponible sur le serveur URL : <http://www.cazy.org/> régulièrement mis à jour. La classification est basée sur les homologies de séquence et regroupe en familles des enzymes ayant des motifs structuraux communs et un mécanisme d'action catalytique analogue.

A. Choix des glyco-enzymes

A.1. Les glyco-enzymes impliquées dans la glucosylation du LPS de *S. flexneri* ?

Pour la synthèse chimio-enzymatique des antigènes O de *S. flexneri* 1b et 3a et en particulier la glucosylation de la *N*-acétyl-D-glucosamine et du L-rhamnose, les premières enzymes à considérer sont les enzymes impliquées dans la glucosylation du LPS de la bactérie. Un **complexe tri-enzymes** intervenant dans ce processus a été identifié. Il s'agit des enzymes GtrA, GtrB et Gtr (type) qui dépendent d'un même opéron porté par un bactériophage tempéré. La fonction précise des protéines transmembranaires correspondantes GtrA, GtrB et Gtr (type) n'est pas entièrement élucidée. Cependant, un modèle illustrant le processus de glucosylation de l'antigène O a été récemment proposé. Gtr-B permettrait de générer l'UndP-glucose (undecaprenol phosphate-glucose). Ce précurseur serait alors transloqué par la GtrA depuis le cytoplasme vers le

périplasme, où le résidu glucosyl serait transféré sur un antigène en voie de synthèse¹¹⁹⁻¹²¹ par une Gtr (type) **spécifique pour chaque sérotype**.

Pour les sérotypes 1a, 2a, 4a, 5a et X, cinq Gtr (type) ont été identifiées mais restent encore inclassées dans la base de données CAZy. **Notons qu'aucune Gtr (type) n'est connue pour les sérotypes 1b et 3a**. En général, ces Gtr (type) **ne partagent pas d'homologie significative**. Cette diversité protéique découle certainement de leur spécificité. Par contre, elles utilisent toutes le UndP-glucose comme substrat. Les connaissances sur ces Gtr (type) restent encore peu avancées¹²²⁻¹²⁴. Notamment, des modèles topologiques des Gtr (type) réalisés par des programmes informatiques de prédiction révèlent qu'elles contiennent de 9 à 11 segments transmembranaires^{122, 124}. Elles sont extrêmement complexes à isoler et à étudier et utilisent un substrat glycolipidique peu disponible. Ces propriétés sont rédhibitoires pour **leur utilisation dans une étape enzymatique aussi bien *in vivo* qu'*in vitro***.

A.2. Vers d'autres glycosyltransférases ?

Les glycosyltransférases (GTs) sont les enzymes **responsables de la synthèse de la majeure partie des disaccharides, oligosaccharides et polysaccharides** rencontrés dans la nature, la biosynthèse assurée par les transglycosidases du groupe des GHs étant très minoritaire. On estime aujourd'hui qu'environ 1 % des cadres ouverts de lecture (ORFs) de chaque génome serait dédié à la synthèse de liaisons osidiques^{125, 126}, la plupart d'entre elles étant assurée par des glycosyltransférases. Elles catalysent le transfert de résidus glucidiques issus d'un sucre activé (le plus souvent par un groupement phosphate, nucléotide-phosphate ou lipide-phosphate) sur des molécules acceptrices spécifiques qu'elles soient glucidiques ou non. Plus de 31 500 glycosyltransférases, dont bon nombre d'entre elles restent encore putatives sont aujourd'hui réparties en 89 familles dans CAZy (plus environ 500 enzymes encore non classées). Certaines seraient apparentées et la notion de clan a été proposée pour certaines familles¹²⁵, bien que la banque de données ne les mentionne pas.

Notons qu'hormis les GTs de *S. flexneri*, **aucune autre glycosyltransférase** n'a été identifiée comme catalyseur de la synthèse des motifs cibles. Pour envisager leur emploi comme outil de glucosylation, il faudrait impérativement avoir recours à l'ingénierie pour les adapter aux substrats cibles. De plus, ces enzymes, généralement associées aux membranes sont **déliçates à purifier et à caractériser**. De plus, les nucléotides-sucres restent **chers, difficilement accessibles et sont parfois instables**. Toutes ces considérations représentent autant d'obstacles au développement et à l'utilisation de ces biocatalyseurs à grande échelle.

A.3. Une préférence pour les glycoside-hydrolases ... plus particulièrement les glucane-saccharases

Les glycoside-hydrolases (GHs) sont **des enzymes très largement répandues, impliquées dans l'hydrolyse d'une liaison osidique**. Elles sont principalement actives pour la dégradation des oligosaccharides ou polysaccharides (glycosidases: E.C.3.2.1.x) mais certaines **sont aussi capables de former des liaisons osidiques (transglycosidases : E.C.2.4.x.y)**. D'autres présentent une activité lyase (EC.4.2.x.y.) ou ont même perdu leur machinerie catalytique pour évoluer vers d'autres fonctions comme transporteur d'acides aminés¹²⁷, inhibiteurs¹²⁸ ou encore agir comme la lectine¹²⁹. A ce jour (décembre 2007), on dénombre 106 familles de GHs, elles-mêmes subdivisées en sous famille. Certaines familles sont regroupées en clans (14 clans : GH-A à GH-N) lorsqu'elles renferment des enzymes apparentées d'un point de vue structural, résultant du fait que les structures des protéines sont mieux conservées que les séquences. Cinq familles ont été supprimées (GHs 21, 40, 41, 60 et 69) et leurs enzymes reclassées. Au total, près de 36 000 glycoside-hydrolases putatives ou caractérisées sont, à ce jour, classées. On compte également 317 GHs n'ayant pas encore été attribuées à l'une des familles définies.

L'utilisation des glycoside-hydrolases, comme outils de glycosylation a été **jugée plus appropriée pour satisfaire les critères de la synthèse enzymatique ciblée**. En effet, ces enzymes sont souvent bon marché et abondantes ; elles sont généralement solubles, nombre d'entre elles peuvent être produites sous forme recombinante et beaucoup sont disponibles commercialement. Elles présentent souvent des spécificités de substrat (donneurs et accepteurs) plus larges que les glycosyltransférases. Au sein des glycoside-hydrolases, les transglycosidases (EC.2.4.x.y) sont présentes dans 23 familles généralement en faible nombre. En particulier, les α -transglucosidases n'apparaissent que dans 5 familles (GHs 13, 57, 70, 77 et 94). De plus, la glycosylation de la *N*-acétyl-D-glucosamine par l'amyloamylase de *Escherichia coli* IFO 3806¹³⁰ (famille GH77) et par les maltose phosphorylases de *Nessleria perflava*¹³¹ et de *Propionibacterium freudenreichii*¹³² (famille GH65) a déjà été décrite. Dans le cas de l'amyloamylase, la glycosylation se produit par formation d'une liaison α -1,4. Par contre, aucun travail ne mentionne la glycosylation d'accepteurs protégés qui pourraient ensuite être rallongés par synthèse chimique.

Parmi les α -transglucosidases, les amylosaccharases (EC.2.4.1.4-GH13), les dextrane-saccharases (E.C.2.4.1.5-GH70) et les alternane-saccharases (E.C.2.4.1.140-GH70) sont des enzymes spécifiques du saccharose, catalysant la synthèse de polymères de glucose (glucanes) par transferts successifs d'unités α -D-glucopyranosyles. **Ces glucane-saccharases** se révèlent être particulièrement attractives pour notre étude, dans la mesure où elles sont aussi capables de

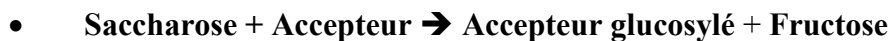
glucosyler une large gamme d'accepteurs saccharidiques, voire des molécules hydroxylées. Chacune d'entre elles possède **une ou plusieurs spécificités de liaison**, permettant ainsi de disposer d'une gamme d'enzymes de spécificité différente : $\alpha(1-2)$ / $\alpha(1-3)$ / $\alpha(1-4)$ et $\alpha(1-6)$. Enfin, le saccharose qu'elles utilisent naturellement comme donneur est **un substrat abondant et peu coûteux**. Ce sont donc ces enzymes qui ont été proposées comme candidates à l'étape enzymatique.

B. Les outils enzymatiques : les glucane-saccharases

Les glucane-saccharases (GS) sont des enzymes d'origine bactérienne capables de synthétiser des glucanes de masse moléculaire élevée par transferts successifs d'unités α -D-glucopyranosyles à partir de saccharose seul, selon la réaction :



En présence de molécules acceptrices exogènes, principalement des monosaccharides, elles peuvent également catalyser leur glucosylation, formant alors de nouveaux oligosaccharides de faible DP (Degré de Polymérisation) :



L'ensemble des GS est regroupé dans les familles 13 et 70 des GH. Alors que la famille GH70 est composée uniquement de GS (73 recensées, 37 identifiées), la famille GH13 (ou famille des α -amylases) contient des enzymes amylolytiques (plus de 4000 recensées) aux activités (21) et spécificités remarquablement variées. En outre, l'amylosaccharase est la seule enzyme de cette famille dotée d'une activité polymérase. Cette capacité à former un polymère à partir du saccharose la rapproche des GS de la famille GH70.

Enfin, notons que ces 2 familles d'enzymes (GH13 et GH70) constituent au même titre que la GH77 (famille de l'amyloamylase ou 4- α -glucanotransférase) le clan GH-H. L'appartenance à ce même clan reflète les caractères communs suivants :

- un même mécanisme de rétention de configuration anomérique de la liaison osidique
- une même machinerie catalytique constituée d'une triade catalytique composée de deux acides aspartiques et d'un acide glutamique
- une même structure du domaine catalytique $(\beta/\alpha)_8$

Aussi, au vu des similarités existant entre les glucane-saccharases des familles GH13 et GH70, nous présenterons tout d'abord ces enzymes sur un plan général en insistant davantage sur les

données connues relatives à la glycosylation d'accepteurs non naturels. Puis nous recentrerons l'exposé sur l'amylosaccharase de *N. polysaccharea*, enzyme la plus utilisée dans l'étude et qui a fait l'objet de la stratégie de remodelage.

B.1. Présentation des glucane-saccharases

B.1.a. Biodiversité des glucane-saccharases

La grande variété de spécificités des glucane-saccharases a amené à classer ces enzymes en fonction de la nature des glucanes qu'elles synthétisent. Les structures de ces polymères sont déduites à l'aide de diverses techniques comme l'oxydation au périodate, la méthylation du polymère, l'analyse des fragments de méthylation, l'acétylyse et la spectroscopie R.M.N. du carbone 13 ¹³³.

Ainsi, les glucane-saccharases peuvent être classées en cinq groupes :

- **Les amylosaccharases (AS) (GH13, E.C. 2.4.1.4).** Elles synthétisent un polymère de type amylose, composé essentiellement de liaisons α -1,4. Les micro-organismes producteurs d'amylosaccharases recensés dans la littérature appartiennent au genre *Neisseria*, mais plus récemment le séquençage de nombreux génomes bactériens a révélé la présence de gènes codant pour des amylosaccharases putatives.
- **Les dextrane-saccharases (DSR) (GH70, E.C. 2.4.1.5).** Elles synthétisent un polymère présentant un minimum de 50% de liaisons osidiques α -1,6 successives avec des ramifications en α -1,2, α -1,3 ou α -1,4. Ce type de polymère, nommé dextrane, est soluble dans l'eau. Les dextrans les plus représentatifs sont ceux synthétisés par *L. mesenteroides* NRRL B-512F, NRRL B-742F et NRRL B-1299
- **Les alternane-saccharases (ASR) (GH70, E.C. 2.4.1.140).** Elles catalysent la synthèse d'un glucane dénommé " alternane ". Il est formé d'un enchaînement de liaisons α -1,6 et α -1,3 alternées, et présente quelques ramifications liées en α -1,6 ou α -1,3. Trois souches productrices d'alternane ont été mises en évidence : *L. mesenteroides* NRRL B-1355, NRRL B-1501 et NRRL B-1498^{133, 134,135}.
- **Les mutane-saccharases (GH70, E.C. 2.4.1.5).** Elles produisent un polymère constitué majoritairement de liaisons α -1,3 dans la chaîne principale, appelé mutane et

généralement insoluble dans l'eau. Elles sont produites par *L. mesenteroides* NRRL B-523, NRRL B-1149 et par différentes souches de *Streptococcus* sp.¹³⁶.

- **Les reuterane-saccharases (GH70, E.C. 2.4.1.5).** Elles catalysent la synthèse d'un glucane dit " reuterane ", composé à plus de 50% de liaisons α -1,4 dans la chaîne principale, le reste étant lié en α -1,6 (chaîne principale et ramification). La découverte de ces polymères est très récente, et seules deux souches productrices de *Lactobacillus reuteri* sont à ce jour recensées : *Lb. Reuteri* 121 et *Lb. Reuteri* " BioGaia " ^{137, 138}.

Les GS de la famille GH70 sont produites par des bactéries lactiques Gram⁺ appartenant aux genres *Leuconostoc*, *Streptococcus*, et *Lactobacillus*¹³⁹. Seuls, les Streptocoques ne sont pas reconnus comme **G.R.A.S.** (Generally Recognized As Safe) par la **F.D.A.** (Food and Drug Administration). Il n'est pas rare de rencontrer des souches produisant plusieurs enzymes, comme, par exemple, dans celles de *S. mutans*¹⁴⁰, *S. sobrinus*, *L. citreum* NRRL B-742¹⁴¹, *L. mesenteroides* NRRL B-1355¹⁴², *L. mesenteroides* NRRL B-1299^{143, 144} ou encore *Lactobacillus reuteri* 121¹⁴⁵.

Les glucane-saccharases dont le gène a été cloné sont recensées dans les **TABLEAUX 4.1** et **4.2**.

TABLEAU 4.1 | Glucane-saccharases de la famille GH13 clonées et caractérisées.

| Souche | Enzyme | Taille (aa) | α -(1,6) | α -(1,4) | α -(1,3) | α -(1,2) | Référence |
|---------------------------------|--------|-------------|-----------------|-----------------|-----------------|-----------------|-----------|
| <i>N. polysaccharea</i> | ASNP | 637 | | 100 | | | 146 |
| <i>Deinococcus radiodurans</i> | ASDR | 644 | | 100 | | | 147 |
| <i>Deinococcus geothermalis</i> | ASDG | 650 | | 100 | | | 148 |

TABLEAU 4.2| Principales glucane-saccharases de la famille GH70 clonées et caractérisées.

| Souche | Enzyme | Taille (aa) | α -(1,6) | α -(1,4) | α -(1,3) | α -(1,2) | Référence |
|---------------------------------|----------|-------------|-----------------|-----------------|-----------------|-----------------|-----------|
| <i>Ln. mesenteroides</i> B-512F | DSR-S | 1527 | 95 | | 5 | | 149 |
| <i>Ln. mesenteroides</i> B-512F | DSR-T5* | 1499 | 50 | | 40 | | 150 |
| <i>Ln. mesenteroides</i> B-1299 | DSR-E | 2835 | 68 | | 3 | 29 | 151 |
| <i>Ln. mesenteroides</i> B-1299 | DSR-A | 1290 | 85 | | 15 | | 152 |
| <i>Ln. mesenteroides</i> B-1299 | DSR-B | 1508 | 95 | | | | 153 |
| <i>Ln. mesenteroides</i> B-1355 | DSR-C | 1477 | 96 | | | | 154 |
| <i>Ln. mesenteroides</i> Lcc4 | DSR-D | 1527 | majoritairement | | | | 155 |
| <i>S. mutans</i> GS 5 | GFT-D | 1430 | 70 | | 30 | | 156 |
| <i>S. oralis</i> | GTF-R | 1575 | 86 | | | | 157, 158 |
| <i>S. downei</i> Mfe 28 | GTF-S | 1365 | 90 | | | | 159 |
| <i>S. sobrinus</i> OMZ176 | GTF-T | 1468 | 79 | | 21 | | 160 |
| <i>S. sobrinus</i> B13N | GTF-U | 1554 | 69 | | 31 | | 161 |
| <i>S. salivarius</i> ATCC 25975 | GTF-K | 1599 | 100 | | | | 162 |
| <i>S. salivarius</i> ATCC 25975 | GTF-M | 1577 | 95 | | | | 162 |
| <i>S. gordonii</i> CH1 | GTF-G | 1577 | 75 | | 23 | | 163 |
| <i>Lb. reuteri</i> 180 | GTF-180 | 1772 | 61 | | 39 | | 145 |
| <i>Lb. sakei</i> Kg15 | GTF-Kg15 | 1561 | 90 | | 10 | | 145 |
| <i>Lb. fermentum</i> Kg3 | GTF-Kg3 | 1595 | 92 | | 8 | | 145 |
| <i>Lb. parabuchneri</i> 33 | GTF-33 | 1463 | 81 | | 19 | | 145 |
| <i>S. salivarius</i> ATCC 5975 | GFT-L | 1449 | 50 | | 50 | | 162 |
| <i>S. salivarius</i> ATCC 25975 | GTF-J | 1517 | | | 90 | | 162 |
| <i>S. mutans</i> GS 5 | GTF-B | 1475 | 12 | | 88 | | 164 |
| <i>S. mutans</i> GS 5 | GTF-C | 1375 | 15 | | 86 | | 165 |
| <i>S. downei</i> Mfe 28 | GTF-I | 1597 | 12 | | 88 | | 166 |
| <i>S. sobrinus</i> 6715 | GTF-Ia | 1592 | | | majoritairement | | 167 |
| <i>Lb. reuteri</i> ML1 | GTF-MLI | 1772 | 35 | | 65 | | 145 |
| <i>Lb. reuteri</i> 121 | GTF-A | 1781 | 43 | 57 | | | 137 |
| <i>Lb. reuteri</i> BioGaia | GTF-Bio | 1781 | 22 | 78 | | | 138 |
| <i>Ln. mesenteroides</i> B-1355 | ASR | 2057 | 45 | | 55 | | 154 |
| <i>Ln. mesenteroides</i> B-1299 | GBD-CD2* | 1745 | | | | 100 | 168 |

B.1.b Données structurales

Avec une masse molaire moyenne de 160 à 180 kDa¹⁶⁹, les GS de la famille GH70 sont de taille beaucoup plus importante que les amylosaccharases de la famille GH13 de masse molaire voisine de 70 kDa. Cette taille assez inhabituelle les rend particulièrement délicates à exprimer sans formes dégradées, à purifier mais aussi à cristalliser. Aussi, jusqu'en 2007, aucune structure tridimensionnelle n'avait été obtenue. La **résolution récente des structures des GTF-A et GTF-180 de *Lactobacillus reuteri*** (données non publiées¹⁷⁰) **constitue une avancée remarquable**, qui permettra à l'avenir, de mieux comprendre les facteurs structuraux impliqués dans la spécificité des glucane-saccharases.

En l'absence de structure 3D, les analyses de structure primaire et les prédictions de structure secondaire ont cependant permis de révéler la proximité structurale et mécanistique des enzymes de la famille GH70 avec celles de la famille GH13, famille pour laquelle de nombreuses structures ont été résolues (50) dont **celle de l'amylosaccharase de *N. polysaccharea*** (ASNP).

B.1.b.1. Le domaine catalytique : une structure prédite en tonneau (β/α)₈

L'observation d'homologies de séquences dans le domaine catalytique des glucane-saccharases et des enzymes de la famille des α -amylases suggère que ces deux familles d'enzymes déploient un mécanisme apparenté¹⁷¹⁻¹⁷³. Des prédictions de structures secondaires par HCA ("Hydrophobic Cluster Analysis"), associées aux comparaisons de séquences^{172,173}, ont permis de conclure que le domaine catalytique des glucane-saccharases adopte une structure en tonneau (β/α)₈ analogue à celle décrite pour les enzymes de la famille GH13.

Cependant, les prédictions de structure secondaire proposées par MacGregor et coll. (1996)¹⁷³ et Devulapalle et coll. (1997)¹⁷² divergent sur l'organisation du tonneau (β/α)₈. Les premiers proposent **une permutation circulaire** des éléments E1-H1-E2-H2-E3 (avec H les hélices α et E les brins β), qui seraient situés après les éléments H3-E4-H5-E5-H6-E6-H7-E7-H8-E8. Devulapalle et coll. suggèrent au contraire un enchaînement des brins et des hélices analogue à celui rencontré chez les α -amylases (E1 à H8). Des travaux plus récents de modélisation du domaine catalytique de GTF-B de *S. mutans*¹⁷⁴ et de dichroïsme circulaire effectués sur la GTF-I de *S. downei*¹⁷⁵ soutiennent l'hypothèse d'une permutation circulaire. Cette dernière semble la plus vraisemblable et justifie l'appartenance de ces glucane-saccharases à une famille distincte de celle des α -amylases (**FIGURE 15**).

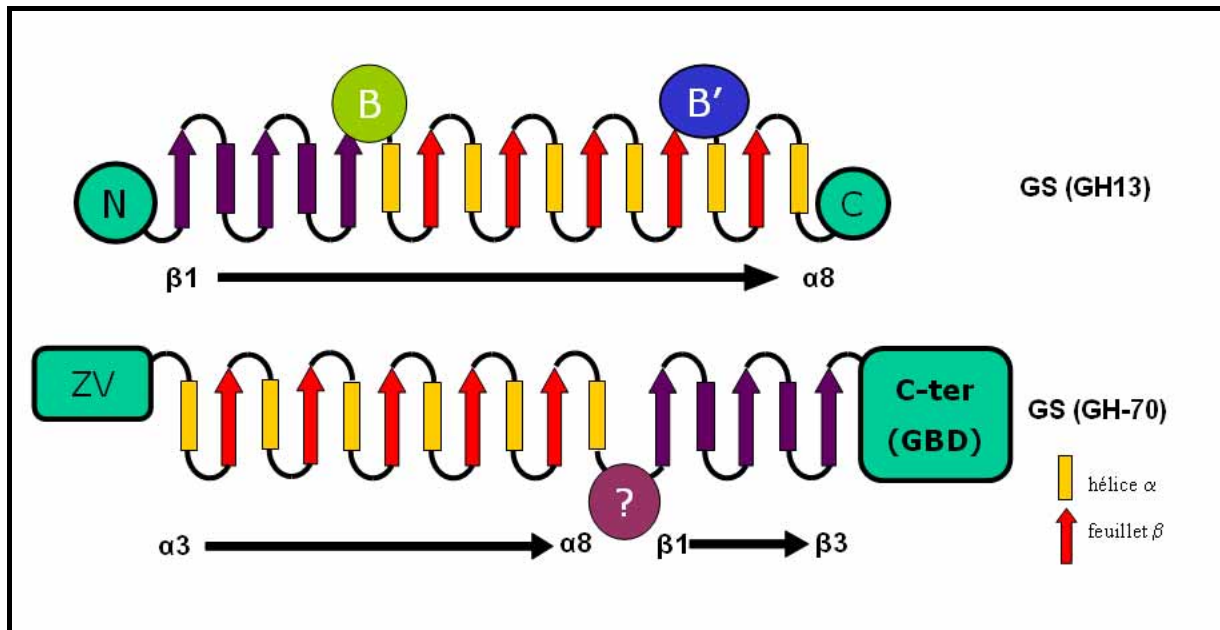


FIGURE 15| Représentation schématique de l'organisation du tonneau (β/α)₈ des GS des familles GH13 et GH70 et des domaines N, B, B' et C avec l'exemple de l'AS de *N. polysacchara* (GH13). (ZV: zone variable, ? : zone située entre la permutation circulaire et C-term : domaine C-terminal).

B.1.b.2. Les acides aminés conservés du site actif

Les alignements de séquences des domaines catalytiques des enzymes des familles GH13 et GH70 ont permis de retrouver **6 des 7 résidus les mieux conservés** dans la famille GH13^{172, 173}. On observe la présence des 3 acides carboxyliques de la triade catalytique (Asp286, Glu328 et Asp393 de l'ASNP), de l'arginine et de l'acide aspartique correspondant, respectivement, à Asp182 et Arg284 de l'ASNP et d'une des deux histidines de stabilisation de l'intermédiaire covalent (His392 de l'ASNP). La seconde serait remplacée, chez les enzymes de la famille GH70, par un résidu glutamine (Gln). L'implication de ces résidus a été confirmée par mutagenèse dirigée chez diverses enzymes^{149, 172, 175-177}, suggérant qu'ils occupent les mêmes fonctions que pour les α -amylases¹⁷³. L'observation de ces mêmes résidus "signature", chez les membres des familles 13 et 70 des glycoside-hydrolases, suggère un mécanisme commun pour ces deux familles d'enzymes, tout du moins au niveau de la formation du complexe glucosyl-enzyme. Ce complexe covalent glucosyl-enzyme a d'ailleurs été mis en évidence chez les glucane-saccharases de *S. sobrinus* dès 1989 par Mooser et coll.^{171, 178}.

B.1.c Mécanisme catalytique

Les GS partagent toutes un même mécanisme dit "de rétention de la configuration anomérique", sous forme α générant ainsi un produit dont la stéréochimie du carbone anomérique est identique à celle du substrat (suite à deux inversions successives). Ce mécanisme a particulièrement bien été étudié chez les enzymes de la famille GH13, notamment chez l'ASNP.

L'hydrolyse de la liaison osidique du saccharose met ainsi en jeu deux acides carboxyliques critiques : un résidu acide/base (Glu) et un nucléophile (Asp)¹¹⁴ (FIGURE 16).

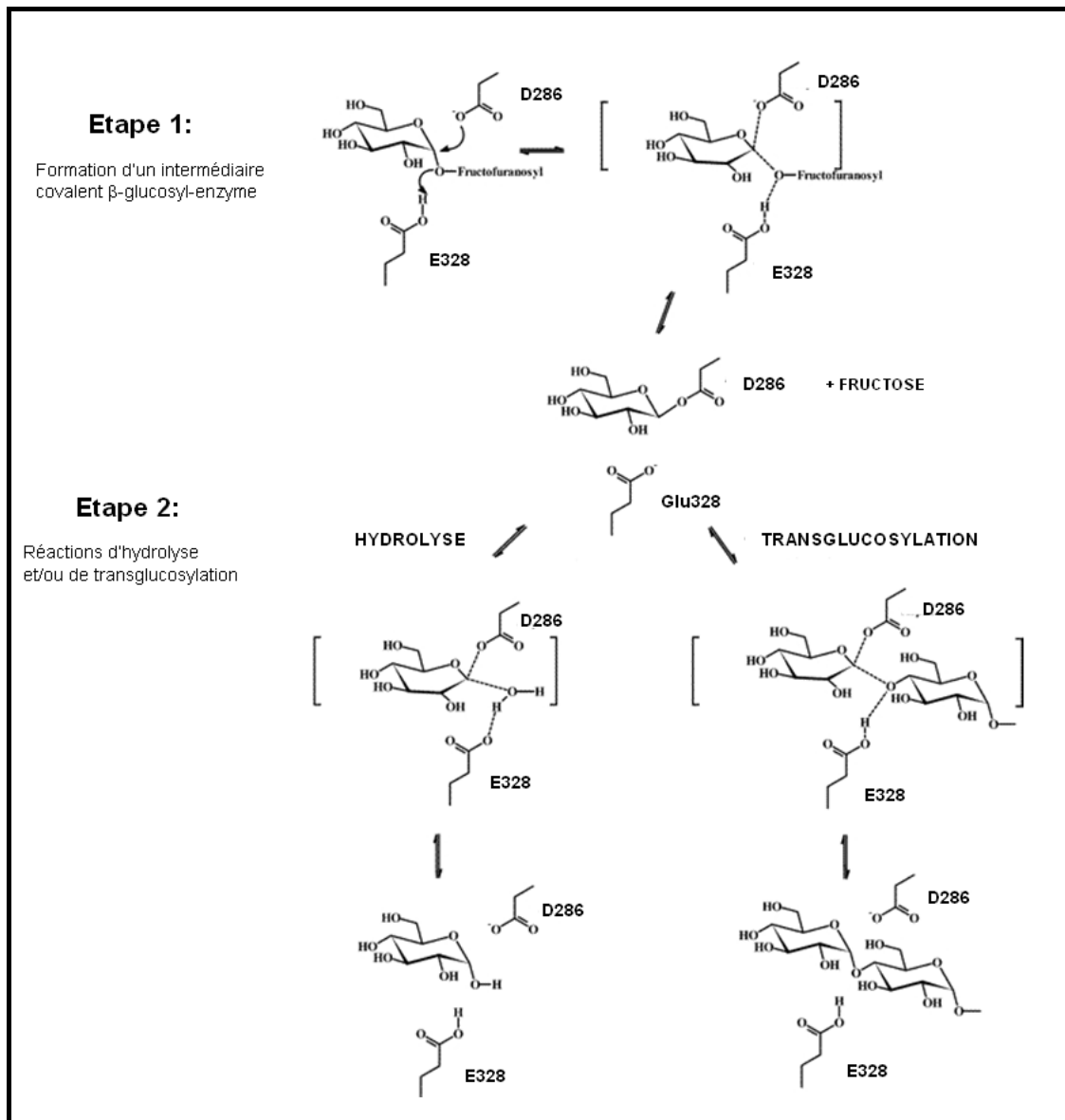


FIGURE 16 | Schéma du mécanisme réactionnel suivi par l'ASNPN.

(D'après Skov et coll.¹⁷⁹)

Dans ce mécanisme, la première étape consiste en l'attaque du résidu nucléophile/base sur le carbone C1 de la partie « glycone » du substrat, générant la formation d'un intermédiaire covalent β -glucosyl-enzyme. Simultanément, le catalyseur acide/base cède un proton à l'oxygène de la liaison osidique, conduisant à la libération de fructose. Au cours de la seconde étape, le résidu acide/base déprotoné agit en tant que base pour activer une molécule d'eau venant attaquer le β -glucosyl-enzyme. Si un autre sucre ou une molécule hydroxylée est présent dans le milieu réactionnel, une réaction de transglucosylation entre en compétition avec la réaction d'hydrolyse, aboutissant à la formation d'un nouvel oligosaccharide. Dans un tel mécanisme, la distance entre

les deux acides carboxyliques est d'environ 5.5 \AA ^{114, 180, 181}. Le site actif est composé d'un troisième acide carboxylique (Asp), formant ainsi une triade catalytique Asp-Glu-Asp. Cet acide aspartique interviendrait dans la stabilisation de l'intermédiaire covalent¹⁸²⁻¹⁸⁴.

Au cours du mécanisme, des états de transition de type ion oxocarbone apparaissent¹⁸⁰ (**FIGURE 16**). L'ion oxocarbone comporte une double liaison partielle entre les atomes O5 et C1, sur lesquels une charge positive est délocalisée, et est caractérisée par une configuration plane des atomes C2, C1, O5 et C5. Cet état est défini comme la structure de plus haute énergie au cours de la réaction et est donc hautement instable¹⁸⁵.

Des expériences de diffraction des rayons X sur des cristaux de complexes enzyme : inhibiteurs fluorés ont permis de piéger l'intermédiaire covalent β -glycosyl-enzyme formé par l'ASNP, procurant une description fine de la catalyse au niveau atomique¹⁸⁶. Ces résultats structuraux entérinent l'existence de cet intermédiaire covalent, initialement proposé par Koshland en 1953 et qui, après avoir longtemps donné lieu à polémique, est aujourd'hui largement admis¹⁸⁷.

B.1.d. Réactions catalysées par les GS

La première réaction catalysée par les glucane-saccharases est la formation d'un intermédiaire covalent glucosyl-enzyme. A partir de cet élément central entrent **en compétition** plusieurs réactions¹⁸⁸. La nature des produits de synthèse dépend alors de la destination du résidu glucosyle (**FIGURE 17**) qui **peut être transféré sur différentes molécules acceptrices**, soit:

- la chaîne de glucane en cours d'élongation, conduisant à la formation des polymères de haute masse molaire,
- une molécule d'eau, conduisant à une réaction d'hydrolyse¹⁸⁹,
- le fructose libéré, conduisant soit à la néo-synthèse de saccharose (échange isotopique), soit à la formation d'isomères du saccharose,
- une molécule acceptrice exogène, le plus souvent de type osidique. On parle alors de **réaction d'accepteur**¹⁹⁰. Les produits issus de cette réaction peuvent par la suite jouer le rôle d'accepteur, aboutissant à la synthèse d'oligosaccharides.

La synthèse de ces produits ne nécessite aucun apport énergétique autre que celui libéré par la rupture de la liaison osidique du saccharose (6.6 kcal/mol), et l'équilibre est très largement déplacé dans le sens de la synthèse (réactions 2, 3, 4 et 5).

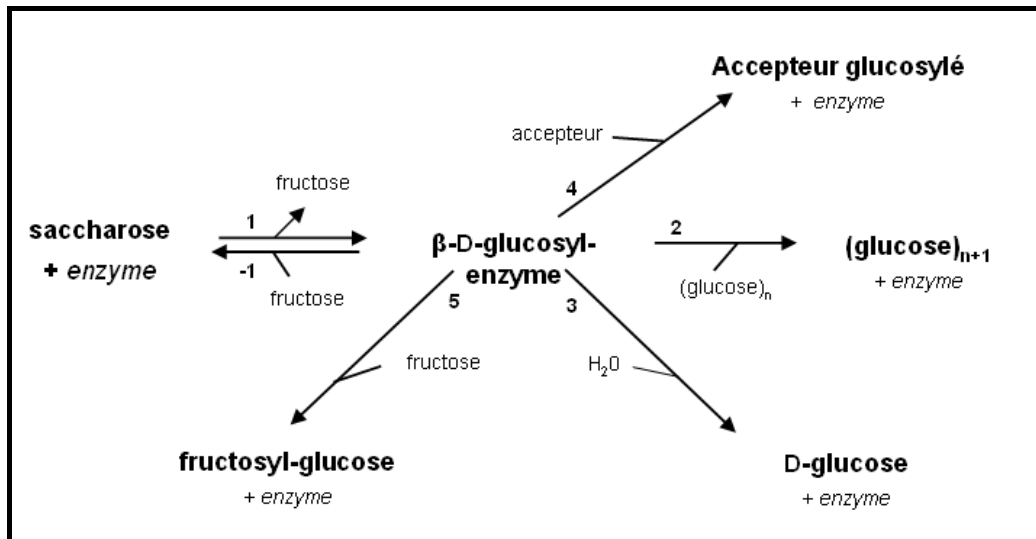


FIGURE 17| Réactions catalysées par les glucane-saccharases.

+1: formation du glucosyl-enzyme; -1: échange isotopique; 2: réaction de polymérisation; 3: hydrolyse; 4: réaction d'accepteur; 5: formation d'isomères du saccharose.

Qu'il s'agisse de réactions de polymérisation ou d'accepteur exogène, le transfert du résidu glucosyle a lieu du côté de l'extrémité non-réductrice des molécules acceptrices. Notamment, dans le cas de l'ASNP, le mode de polymérisation correspond à une **élongation non-processive**, au cours de laquelle les chaînes en cours d'élongation sont libérées dans le milieu après chaque transfert d'unités glucosyles¹⁹¹. Chez les GS de la GH70, ces chaînes fixées à des zones d'ancrage de la protéine ne seraient que partiellement libérées, favorisant leur disponibilité pour de nouvelles réactions de glucosylation. Un tel mécanisme, dit **semi-processif**, a été proposé, justifiant ainsi la formation rapide de chaînes de glucane de très haute masse molaire¹⁹².

B.1.d.1. Les GS, catalyseurs de réactions de glucosylation d'accepteurs exogènes

Alors qu'il n'existe qu'un nombre très limité de donneurs de glucosyle référencés dans la littérature (principalement le saccharose mais aussi quelques dérivés tels que le fluorure α -D-glucopyranoside ou le p-nitrophenyl- α -D-glucopyranoside^{193, 194}), une large gamme de molécules peut jouer le rôle d'accepteur. Ces accepteurs sont classiquement introduits dans le milieu réactionnel en conditions équimolaires par rapport au donneur, le saccharose. Deux types d'accepteurs émergent de ces travaux :

- **les accepteurs forts** : ce sont des accepteurs facilement glucosylés par les GS. Les produits de glucosylation générés deviennent, à leur tour, des accepteurs potentiels, conduisant dans certains cas à la formation de familles d'oligosaccharides¹⁹⁵. De tels accepteurs se **caractérisent par leur fort pouvoir activateur**, accélérant de par leur présence, la vitesse de consommation de saccharose. Ainsi, en présence de D-glucose et

de maltose, l'activité initiale de l'ASN est augmentée respectivement d'un facteur 7,5 et 6,4. Elle l'est d'un facteur 100 en présence de glycogène utilisé en conditions saturantes (30g/L), le meilleur accepteur de l'ASN connu à ce jour¹⁹⁶. La glycosylation de ces accepteurs est alors **largement privilégiée, au détriment de la polymérisation**¹⁹⁷. Cette propriété à diminuer le rendement en polymère est notamment exploitée pour évaluer le pouvoir accepteur de ces molécules exogènes^{198, 199}. Il est par ailleurs possible d'influencer la distribution de taille et le rendement en oligosaccharides en jouant sur le ratio des concentrations en saccharose/ accepteur^{200,201, 202}.

- **les accepteurs faibles** : ce sont des accepteurs difficilement reconnus par les GS, formant principalement un produit mono- voire diglucosylé. Le rendement en nouveaux oligosaccharides est faible. Ces accepteurs n'auront pas d'effet activateur, la réaction de polymérisation restant prépondérante¹⁹⁰. Introduits en large excès, ils agiront même en tant qu'inhibiteurs modérés²⁰³. Classé dans cette catégorie d'accepteurs, le fructose est converti en isomères de saccharose, généralement le leucrose²⁰³.

Par ailleurs, l'accepteur exogène différant de l'accepteur naturel (D-Glucose), la liaison osidique créée entre le premier résidu glucosyle et l'accepteur **ne correspond pas forcément à la spécificité de liaison de l'enzyme envers son substrat naturel**. Elle dépendra du positionnement de cet accepteur dans le site actif de l'enzyme. A titre d'exemple, l'amylosaccharase de *N. polysacchara* (ASN), de spécificité α -1,4 forme majoritairement avec le L-arabinopyranose un produit monoglucosylé en α -1,3²⁰⁴. Cette liaison osidique pourra éventuellement être orientée par des accepteurs protégés (par un groupement tosyl) bloquant le site préférentiel de glycosylation²⁰⁵. Cependant, les glucane-saccharases conservent généralement leur spécificité de synthèse lors des transferts de résidus glucosyles qui suivront le premier, puisque cette fois-ci, la réaction de glycosylation aura lieu sur l'unité glucosyle précédemment greffée (partie réductrice des produits générés). Ainsi, en présence de maltose, l'ASN synthétise une série de malto-oligosaccharides (liaisons α -1,4 uniquement)¹⁴⁶. La dextrane-saccharase DSR-S de *L. mesenteroides* NRRL B-512F, elle, synthétise une série d'oligosaccharides présentant une molécule de maltose côté réducteur, et environ 8 à 12 unités glucosyle liées uniquement en α -1,6 côté opposé¹⁹⁷. L'alternane-saccharase pourra en plus synthétiser des liaisons α -1,3, la reuterane-saccharase des liaisons α -1,4 et la dextrane-saccharase de *L. mesenteroides* NRRL B-1299 des liaisons α -1,2.

Une centaine de molécules hydroxylées de nature variée ont été identifiées comme accepteur d'une GS, témoignant ainsi de la plasticité de ces enzymes au niveau de leur site de reconnaissance de l'accepteur :

- les **mono- et disaccharides** classiques (D-Glucose¹⁹⁹, maltose¹⁹⁹, isomaltose¹⁹⁹ ...) et moins conventionnels (L-glucose²⁰⁶, D-Tagatose²⁰⁶, cellobiose²⁰⁷, gentiobiose¹⁹⁹)
- les **oligosaccharides** : panose¹⁹⁹, acarbose²⁰⁸ ...
- les **polysaccharides** : glycogène²⁰⁹, amylopectine²⁰⁹ ...
- les **sucres fonctionnalisés** : l'anhydrofructose²⁰², le fructose dianhydride²¹⁰, les alkylglucosides^{199, 211}, les methyl-hexopyranosides¹⁹⁸, l'acide arabinonique²¹⁰, les aldoses²¹⁰, les thioglucosides²⁰⁵ ...
- les **alditols** : D-glucitol²¹⁰, D-mannitol²¹⁰, maltitol^{199, 210}
- les **alcools** : méthanol, éthanol, 2-chloroéthanol, propanol et isopropanol¹⁵⁸
- les **composés aromatiques** : catéchines^{212,213}, mono- ou polyphénols (salicine²¹⁴, flavonoides²¹⁵)
- les **dérivés d'acides aminés**¹⁵⁸

A titre d'exemple, la **FIGURE 18** illustre, avec 46 accepteurs glucidiques²¹⁶ testés, le potentiel de l'AS de *Neisseria polysaccharea* à reconnaître des composés autres que son accepteur naturel.

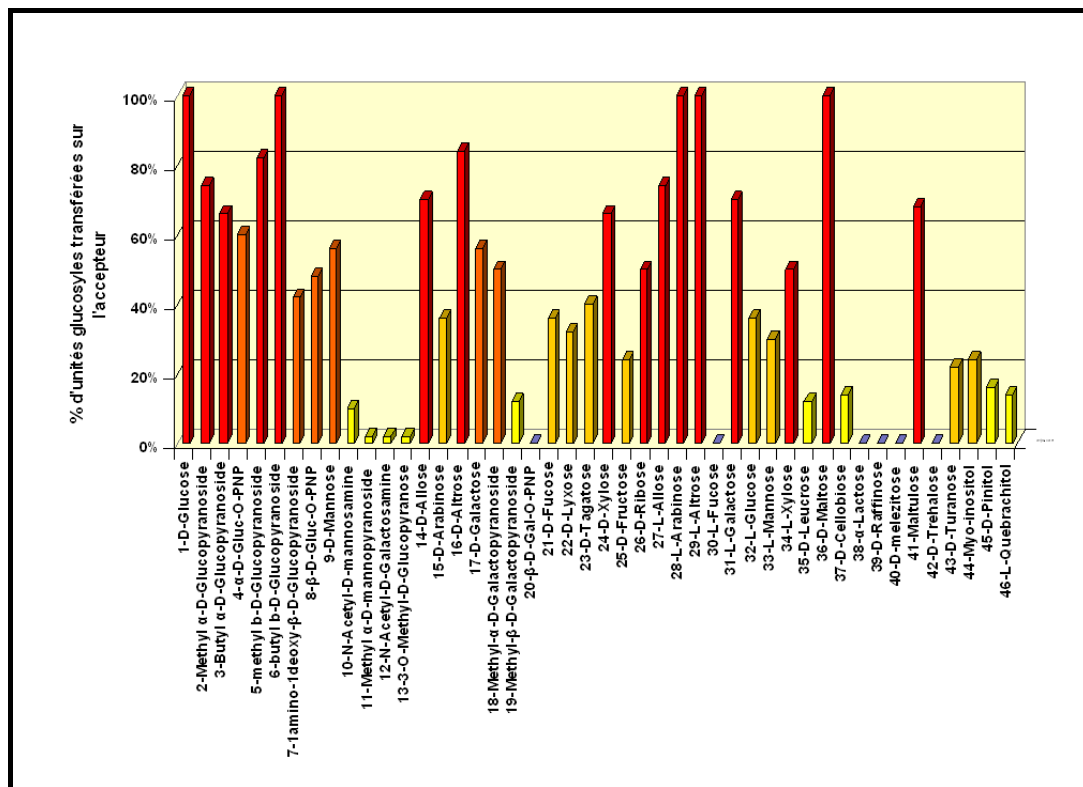


FIGURE 18| Glucosylation d'accepteurs glucidiques catalysée par l'ASNP.
(Conditions de réaction: saccharose=146mM/ accepteur=292mM)

Outre l'accès à de nouveaux composés et à des données fonctionnelles supplémentaires sur le potentiel de ces GS, la glucosylation de telles molécules suscite un intérêt applicatif particulier dans :

- la synthèse de **composés prébiotiques**. La dextrane-saccharase de *L. mesenteroides* NRRL B-1299 est ainsi utilisée à l'échelle industrielle pour la synthèse des α -1,2 GOS (glucooligosaccharides) aux applications prébiotiques. Dans un autre registre, l'ASNIP glucosyle des glucanes contenant des liaisons α -1,4 et α -1,6 (glycogène, amylopectine, limite-dextrines), en rallongeant leurs ramifications²⁰⁹. La teneur en amidon résistant des polysaccharides accepteurs est ainsi augmentée, et peut être modulée en fonction de la structure de l'accepteur et du rapport donneur/accepteur utilisé pour la glucosylation. Cet amidon arrive intact au niveau du gros intestin sans avoir été hydrolysé auparavant par les enzymes de l'estomac et de l'intestin grêle. L'amidon résistant joue donc le rôle de fibre alimentaire. Il est qualifié, tout comme les α -1,2 GOS de prébiotique puisqu'il stimule dans le colon la croissance de bactéries bénéfiques à l'organisme. Chez l'homme, les effets physiologiques potentiels sont la prévention du diabète de type II, du cancer du colon, des maladies cardio-vasculaires et de l'hypercholestérolémie, ainsi que le traitement de diarrhées infectieuses²¹⁷.
- la synthèse de **fragments antigéniques** à visée thérapeutique. L'alternane-saccharase de *L. mesenteroides* NRRL 1355 en présence d' α -D-méthyl galactoside permet l'obtention de α -D-glucofuranosyl-(1 \rightarrow 4)- α -D-méthyl galactoside¹⁹⁸, de structure proche du segment répété du polysaccharide antigénique O-Ag de *Salmonella* sp²¹⁸.
- **l'amélioration de certaines propriétés fonctionnelles** de molécules à visées biomédicales et/ou agro-alimentaires. En effet, la fixation de dérivés glucidiques sur certaines macromolécules permet généralement d'améliorer efficacement leur solubilité et leurs propriétés émulsifiantes ou moussantes. Plusieurs exemples illustrent cette application: Richard et coll. sont parvenus à synthétiser des α -1,6 et α -1,2 anhydrofructo-D-glucooligosaccharides de DP3 à partir d'anhydrofructose et en présence de préparations enzymatiques issues de *L. mesenteroides* NRRL B-512F, et B-1299 respectivement. Ces composés sont des anti-oxydants potentiels qui pourraient être utilisés en agro-alimentaire^{219,202}. Citons également la glucosylation de catéchines- anti-oxydant médical et alimentaire- par l'activité glucane-saccharase de *S. sobrinus*^{212,213} et celle de polyphénols par les GTF-D de *S. mutans* GS-5²²⁰, DSR-S de *L. mesenteroides* NRRL B-512F et ASR de *L. mesenteroides* NRRL B-23192²¹⁵. Enfin, mentionnons la glucosylation d'alkylpolyglucosides, employés en tant que détergents et agents moussants biodégradables par différentes glucane-saccharases de *L. mesenteroides*²¹¹.

Les glucane-saccharases constituent donc une alternative d'intérêt majeur à la synthèse par voie chimique d'oligosaccharides et glucoconjugués originaux. De plus, alors que la synthèse chimique

emploi des solvants nocifs à l'environnement, la synthèse enzymatique s'effectue principalement en milieu aqueux. On peut donc parler de « **chimie verte** ». De même, lorsque la glucosylation doit porter sur un composé non soluble dans l'eau, les glucane-saccharases démontrent également leur flexibilité d'utilisation, puisque certaines d'entre elles présentent une activité résiduelle non négligeable dans certains solvants miscibles dans l'eau^{215, 220}.

B.1.d.2 Les GS, catalyseurs de réactions des accepteurs cibles

Une seule étude, ayant donné lieu à un brevet²²¹ destiné à la synthèse d'oligosaccharides prébiotiques fait référence à la glucosylation du L-rhamnose et de la *N*-Acétyl-D-glucosamine. Ces travaux concluent que ces 2 accepteurs, n'affectant que très faiblement la production de polymère, sont peu reconnus par la GS testée, l'ASR de *L. mesenteroides* NRRL B-23297. De plus, les produits de glucosylation n'ont ni été isolés, ni caractérisés¹⁹⁹.

B.2. Les amylosaccharases

Au vu de la remarquable diversité des enzymes constituant la famille 13 des GHs, une division de cette famille en 40 sous familles a récemment été proposée²²². Basée sur la similarité de séquences de ces enzymes, cette classification apparaît comme un premier outil fiable dans la prédiction de fonction des GH13 puisqu'elle s'accompagne d'une forte corrélation entre les activités (et parfois les spécificités) et les sous-familles définies. A ce titre, l'amylosaccharase constitue avec la saccharose-hydrolase **la sous famille GH13.4**.

B.2.a. Organismes producteurs

Jusqu'à la fin des années 1990, les amylosaccharases recensées dans la littérature étaient exclusivement produites par des bactéries non-pathogènes du genre *Neisseria* (genre connu particulièrement pour ses deux espèces pathogènes *N. meningitidis* et *N. gonorrhoeae*). Hehre et Hamilton ont été les premiers, en 1946, à découvrir une amylosaccharase intracellulaire chez *Neisseria perflava*. Plus tard, la présence d'une amylosaccharase intracellulaire a été révélée chez six autres espèces de *Neisseriae* (*N. canis*, *N. cinerea*, *N. denitrificans*, *N. sicca* et *N. subflava*)²²³ appartenant à la flore de la cavité buccale et potentiellement impliquées dans la formation de la plaque dentaire²²⁴. Enfin, une amylosaccharase supposée extracellulaire a été identifiée chez l'espèce *N. polysaccharea* qui provient du rhino-pharynx d'enfants sains d'Europe et d'Afrique²²⁵. Cette enzyme, dont le gène a été cloné puis séquencé en 1999, est la première amylosaccharase recombinante caractérisée¹⁴⁶. Sa structure tridimensionnelle a été décrite en 2002¹⁷⁹.

Le séquençage récent de nombreux génomes bactériens et leur analyse par le logiciel BLAST a révélé **la présence de 34 gènes d'amylosaccharases putatives**, dont la séquence protéique

présente entre 33 et 97 % d'identité avec celle de l'AS de *N. polysaccharea* et le domaine catalytique varie entre 37 et 98% (FIGURE 19). Toutes possèdent les résidus catalytiques identifiés chez cette dernière par mutagenèse dirigée (cf partie III.B.2.d.2., page 69).

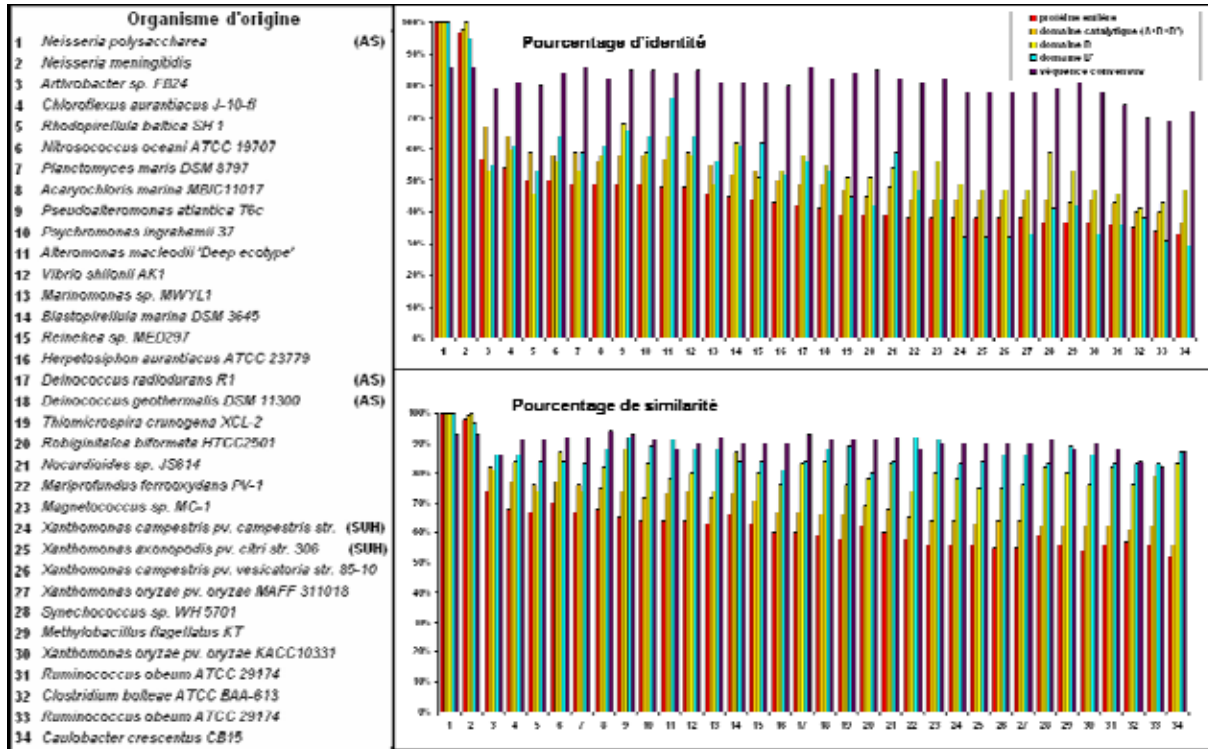


FIGURE 19| Amylosaccharases hypothétiques identifiées dans les génomes séquencés via le logiciel BLAST (analyse effectuée en novembre 2007).

Le pourcentage d'identité et de similarité (global, au niveau du domaine catalytique, du domaine B et du domaine B') est calculé en prenant l'amylosaccharase de *Neisseria polysaccharea* comme référence. Celui relatif à la séquence consensus est déterminé à partir de la séquence consensus globale des GH13.4.

Ceci a conduit à la **caractérisation de 2 nouvelles amylosaccharases recombinantes** issues de *D. radiodurans*¹⁴⁷ (2005) et de *D. geothermalis*¹⁴⁸ (2007). En revanche, la caractérisation des enzymes identifiées chez les souches du genre *Xanthomonas* a montré qu'il s'agissait de **saccharose-hydrolases (SUHs)**^{226, 227}. Selon Pizzut-Serin et coll. (2005)¹⁴⁷, la structure du domaine B' pourrait jouer un rôle important dans l'activité polymérase. Les saccharose-hydrolases possèdent un domaine B' présentant un faible pourcentage d'identité (32%) avec celui de l'amylosaccharase de *Neisseria polysaccharea* (ASNP). Cependant, leur pourcentage de similarité reste très élevé (83%). Il apparaît donc difficile de conclure sur les éléments structuraux responsables de l'activité polymérase des amylosaccharases. Ils pourraient être attribués au domaine B' et/ou à des acides aminés clés impliqués dans l'activité transglucosidase de cette enzyme (cf partie III.B.2.d.2., page 69). Enfin, notons que les 2 domaines B et B', boucles d'extension du domaine catalytique sont des domaines très conservés en terme de similarité de

séquences au sein de la famille GH13_4. A ce jour, l'AS de *Neisseria polysaccharea* est la plus étudiée et la mieux caractérisée de ces enzymes.

B.2.b. Séquences consensus de la famille GH13.4

En réalisant un alignement des 34 séquences d'amylosaccharases ou saccharose-hydrolases putatives à l'aide du programme MUSCLE http://phylogenomics.berkeley.edu/cgi-bin/muscle/input_muscle.py et en le soumettant ensuite au logiciel WebLogo (<http://weblogo.berkeley.edu/logo.cgi>), on obtient une représentation très visuelle des motifs conservés, appelée séquence logo (**FIGURE 20**). Celle-ci met en lumière un grand nombre de résidus strictement conservés sur les 34 séquences analysées. Notamment, elle permet de visualiser **les 11 motifs consensus** définis ci-après pour cette famille. Tous ces motifs couvrent l'ensemble des résidus clés de l'AS de *Neisseria polysaccharea* qui ont été identifiés. Ces résidus seront présentés par la suite.

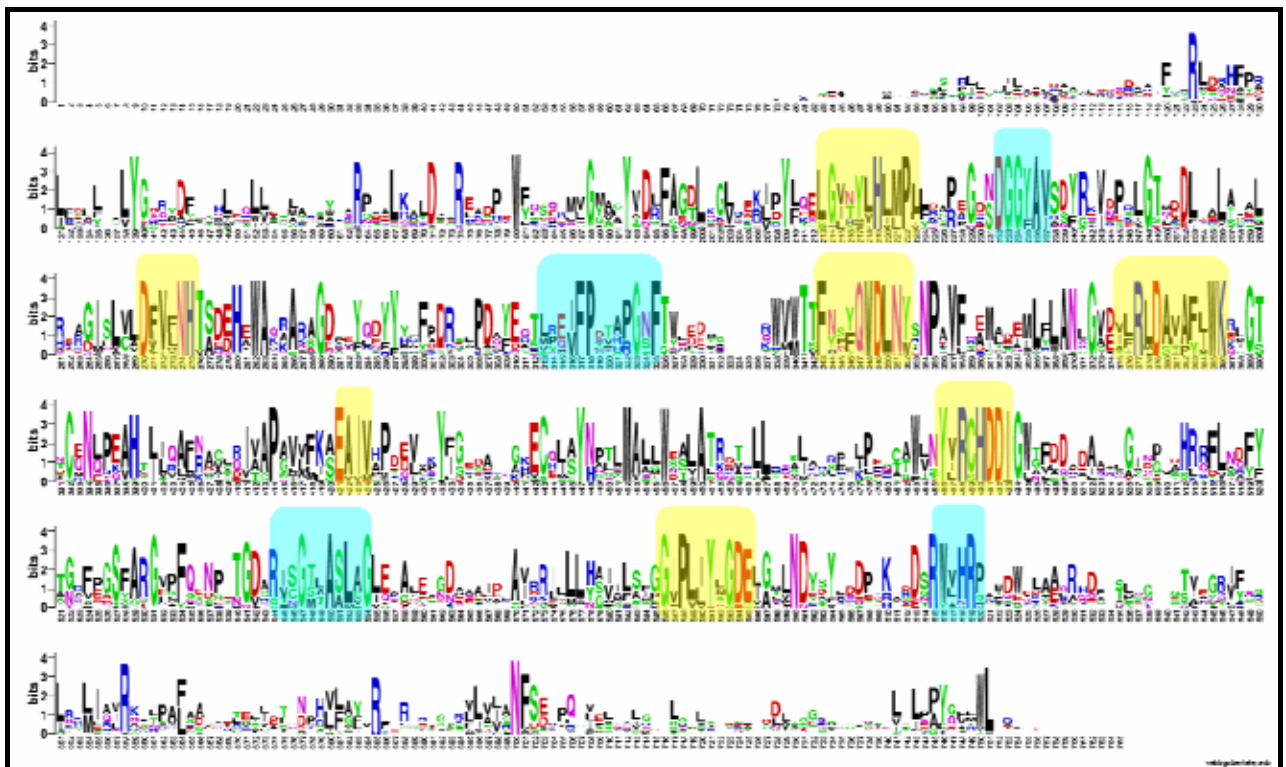


FIGURE 20 | Séquence logo des amylosaccharases et saccharose-hydrolases (famille GH13.4)

Les 7 séquences consensus relatives à la famille GH13 sont surlignées en jaune; les 4 propres à la famille GH13.4 le sont en bleu.

B.2.b.1. Les sept séquences « signature » de la famille GH13

Les alignements de séquences des enzymes de la famille 13 des glycoside-hydrolases ont permis de révéler l'existence de 4 séquences consensus (**I à IV**), positionnées près de l'extrémité des brins $\beta 3$ (**I : 182-187**), $\beta 4$ (**II : 282-293**), $\beta 5$ (**III : 328-331**) et sur la **boucle 7 (IV : 388-395)**²²⁸⁻²³⁰. Ces séquences renferment les **7 résidus les plus conservés** de la famille GH13: les trois acides carboxyliques catalytiques (Asp286, Glu328 et Asp393 selon la numérotation de l'ASNP) (régions II, III et IV), les deux histidines de stabilisation de l'intermédiaire (His187 et His392) (région I et région IV) et les deux résidus impliqués dans le maintien de l'intégrité structurale (Asp182 et Arg284) (régions **I** et **II**). Seuls les trois résidus constituant la triade catalytique (deux acides aspartiques et un acide glutamique) sont strictement conservés dans la famille²²⁸.

Une séquence consensus additionnelle (région **V : 250-258**), proposée par Janecek²²⁸, est située près de l'extrémité C-terminale du domaine B. Elle renferme un acide aspartique impliqué dans la fixation de l'ion calcium trouvé invariablement chez les α -amylases^{231, 232}. Janecek^{228, 233} fait état de 2 autres séquences consensus au niveau des brins $\beta 2$ (région **VI : 125-134**) et $\beta 8$ (région **VII : 480-489**) qui contiennent respectivement, une glycine et une proline (Gly126 et Pro134), et une autre glycine (Gly480). Ces 3 résidus, fréquemment rencontrés chez les enzymes de la famille GH13, interviendraient dans le maintien de l'intégrité structurale de ces enzymes.

Six de ces 7 motifs sont situés sur des brins β du tonneau catalytique $(\beta/\alpha)_8$, le 7^{ème} appartenant au domaine B des ASs ou SUHs (*cf partie III.B.2.d., page 70*).

B.2.b.2. Les quatre séquences consensus additionnelles de la famille GH13.4

Trois zones additionnelles propres aux amylosaccharases et aux saccharose-hydrolases ont été décrites²³⁴. Elles sont situées sur les **boucles 2 (région VIII : 144-149) et 8 (région X : 509-514)** du tonneau $(\beta/\alpha)_8$. On y retrouve les acides aminés clés conservés Asp144 et Arg509 formant le pont salin (*cf. partie III. D.2.d.2., page 70*). La troisième zone remarquable (région **IX : 225-237**) est située au niveau du domaine B. Elle renferme notamment l'arginine 226, dont le rôle sera explicité par la suite (*cf. partie III.B.2.d.2.b., page 74*).

A ces trois motifs, nous avons ajouté une quatrième zone (région **XI : 446-456**) localisée sur le domaine B', domaine spécifique à cette famille. Elle contient notamment l'arginine 446 impliquée dans l'activité de transglucosylation de l'ASNP (*cf. partie III.B.2.d.2.a., page 72*).

En utilisant l'ensemble de ces 11 séquences consensus, on obtient une séquence consensus globale de la famille GH13.4 présentée en **FIGURE 21**. Sur la base de cette séquence composée de 96 résidus, les 34 amylosaccharases ou saccharose-hydrolases partagent une moyenne de 79% d'identité (de 67% à 86%) et de 90% de similarité (de 83 à 93%).

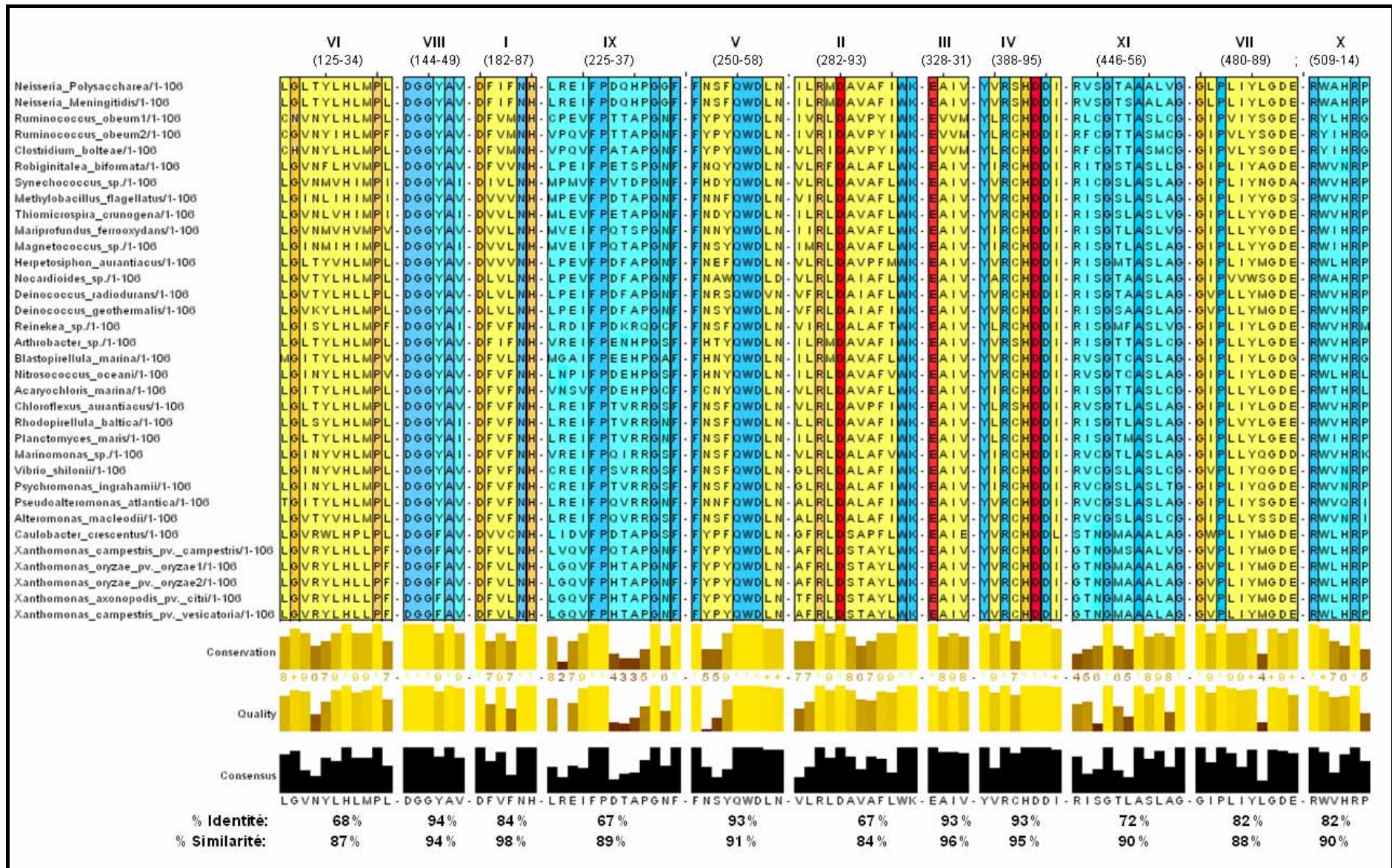


FIGURE 21| Alignement des séquences consensus des GH13.4 et identification des résidus conservés

En jaune, les 7 séquences consensus de la famille GH13; en bleu cyan, les 4 séquences consensus de la sous famille GH13.4.

En rouge, la triade catalytique; en orange, les résidus strictement ou hautement conservés de la famille GH13 ; en bleu ciel, les résidus strictement conservés de la sous famille GH13.4.

Les 3 dernières lignes indiquent respectivement la séquence consensus des 11 motifs ainsi que le % d'identité et de similarité de chaque motif.

B.2.c.Caractéristiques biochimiques de l'ASNP

B.2.c.1. Réactions catalysées à partir de saccharose seul

A partir de saccharose, l'ASNP catalyse la formation (i) d'amylose insoluble, (ii) de maltooligosaccharides, (iii) de fructose, (iv) de glucose issu de l'hydrolyse du saccharose et (v) de turanose et de tréhalulose^{235,191} (FIGURE 22). Le milieu de synthèse est donc caractérisé par la présence de **deux phases, l'une soluble et l'autre insoluble**. La fraction soluble présente des maltooligosaccharides de degré de polymérisation compris entre 2 et 25 et des traces de DP supérieur ; la fraction insoluble est composée de chaînes d'amylose pouvant atteindre un DP de 90. La répartition des produits formés en fin de réaction à partir de saccharose 100 mM et analysés en chromatographie d'exclusion de taille est présentée en FIGURE 22.

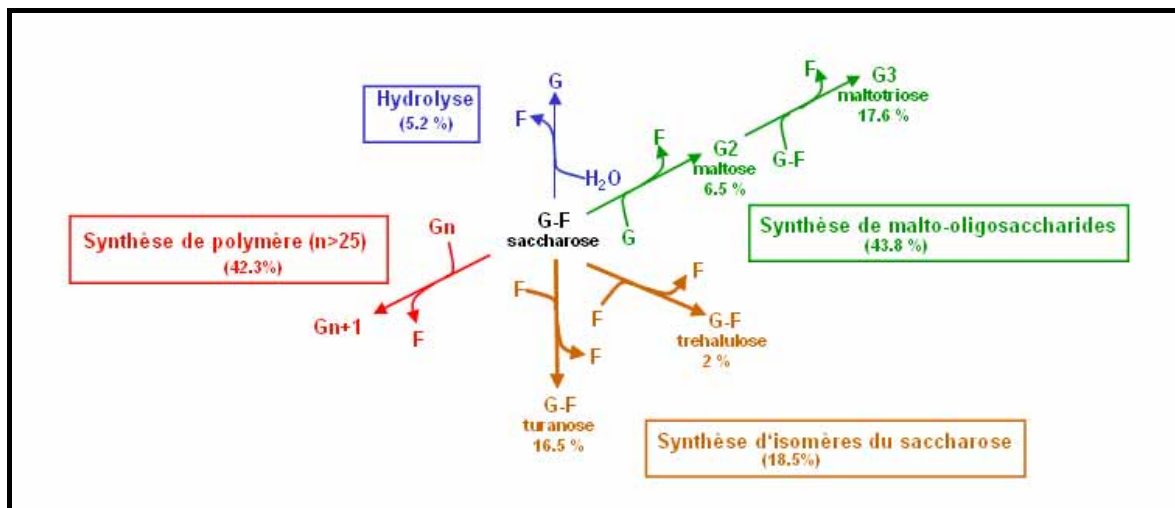


FIGURE 22| Réactions catalysées par l'ASNP à partir de saccharose seul (100 mM) et répartition des produits.

Le mécanisme de synthèse d'amylose a été élucidé grâce à l'étude de la structure tridimensionnelle de l'ASNP et par l'analyse biochimique du suivi cinétique de la formation du polymère¹⁹¹. L'étude des composés des fractions solubles et insolubles a permis de révéler que la formation d'amylose résulte de **l'élongation non-processive des maltooligosaccharides** produits par l'enzyme.

B.2.c.2. Paramètres cinétiques

La production de l'ASNP et sa purification jusqu'à homogénéité ont permis la caractérisation fine des paramètres cinétiques. La vitesse de consommation du saccharose (au même titre que l'évolution de la vitesse de formation du glucose et du polymère) par l'AS de *N. polysaccharea* ne présente pas un profil Michaelien classique²³⁵. Les auteurs ont proposé de modéliser le comportement de l'enzyme par **deux équations de Michaelis-Menten**, l'une ou l'autre étant employée pour des concentrations en saccharose supérieures ou inférieures à une valeur seuil (20 mM). Ce comportement atypique pourrait être dû à la présence d'un second site, non-catalytique,

de fixation du saccharose. Son occupation, aux fortes concentrations en saccharose, induirait des modifications conformationnelles favorables à l'activité de l'enzyme, et en particulier à la réaction de polymérisation. Les valeurs des constantes cinétiques apparentes sont présentées dans le **TABLEAU 5**. Elles révèlent que l'amylosaccharase présente une faible efficacité catalytique. Par ailleurs, une diminution de l'activité initiale a été observée pour des concentrations en saccharose supérieures à 300 mM. Ce comportement peut être attribué à une inhibition par excès de substrat ou une limitation du transfert de matière due à la viscosité du milieu réactionnel²³⁴.

TABLEAU 5 | Valeurs des constantes cinétiques apparentes des vitesses initiales de consommation du saccharose (ViS), d'hydrolyse du saccharose (ViG) et de polymérisation (ViG_n)
(D'après De Montalk et coll.¹⁴⁶)

| [Saccharose] initiale | Constantes cinétiques | V _{iS} | V _{iG} | V _{iG_n} |
|-----------------------|-----------------------------------|--|---|---|
| < 20 mM | K_m V_{max}^b k_{cat} | 1,9 mM 470 U/g 0,6 s ⁻¹ | 1,7 mM 288 U/g 0,3 s ⁻¹ | 1,9 mM 147 U/g 0,2 s ⁻¹ |
| > 20 mM | K_m V_{max}^b k_{cat} | 50,2 mM 1100 U/g 1,3 s ⁻¹ | 38,7 mM 472 U/g 0,6 s ⁻¹ | 387 mM 1620 U/g 1,9 s ⁻¹ |

^b la vitesse maximale est exprimée en unité par gramme de protéine pure. L'unité étant, selon le cas, des micromoles de saccharose consommé par minute ou des micromoles de glucose libéré, ou incorporé dans le polymère, par minute.

B.2.d. Structure de l'amylosaccharase de *Neisseria polysaccharea*

B.2.d.1. Structure tridimensionnelle de l'ASN_P sauvage

La séquence du gène rapportée par Potocki de Montalk et coll. (1999)¹⁴⁶ comporte 1911 paires de bases, codant pour une protéine de 636 acides aminés (environ 70 kDa). L'optimisation du procédé de purification a permis de purifier l'ASN_P recombinante jusqu'à homogénéité, en quantité suffisante pour la détermination par cristallographie et diffraction aux rayons X de sa structure tridimensionnelle. Celle-ci a été obtenue en 2001 avec une résolution de 1,4 Å¹⁷⁹. Elle présente une molécule de TRIS liée au site actif. Il s'agit de **la seule structure d'amylosaccharase disponible à ce jour**. L'ASN_P présente une organisation structurale en cinq domaines nommés N, A, B, B' et C (**FIGURE 23**). Les trois domaines A, B et C sont communs aux α -amylases. Le domaine A, le domaine catalytique, est organisé en tonneau (β/α)₈ comme pour toutes les enzymes de la famille GH 13. Il comprend les résidus 98-184, 261-395, et 461-550. Le domaine B correspond à la boucle 3 de ce tonneau. Il s'étend des résidus 185 à 260 et comporte 2

feuilletés β anti-parallèles incluant chacun une hélice α . Enfin, le domaine C (résidus 555 à 628) comporte exclusivement des brins β qui adoptent une structure en sandwich ou en clé grecque.

L'ASNP possède **deux domaines supplémentaires qui lui sont spécifiques** : les domaines B' et N. Le domaine B' (résidus 395-460) résulte, comme le domaine B, d'une extension du tonneau $(\beta/\alpha)_8$. Il s'agit de la boucle 7, comportant 2 hélices α suivies d'un brin β et d'une autre hélice α . Le domaine N-terminal (résidus 1-90) est composé de 6 hélices α . Deux de ces hélices soutiennent α_3 et α_4 du tonneau. Elles contribueraient au maintien de l'intégrité structurale. Par ailleurs, l'ASNP ne contient pas de pont di-sulfure malgré la présence de 6 résidus cystéine dans sa séquence polypeptidique.

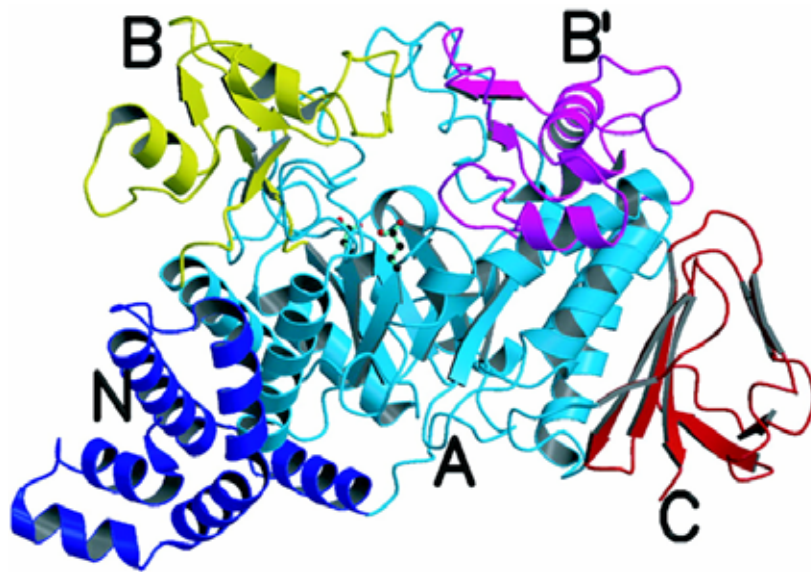


FIGURE 23| Structure tridimensionnelle de l'amylosaccharase de *N. polysaccharea*
Légendes des domaines: N (bleu sombre), A (bleu clair), B (jaune), B' (magenta) et C (rouge). Les résidus catalytiques Asp286 et Glu328, respectivement localisés à la fin des feuilletés β -4 et β -5 dans le domaine A, sont également indiqués.
 (D'après Skov et coll.¹⁷⁹)

Enfin, signalons que la structure de la SUH de *Xanthomonas campestris pv. campestris* a été résolue en 2004²³⁶. Cette enzyme, qui contient les 5 domaines précédemment cités est structurellement très proche de l'ASNP. En revanche, elle présente un site actif dont la topologie est beaucoup plus ouverte que celui de l'ASNP et, de ce fait, pourrait justifier son absence d'activité transférase voire polymérase.

B.2.d.2. Facteurs structuraux responsables de la spécificité catalytique

L'arrimage d'un substrat au niveau du site actif repose sur une **complémentarité stéréospécifique et électronique** mettant en jeu deux principaux types d'interactions : les liaisons hydrogène et les interactions de van der Waals²³⁷. Lorsque le substrat est glucidique, un phénomène important, le « **stacking** » (empilement ou sandwich), participe également à son orientation. Il s'agit de l'ensemble des forces mises en jeu (van der Waals et hydrophobes) entre

un cycle aromatique (chaîne latérale de résidus tels que la tyrosine, la phenylalanine ou le tryptophane) et les atomes de carbone du cycle de la molécule de sucre, dont les plans sont disposés parallèlement. Ces résidus aromatiques, encombrants, constituent des points d'ancrage clé pour le substrat, permettant son arrimage ou pré-arrimage ou son orientation vers le site catalytique.

Lors de cet arrimage, on considère isolément les interactions mises en jeu au niveau de chaque unité de sucre. On parle de **sous-sites de fixation**, nomenclature utilisée principalement pour les enzymes de dépolymérisation comme les glycoside-hydrolases mais aussi les protéinases et les nucléases²³⁸⁻²⁴¹. Un sous-site est formé d'un ou de plusieurs résidus et contribue plus ou moins fortement à la fixation de la molécule entière. Le nombre de sous-sites et leur énergie d'interaction peuvent se mesurer expérimentalement et sont intimement liés aux propriétés catalytiques des enzymes et à leur spécificité de substrat.

Une nomenclature a été proposée pour les sous-sites de fixation des glycoside-hydrolases²⁴¹. Leur nombre peut varier de 2, dans le cas des monoglycosidases et disaccharidases, à 10 dans le cas de certaines α -amylases. Les sous-sites sont numérotés de (-n) à (+n), -n représentant la partie non-réductrice et +n la partie réductrice. La rupture de la liaison osidique s'opère entre les positions (-1) et (+1). Dans le cas de l'ASNP, les sous-sites (-n) à (-2) sont inexistants (**FIGURE 24**).

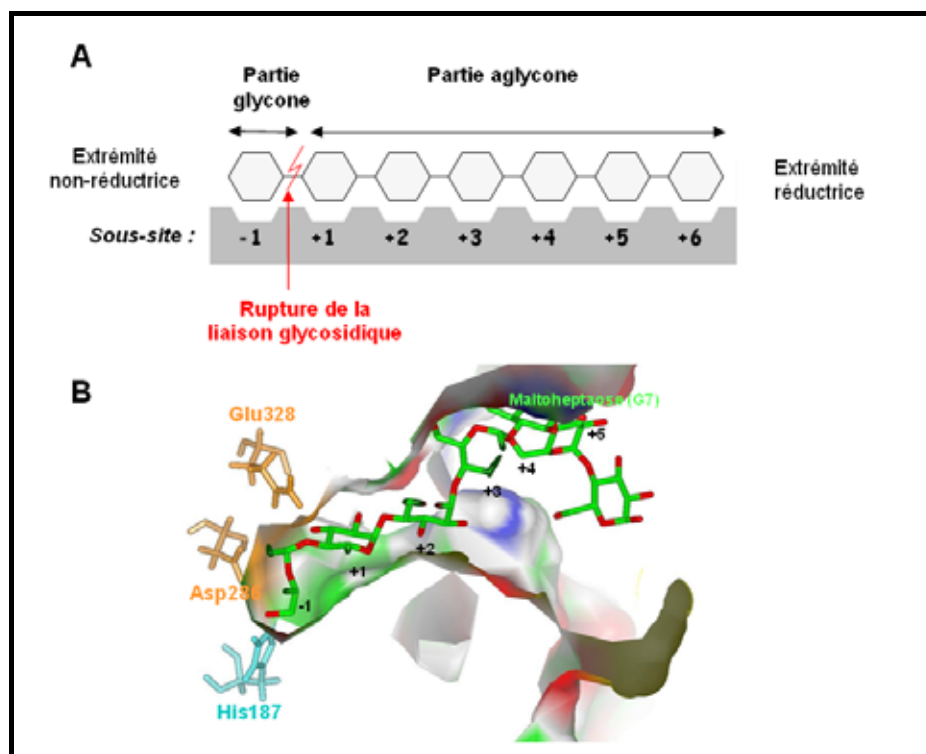


FIGURE 24| Représentation des sous-sites de liaison de l'ASNP

A: Représentation schématique d'après la nomenclature proposée par Davies et coll.²⁴¹
 B: Représentation en coupe du site actif de l'AS complexée avec le maltoheptaose (G7)

B.2.d.2.a Architecture du site actif

L'analyse des complexes E328Q (mutant inactif) : saccharose (**FIGURE 25A**) et AS : glucose²⁴² (**FIGURE 25B**) associée à des travaux de mutagenèse dirigée^{243, 244} a permis d'identifier les acides aminés clés du site actif, responsables de la spécificité de l'enzyme envers son substrat.

Ainsi, l'architecture du site actif de l'ASNP est définie par les **sous-sites (-1) et (+1)** occupés respectivement par les résidus glucosyle et fuctosyle du saccharose:

- **Le sous-site (-1) assure le maintien de l'unité glucosyle** (qui sera transférée) par un ensemble d'interactions fortes (liaisons hydrogène et stacking) impliquant **un pont salin et deux résidus aromatiques**. Le pont salin situé en amont de ce sous-site est formé par les résidus **Asp144 et Arg509**. Cet élément remarquable de l'ASNP empêche l'accès du substrat vers d'éventuels sous-sites (-2) ou (-3), tels que ceux rencontrés chez les α -amylases. En délimitant le fond de ce sous-site, il confère au site actif une topologie en forme de poche (**FIGURE 24**). Une telle architecture est également due au domaine B', qui couvre partiellement l'entrée du site actif. Situés de part et d'autre du cycle glucosyle, les 2 résidus aromatiques **Tyr147 et Phe250** renforcent la stabilisation de l'intermédiaire covalent par un double stacking de type "sandwich". A ces résidus essentiels à l'arrimage du substrat, s'ajoutent les histidines **His187 et His392**, qui assistent la catalyse au même titre que le troisième acide carboxylique **Asp393** de la triade catalytique. Cet acide aminé, Asp 393, est situé à la jonction des sous-sites (-1) et (+1). Mentionnons que les 2 autres acides aminés catalytiques **Asp286** (le nucléophile) et **Glu328** (donneur de proton) sont localisés de part et d'autre de Asp393 ie. au niveau des sous-sites (-1) et (+1) respectivement. La mutation individuelle de tous ces résidus clés^{243, 244} conduit à une perte très conséquente voire totale d'activité, confirmant leur rôle dans la fixation du saccharose, dans la formation ou la stabilisation de l'intermédiaire covalent.

- **Le sous-site (+1) ne maintient que très faiblement le cycle fructosyle**. Les résidus **Asp394 et Arg446** assurent l'essentiel des interactions avec ce résidu mais ne contribuent pas de façon majeure à la spécificité envers le saccharose. En effet, la fragilisation du réseau de liaisons hydrogène au niveau du sous-site (+1) entraînée par la mutation de Asp394 ou Arg446 en alanine affecte l'activité envers le saccharose mais ne la détruit pas complètement²⁴⁴. Après formation de l'intermédiaire covalent glucosyl-enzyme, le cycle fructosyle est probablement facilement exclu du site actif, laissant ainsi place à l'arrimage des molécules acceptrices. Le rôle de ces résidus Asp394 et Arg446 est alors beaucoup plus déterminant dans cette 2^{ème} étape (étape de transfert) puisqu'ils assureront en grande partie la fixation des résidus glucosyle accepteurs et leur positionnement correct. A ce titre, ils jouent une fonction clé dans la réaction de transglucosylation, étant responsables de la spécificité de produit de l'ASNP²⁴⁴. Leur mutation affecte essentiellement la réaction de transglucosylation et favorise la réaction d'hydrolyse du

saccharose au détriment de la polymérisation²⁴⁴. Notons aussi que l'arginine 446 présente dans les ASs est remplacée par une glycine dans les SUHs du genre *Xanthomonas*. Ce résidu semble donc être un des éléments à considérer dans la différence d'activité de ces 2 classes d'enzymes.

Alors que les résidus du sous-site (-1) interagissent très spécifiquement avec le substrat donneur, le sous-site (+1) présente une topologie beaucoup plus tolérante à la fixation de molécules acceptrices autres que l'accepteur naturel. Cette flexibilité est illustrée par la capacité de cette enzyme à glucosyler une large gamme de dérivés saccharidiques. (cf. partie III.B.1.d.1. page 61).

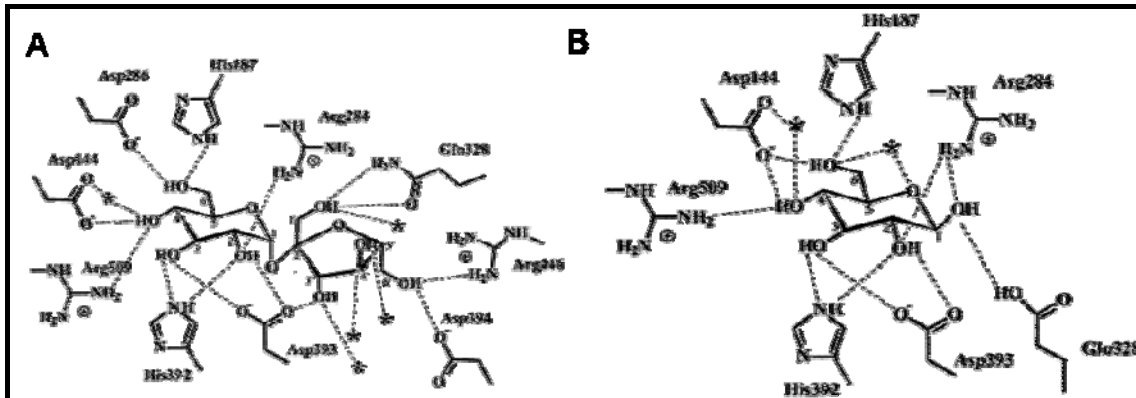


FIGURE 25 | Représentation schématique des interactions (A) entre le mutant inactif E328Q et le saccharose et (B) entre l'ASNPN et le D-Glucose.
(D'après Mirza et coll.²⁴²)

Les liaisons hydrogène sont représentées en pointillées et les * représentent les molécules d'eau.

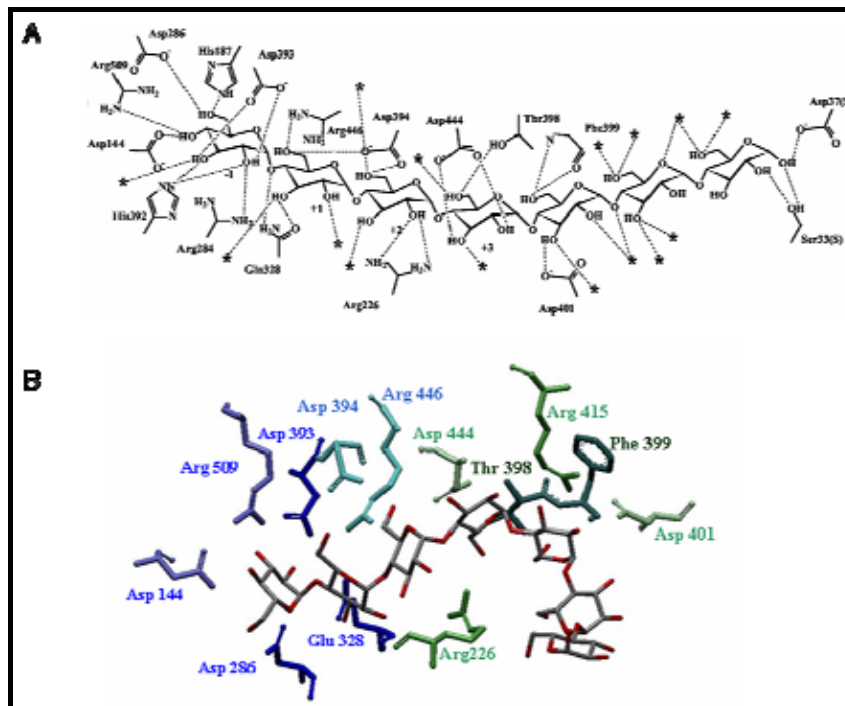


FIGURE 26 | Interactions du site OB1.
(D'après Skov et coll.²⁴⁵)

(A) Représentation schématique des liaisons hydrogène (les * représentent les molécules d'eau et les interactions du site +6 proviennent d'une autre molécule d'ASNPN)
(B) Représentation spatiale de certains résidus (bleu : résidus aussi impliqués dans la fixation du saccharose, vert : autres résidus du site OB1).

B.2.d.2.b. De multiples sous-sites multiples de fixation des oligosaccharides

La co-cristallisation du mutant inactif E328Q de l'ASNP avec du maltoheptaose (susceptible de simuler une chaîne en cours d'élongation) a permis de mettre en évidence les sites accepteurs (+1) à (+6) (site OB1)²⁴⁵. La résolution du complexe E328Q : maltoheptaose montre que l'oligosaccharide prend place au fond du site actif, l'extrémité non réductrice de la chaîne occupant le site (-1) dans une conformation identique à celle observée dans les complexes AS : glucose et E328Q : saccharose²⁴². Les autres unités glucosyle prennent place dans le site aglycone (+n). L'analyse des sous-sites +1, +2, +3, +4 et +5 ne révèle aucune interaction de "stacking". Les résidus Arg226 (+2), Asp444 et Thr398 (+3), Phe399 et Asp401 (+4) sont impliqués dans des liaisons hydrogène avec le maltoheptaose. Les **sous-sites de (+1) à (+3)** sont cependant **considérés comme faibles**, notamment du fait de **l'encombrement dû à Arg226**. Le remplacement de cette arginine par un résidu moins encombrant (alanine) permet de diminuer la gêne stérique au niveau du sous-site (+2) et d'augmenter la glucosylation du maltose et du maltotriose. Cette mutation convertit l'AS en une **polymérase plus performante (FIGURE 26)**. L'arginine 415 du sous-site (+4) serait en revanche considérée comme un site d'ancrage fort pour les molécules acceptrices.

C. Techniques d'Ingénierie et Applications**C.1. Ingénierie des protéines : Pourquoi ?**

Le terme "ingénierie des protéines" s'applique à la modification volontaire de la structure d'une protéine induisant une altération de ses propriétés fonctionnelles. Cette discipline est destinée à améliorer la performance des enzymes ciblées. Il s'agit notamment **d'une étape quasi-incontournable** pour adapter les propriétés de ces biocatalyseurs aux **contraintes et demandes des procédés industriels** :

- **Conditions de synthèse** (pH, température et solvants organiques). Suivant leur domaine d'applications, les enzymes devront être optimisées de façon à fonctionner dans des conditions différentes de leur environnement naturel.
- **Rendement de synthèse** qui se situe à plusieurs niveaux: le taux d'expression de l'enzyme dans son hôte, sa stabilité, son efficacité et son énantiosélectivité sont autant d'exemples de caractéristiques à améliorer pour atteindre des rendements de production optimaux.
- **Spécificités de synthèse** : modifier l'activité de l'enzyme envers des substrats non naturels (spécificité de substrat, élimination de fonctions enzymatiques indésirables)

revêt un intérêt particulier pour élargir le domaine d'applications d'une enzyme et accéder à de nouveaux composés.

C.2. Ingénierie des protéines : Comment ?

L'ingénierie des protéines a débuté par l'utilisation de méthodes chimiques ou enzymatiques visant à modifier la protéine elle-même. Dans les années 1980, l'introduction des outils d'ingénierie génétique a largement contribué au développement de cette discipline en utilisant comme cible de modification le gène codant pour la protéine étudiée. Par la suite, nous ne nous focaliserons que sur les modifications générées par génie génétique.

A l'heure actuelle, trois approches d'ingénierie sont suivies :

- **La conception rationnelle (" rational design ")** : cette approche définit des cibles précises de mutagenèse dirigée (du résidu au domaine entier). **Le succès de cette stratégie repose essentiellement sur le choix de ces résidus.** Or, ces cibles restent particulièrement difficiles à prédire, lorsque l'amélioration porte sur l'adaptation à des températures, à des pH ou à des solvants inhabituels. Cette approche sera donc plus appropriée à des fonctions en lien direct avec le site actif de l'enzyme (énantioselectivité, efficacité, spécificité de substrat). Les connaissances structurales relatives à la protéine étudiée combinées aux alignements de séquence de protéines appartenant à la même famille serviront à prédire au mieux ces cibles. Notamment, la résolution de structures tridimensionnelles de protéines, seules ou complexées avec leurs ligands a largement contribué à la compréhension des relations structure/fonction des protéines. Plus précise que l'alignement de séquences, la structure 3D permet l'identification des résidus ou régions impliqués dans la catalyse et des liaisons en interaction avec le substrat. En complément des données structurales, la méthode nommée " scanning mutagenesis " est une option pour explorer l'effet de la mutation sur une zone de résidus potentiellement candidats. Il s'agit de substituer chacun de ces résidus individuellement par un seul acide aminé et d'analyser fonctionnellement le mutant exprimé. L'alanine est généralement utilisé comme substitut dans la mesure où il ne modifie pas l'orientation de la chaîne peptidique et n'introduit aucune interaction stérique ou électrostatique. On parle alors d'" alanine scanning²⁴⁶ ". Cette méthode permet aussi d'évaluer l'implication de chaque acide aminé dans une fonction précise.

- **L'ingénierie combinatoire ou évolution dirigée** est une approche expérimentale qui vise à modifier les propriétés des protéines par la mise en pratique des principes de l'évolution naturelle que sont la variation génétique et la sélection phénotypique. Elle ne requiert aucune connaissance structurale de la protéine d'intérêt et peut s'appliquer théoriquement à la recherche de toute

nouvelle propriété. Dans un premier temps, des banques de variants sont construites par mutagenèse aléatoire à partir de la séquence ADN codant pour la protéine d'intérêt. Il s'agit, dans cette première étape, **de créer la diversité**. Dans un second temps, ces banques sont soumises à des protocoles de sélection et/ou de criblage de manière à isoler les mutants qui présentent un gain pour la propriété ciblée. Cette deuxième étape consiste donc à **trier la diversité** en testant simultanément (sélection) ou individuellement (criblage) un grand nombre de variants sur la propriété désirée. En utilisant les mutants améliorés pour la construction de nouvelles banques, ce schéma expérimental peut être reproduit de façon itérative jusqu'à ce que le niveau de performance choisi par l'expérimentateur soit atteint.

Lors de la première étape, les méthodes permettant de créer de la diversité sont multiples. Citons les agents mutagènes chimiques²⁴⁷ (bisulfite de sodium, hydroxylamine, méthoxylamine...) et physiques^{248, 249} (rayonnements UV), les souches mutatrices²⁵⁰ et les ADN polymérases infidèles. Ces dernières sont à la base de la technique de mutagenèse aléatoire la plus couramment utilisée actuellement dans les projets d'évolution dirigée : la PCR à erreur (ou "error prone PCR"). Cette technique consiste à employer la technique de PCR dans des conditions réactionnelles réduisant la fidélité de l'ADN polymérase. L'enzyme la plus utilisée dans les protocoles d'epPCR est la *Taq* polymérase. Son taux d'erreurs varie de 10^{-4} (1 erreur pour 10^4 bases répliquées) à 10^{-6} (1 erreur pour 10^6 de bases) en fonction des conditions réactionnelles²⁵¹⁻²⁵³. Les principaux paramètres qui permettent de diminuer la fidélité de la *Taq* polymérase sont l'ajout de $MnCl_2$ ²⁵⁴ (à la place de $MgCl_2$), l'utilisation de concentrations biaisées en dNTPs^{255, 256}, l'augmentation de la concentration en enzyme ou celle du temps de polymérisation²⁵⁵ ou encore l'utilisation d'analogues de nucléoside triphosphate²⁵⁷.

Cette première étape consiste à générer au mieux l'espace des séquences possibles. A titre d'exemple, le nombre de mutants substitués par 5 résidus dans une protéine de 100 acides aminés est d'environ 2×10^{14} . Si cette taille de banques peut être difficilement atteinte (10^4 à 10^{12} en fonction des étapes expérimentales (clonage, transformation de microorganismes etc.), l'exploration de cette diversité sera avant tout limitée par les techniques de sélection et/ou de criblage, qui permettent d'analyser au maximum 10^{12} variants. Aussi, lorsque la propriété recherchée le permet, la **mutagenèse aléatoire par cassette** est une solution envisageable pour explorer des zones plus restreintes comme le domaine catalytique. La recherche du mutant le mieux adapté peut également être accélérée en recombinant les variants positifs identifiés au premier cycle par la technique du "**DNA shuffling**"²⁵⁸. Cette **redistribution de l'ADN** peut également permettre de construire des banques de variants hybrides. Appliquée à des séquences issues de protéines homologues (au minimum 60 à 70% d'identité), le "**family shuffling**"²⁵⁹ vise à générer des hybrides actifs et parmi eux, des variants présentant des propriétés améliorées voire

nouvelles par rapport aux enzymes sauvages. Dans tous les cas, les variants ayant accumulé les mutations bénéfiques pour la propriété ciblée pourront ensuite être isolés par sélection et/ou criblage et subir un nouveau cycle de recombinaison/sélection/criblage.

Si la création de la diversité n'est pas un facteur limitant, son tri l'est davantage. L'efficacité d'une telle approche reposera en majeure partie sur les méthodes de sélection et/ou de criblage propres à chaque enzyme et propriété à améliorer. Ces méthodes doivent être suffisamment **rapides (haut-débit) et sensibles pour détecter efficacement les variants positifs.** Cette problématique doit être effectivement prise en compte dans notre étude puisqu'une seule technique de très haut-débit portant sur l'amélioration de la transglycosylation d'accepteurs et applicable à d'autres glycosyltransférases est connue à ce jour. En effet, en 2006, l'équipe de Withers a développé une méthode de sélection²⁶⁰ pour l'évolution dirigée des sialyltransférases (STs). Basée sur l'utilisation de la cytométrie en flux (**FACS** : Fluorescence-Activated Cell Sorting), leur technologie permet d'identifier par tri cellulaire à haut-débit les cellules qui expriment des variants optimisés pour la transglycosylation de divers accepteurs fluorescents. En criblant une large banque ($>10^6$), les auteurs isolent un variant extrêmement performant pour la glycosylation du D-Galactose et du 3SH-D-Lactose munis d'une étiquette fluorescente (facteur 400 par rapport aux accepteurs sans étiquette). Notons cependant que son principal inconvénient porte sur l'accepteur criblé qui **contient forcément une étiquette fluorescente, généralement volumineuse.**

- La **conception semi-rationnelle** qui vise à combiner les bénéfices de la conception rationnelle et de l'évolution dirigée. Elle exploitera au mieux les outils propres à ces 2 approches, à savoir les données structurales et fonctionnelles de l'enzyme et les techniques de criblage. Dans ce contexte, la **mutagenèse dirigée à saturation** est un réel apport pour explorer **intégralement** des zones cibles. Cette méthode met en œuvre des oligonucléotides synthétiques dégénérés qui introduisent **toutes les mutations possibles** sur un ou des codons choisis au préalable, voire sur l'ensemble des codons de la séquence parentale. Ce type de stratégie permet de contourner la dégénérescence du code génétique. L'approche la plus simple consiste à employer de façon équimolaire les quatre nucléotides (NNN où N = A, C, G ou T) au niveau des codons à modifier²⁶¹. Cependant, cette stratégie favorise la conservation des nucléotides originaux qui forment le plus grand nombre d'appariements de type Watson-Crick et s'hybrideront donc préférentiellement sur le brin matrice. Par ailleurs, l'utilisation d'un mélange équimolaire des quatre nucléotides entraîne également un autre biais en faveur des acides aminés possédant le plus d'entrées dans le code génétique. Pour réduire ce biais²⁶², l'utilisation de codons alternatifs, NN(G/T) ou NN(G/C), dans lesquels la troisième base est partiellement dégénérée, est souvent

envisagée, car ils permettent de représenter les 20 acides aminés par 32 codons différents (au lieu de 64) et ne génèrent qu'un seul codon stop (UAG). Enfin, la présence résiduelle de clones de type sauvage doit être également considérée. Aussi, en tenant compte de ces 3 paramètres (dégénérescence de la diversité génétique par rapport à la diversité protéique, biais d'incorporation introduits lors de l'étape de mutagenèse et taux de molécule sauvage résiduel), l'expérimentateur devra **surcribler la microbanque générée d'au moins un facteur 10 pour en explorer l'entière diversité protéique**. Ainsi, pour 1 position mutée (20 résidus possibles), 100 mutants devront être testés ; pour 2 positions mutées simultanément (400 résidus possibles), ce chiffre s'élève à 4000.

La mutagenèse à saturation pourra être employée pour muter chaque résidu individuellement sur une zone définie (technique **GSMM** : Gene Site Saturation Mutagenesis^{TM 263}) ou pour muter simultanément plusieurs résidus (technique **MM** : Massive Mutagenesis^{TM 264}). Elle pourra également s'intégrer à des techniques de recombinaison (shuffling synthétique²⁶⁵, techniques **DOGS**²⁶⁶ : Degenerate Oligonucleotide Gene Shuffling, et **ADO**²⁶⁷ : Assembly of Designed Oligonucleotides). Les amorces utilisées sont alors composées d'une partie non dégénérée correspondant à une zone conservée des séquences parentales et d'une partie dégénérée correspondant aux zones à explorer.

Enfin, nous proposons d'illustrer les apports de la conception semi-rationnelle par deux exemples. Les auteurs exploitent judicieusement leurs outils d'ingénierie et les données acquises au cours de l'évolution de leur enzyme pour générer des catalyseurs extrêmement performants :

Exemple de conception semi-rationnelle à dominante aléatoire :

Fox et coll. améliorent d'un facteur 4000 l'activité d'une déshalogénase²⁶⁸ en intégrant à l'approche aléatoire la stratégie d'optimisation **ProSAR** (**Protein Sequence Activity Relationship**). Cette stratégie est basée sur une analyse statistique qui attribue à chaque mutation identifiée (50 par cycle) un coefficient d'impact sur l'activité. Les mutations les plus bénéfiques sont maintenues dans le cycle suivant. Pour créer davantage de diversité à chaque cycle, plusieurs bibliothèques obtenues par des techniques différentes (mutagenèses aléatoire, à saturation, dirigée et recombinaison) sont créées et criblées visuellement sur support solide (~60 000 clones par cycle). Au final, la diversité explorée par cette approche se reflète par le nombre de mutations des meilleurs mutants (>35).

Exemple de conception semi-rationnelle à dominante rationnelle :

Les travaux du groupe de Keasling proposent une exploration fine du site actif²⁶⁹ d'une γ -humulène synthase, qui, à partir d'un seul substrat (le farnésyl diphosphate) produit 52 sesquiterpènes (8 étudiés) en différentes proportions. En mutant par mutagenèse à saturation et

individuellement les 19 résidus qui composent le site actif, les auteurs déterminent comment chaque résidu contribue à la haute flexibilité de cette enzyme et altère la distribution des 8 produits. A partir de cette base de données, ils formulent un algorithme, qui leur permettra de définir le jeu de mutations à combiner pour favoriser la synthèse d'un des 8 produits. Ils construisent alors 8 mutants, qui synthétisent spécifiquement 1 des 8 produits (rendement de 15 à 86%).

C.3. L'ingénierie des glycoside-hydrolases en vue de la modification de leur activité

Ce dernier paragraphe est consacré au remodelage des glycoside-hydrolases qui a mis à profit les 3 types d'ingénierie précédemment décrits pour en modifier leur activité.

C.3.a. Les glycosynthases, les thioglycoligases et les thioglycosynthases

Comme nous l'avons mentionné auparavant, la majorité des glycoside-hydrolases sont des glycosidases qui, limitées par leur forte activité hydrolytique, restent peu performantes pour la synthèse de glyco-conjugués. Des travaux d'ingénierie⁶⁷, aussi bien rationnels qu'aléatoires ont levé ce verrou et ont également permis d'élargir la spécificité de substrat de ces enzymes.

C.3.a.1. Leur conception : une approche rationnelle

Les glycoside-hydrolases utilisées en milieu hydro-organique conduisent à la synthèse d'oligosaccharides. Toutefois, les produits obtenus présentent généralement une affinité accrue pour l'enzyme par rapport au substrat de départ. Ils deviennent alors des substrats pour l'enzyme, qui les convertira en mélanges complexes de produits. Ce problème peut être désormais solutionné par l'utilisation de glycosynthases.

En 1998, Withers et coll. ont présenté ce concept²⁷⁰ basé sur la modification du fonctionnement d'une enzyme par la **mutation d'un des résidus catalytiques (FIGURE 27)**. Ainsi, les glycosynthases sont des glycoside-hydrolases dont l'acide aminé nucléophile est remplacé par un résidu non-nucléophile (alanine, glycine ou sérine) par mutagenèse dirigée. Activé en position anomérique par un bon groupe partant (généralement de type fluorure), le donneur de glycosyle se positionnera convenablement dans le site actif de l'enzyme et mimera l'intermédiaire glycosyl-enzyme. Le deuxième résidu catalytique, responsable de la catalyse acide/base est conservé, permettant d'activer l'hydroxyle de l'accepteur. Privées de leur résidu nucléophile, ces enzymes sont incapables d'hydrolyser le substrat ou tout autre produit accumulé dans le milieu. Cette mutation augmente donc leur capacité à réaliser des réactions de transglucosylation tout en conservant leur stéréosélectivité.

Depuis, ce principe a été appliqué avec succès à d'autres exoglycosydases de type β -D avec rétention de configuration, mais aussi à des endoglycosydases, qui présentent une meilleure régiosélectivité. En 2002, Okuyama et coll.²⁷¹ décrivent la première glycosynthase de type α -D agissant également avec un mécanisme de rétention de configuration. Plus récemment, en 2005, Honda et Kitaoka²⁷² étendent ce concept aux β -D-glycoside hydrolases procédant selon un mécanisme d'inversion de configuration. (**TABLEAU 6**).

Ces nouveaux outils de synthèse peuvent atteindre des rendements de glycosylation très performants (de 60 à 95%)^{270, 273-276}. Ils se distinguent également par la nature du groupe glycosyle (D-glucose, D-galactose, D-mannose, D-xylose, acides glucuronique et galacturonique) qu'ils transfèrent et par leur régiospécificité (β -1,3/4/6 et α -1,4) (**TABLEAU 6**). De ce fait, cette biodiversité peut être exploitée pour la synthèse d'oligosaccharides d'intérêt. Citons, à titre d'exemple, les motifs glycosidiques (β -1,3)²⁷⁷ retrouvés dans de nombreux antigènes ou molécules de signalisation chez les plantes et les conjugués d'acides glucuronique et galacturonique²⁷⁸, composants essentiels des systèmes vivants, qu'ils soient bactériens et végétaux (parois cellulaires) ou animaux (tissus conjonctifs (glycosaminoglycanes)).

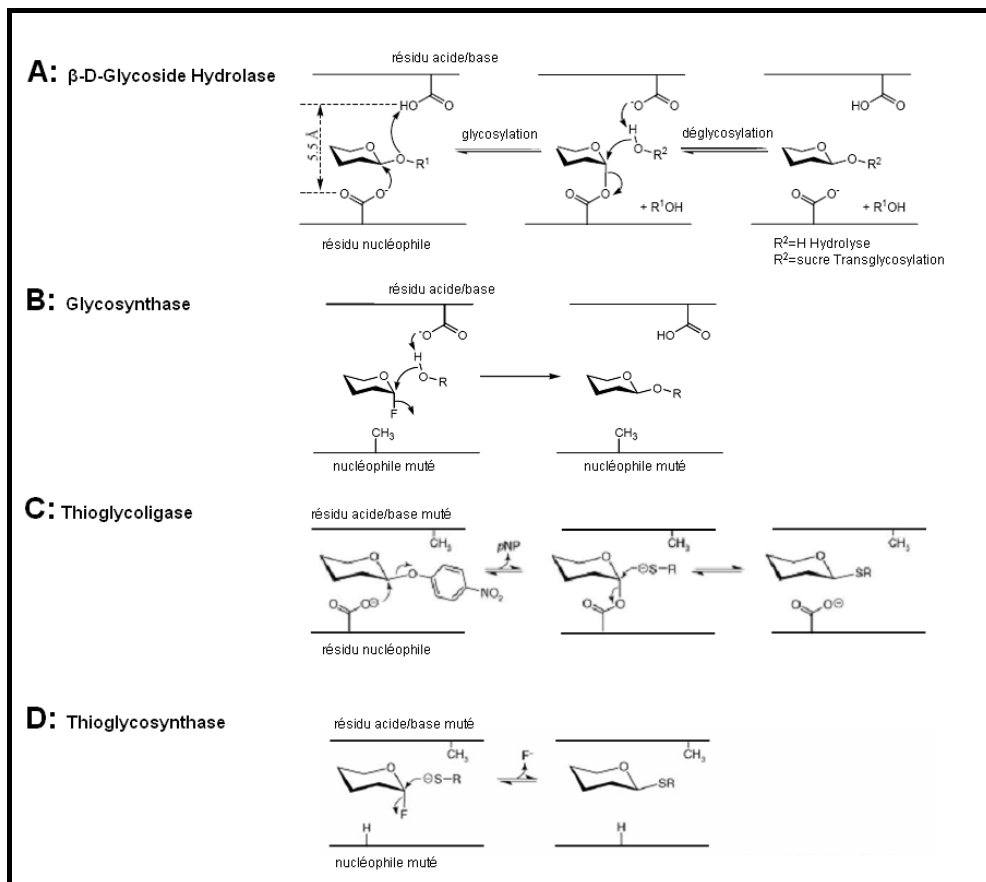


FIGURE 27 | Mécanismes d'action d'une glycoside-hydrolase agissant avec rétention de configuration (cas d'une β -D-glycoside hydrolase) (**A**), d'une glycosynthase (**B**), d'une thioglycoligase (**C**) et d'une thioglycosynthase (**D**)

TABLEAU 6 | Glycosynthases, thioglycoligases et thioglycosynthases répertoriées (décembre 2007).

(Abréviations : Glc, D-Glucose; Gal, D-Gal; Fuc, D-Fucose; Man, D-Mannose; Xyl, D-Xylose; X, Xyl(α -1,6)Glc β ; L, Gal(β -1,2)Xyl(α -1,6)Glc β ; -F, Fluorure; -oNP, ortho nitrophényl; -pNP, para nitrophényl; - DNP, dinitrophényl; -S, thio)

| Organisme | Enzyme native | GH | Activité catalytique | Mutations | liaisons formées | Donneurs (D) Accepteurs (A) | Mécanisme | Endo/Exo | Refs |
|---|---|------|---|-----------------------------------|--------------------------------|---|-----------|----------|------------------------|
| <i>Agrobacterium</i> sp. | β -glucosidase (Abg) | GH1 | Glycosynthase Thioglycoligase Thioglycosynthase | E358A/G/S E171A E358G/E171A | β -1,3/4 | D: Glc α F; Gal α F D: Glc β -DNP A: 4SGlc β -pNP D: Glc α F A: 4SGlc β -pNP | RET | Exo | 270, 273 279 280 |
| <i>Sulfolobus solfataricus</i> | β -glucosidase (Ss- β -Gly) | GH1 | Glycosynthase | E387A/G | β -1,3/4/6 | D: Glc α F; Glc β -oNP; Gal β -oNP; Fuc β -oNP | RET | Exo | 281, 282 |
| <i>Thermosphaera aggregans</i> | β -glycosidase (Ta- β -Gly) | GH1 | Glycosynthase | E386G | β -1,3/4/6 | D: Glc β -oNP; Gal β -oNP | RET | Exo | 283 |
| <i>Pyrococcus furiosus</i> | β -glycosidase (CelB) | GH1 | Glycosynthase | E372A | β -1,3/4/6 | D: Glc β -oNP | RET | Exo | 283 |
| <i>Thermus thermophilus</i> | β -glycosidase (Tt- β -Gly) | GH1 | Glycosynthase | E338A/G/S | β -1,3 | D: Glc α F; Gal α F | RET | Exo | 277 |
| <i>Streptomyces</i> sp. | β -glucosidase | GH1 | Glycosynthase | E383A | β -1,3/4 | D: Glc α F; Gal α F | RET | Exo | 284 |
| <i>Oriza sativa</i> (rice) | β -glucosidase (BGlu1) | GH1 | Glycosynthase | E414A/G/S | β -1,3/4 | D: Glc α F | RET | Exo | 285 |
| <i>Cellulomonas fimi</i> | β -mannosidase | GH2 | Glycosynthase Thioglycoligase | E519A/S E429A | β -1,3/4 | D: ManaF D: Man β -DNP A: 4SMan β -pNP | RET | Exo | 275 279 |
| <i>Escherichia coli</i> | β -galactosidase (LacZ) | GH2 | Glycosynthase | E537S | β -1,6 | D: Gal α F | RET | Exo | 276 |
| <i>Thermotoga maritima</i> | β -glucuronidase | GH2 | Glycosynthase Thioglycoligase | E476A E383A/Q | β -1,3 β -1,3/4 | D: GlcUA α F; GalUA α F D: GlcUA β -pNP; GalUA β -pNP A: 4SGlc β -pNP. 3SGlc β -pNP | RET | Exo | 278 |
| <i>Rhodococcus</i> sp. | endoglycoceramidase (EGCII) | GH5 | Glycosynthase | E351A/G/S | β -1,1 | D: fluorures d' α -glycosyle (DP2, 3 et 5) | RET | Endo | 286 |
| <i>Humiloca insolens</i> | β -1,4-glucanase (cel7B) | GH7 | Glycosynthase | E197A | β -1,3/4 | D: Glc(β -1,4)Glc α F; Gal(β -1,4)Glc α F | RET | Endo | 287, 288 |
| <i>Bacillus halodurans</i> | reducing-end exo-oligoxylanase | GH8 | Glycosynthase | D263N/C | β -1,4 | D: Xyl(β -1,4)Xyl α F | INV | Exo | 272 |
| <i>Cellulomonas fimi</i> | endo- β -1,4-xylanase | GH10 | Glycosynthase | E235G | β -1,4 | D: Xyl(β -1,4)Xyl α F | RET | Endo | 289 |
| <i>Cellulomonas fimi</i> <i>Thermotoga maritima</i> <i>Clostridium stercorarium</i> <i>Bacillus halodurans</i> | endo- β -1,4-xylanases | GH10 | Glycosynthases | E233G E259G E293G E301G | β -1,4 | D: Xyl(β -1,4)Xyl α F | RET | Endo | 290 |
| <i>Bacillus licheniformis</i> | β -1,3-1,4-glucanase (lichenase) | GH16 | Glycosynthase | E314A | β -1,4 | D: Glc(β -1,3)Glc α F | RET | Endo | 274, 291 |
| <i>Pyrococcus furiosus</i> | β -1,3-glucanase (LamA) | GH16 | Glycosynthase | E170A | β -1,3/4 | D: Glc(β -1,3)Glc α F | RET | Endo | 292 |
| <i>Populus tremula x tremuloides</i> | xyloglucan endo-transglycosylase | GH16 | Glycosynthase | E85A | β -1,4 | D : XXXGlc α F ; XLLGlc α F | RET | Endo | 293 |
| <i>Hordeum vulgare</i> | β -1,3-glucanase | GH17 | Glycosynthase | E231G | β -1,3 | D: Glc(β -1,3)Glc α F ; Gal(β -1,3)Glc α F | RET | Endo | 294, 295 |
| <i>Cellvibrio japonicus</i> | β -mannase (Man26A) | GH26 | Glycosynthase | E320G | β -1,4 | D: Man(β -1,4)ManaF | RET | Endo | 296 |
| <i>Shizosaccharomyces pombe</i> | α -glucosidase | GH31 | Glycosynthase | D481G | α -1,4 | D: Glc β F; Glc α -oNP; Xyl α -oNP | RET | Exo | 271 |
| <i>Escherichia coli</i> | α -xylosidase | GH31 | Thioglycoligase | D482A | α -1,4/6 | D: Xyl α F A: 4SGlc β -pNP | RET | Exo | 297 |
| <i>Sulfolobus solfataricus</i> | α -glucosidase | GH31 | Thioglycoligase | D416A | α -1,4 | D: Glc α F A: 4SGlc β -pNP | RET | Exo | 297 |
| <i>Geobacillus stearothermophilus</i> | β -xylosidase | GH52 | Glycosynthase | E335G | β -1,4 | D: Glc α F | RET | Exo | 298 |

En **supprimant la fonction acide/base du 2^{ème} résidu catalytique**, Jahn et coll.²⁷⁹ créent, en 2003, une nouvelle classe de glycosidases : **les thioglycoligases** qui sont capables de catalyser la formation de liaisons thioglycosidiques. La première étape est apparentée au mécanisme de rétention de configuration des GH. Elle consiste en l'attaque du résidu nucléophile/base sur le donneur de glycosyle activé en position anomérique par un bon groupe partant (généralement de type aryl), et conduit à la formation d'un intermédiaire glycosyl-enzyme. Lors de la seconde étape, l'accepteur thiol fortement nucléophile ne requiert pas de catalyse basique pour réagir avec l'intermédiaire ; il générera un thioglycoside. (**FIGURE 27**).

En 2004, Jahn et coll.²⁸⁰ construisent des double-mutants de glycosidases en inactivant la fonction des deux résidus catalytiques. En couplant les fortes réactivités du donneur des glycosynthases (donneur activé de type fluorure de glycosyle) et de l'accepteur des thioglycoligases (nucléophilie de l'accepteur thiol), ces **thioglycosynthases** sont capables de catalyser la condensation de ces deux espèces (**FIGURE 27**).

En l'absence d'un ou de leurs résidus catalytiques, ces nouvelles classes de glycosidases conservent un remarquable pouvoir catalytique. Les valeurs des k_{cat} sont, en effet, de l'ordre de $\sim 200-500 \text{ min}^{-1}$, $\sim 50 \text{ min}^{-1}$ et $\sim 9 \text{ min}^{-1}$ pour les glycosynthases^{273, 299}, glycoligases²⁷⁹ et thioglycosynthases²⁸⁰ provenant de la β -glucosidase d'*Agrobacterium* sp.

C.3.a.2. Leur optimisation : une approche aléatoire

Depuis la découverte des glycosynthases, peu de travaux d'ingénierie rationnelle visant à améliorer les performances de ces enzymes ont abouti. Seuls, Blanchard et coll.²⁸⁸ décrivent en 2006, un changement de régiosélectivité. Guidés par des méthodes de docking moléculaire, ils introduisent une seconde mutation au niveau du sous-site (+1) (H209A/G) de la glycosynthase Cel7B E197A de *Humicola insolens*.

En revanche, les premiers travaux de l'évolution dirigée basés sur des méthodes efficaces de criblage sont remarquables et annoncent des résultats encore plus prometteurs :

- **Criblage *in vivo* sur support solide** : l'équipe de Withers a mis au point un système de criblage³⁰⁰ pour l'identification de glycosynthases plus performantes. Ce crible est basé sur une construction à deux plasmides : l'un contient le gène d'intérêt soumis au préalable à des cycles de mutagenèse aléatoire et l'autre possède le gène codant pour l'endo-cellulase (Cel5A). Réalisé sur support solide, le crible repose sur l'action de l'endo-cellulase, qui clivera uniquement les produits de synthèse des glycosynthases pour libérer un substrat chromogénique. Seuls, les clones

présentant les intensités de fluorescence les plus fortes (~3%) sont sélectionnés et soumis à un criblage en milieu liquide suivant le même principe. Appliquée à l'évolution dirigée de la glycosynthase d'*Agrobacterium* sp. (AbgE358G), cette technique a permis *via* deux cycles de mutagenèse d'obtenir deux mutants améliorés d'un facteur 7 et 27 en terme d'efficacité²⁹⁹. En outre, le meilleur mutant présente de nouvelles spécificités de substrat et élargit ainsi la liste des donneurs (Xyl α F) et accepteurs (Xyl β -pNP, Man β -pNP) des glycosynthases²⁹⁹.

Ce crible reste, néanmoins, difficilement adaptable à des glycosynthases de spécificité différente. Dans ce contexte, le groupe de Cornish propose une alternative à ce crible :

- **Sélection *in vivo* par complémentation chimique** : Lin et coll. ont adapté le système du "triple hybride" à une méthode de sélection des glycosynthases chez la levure *S. cerevisiae*³⁰¹. Dans cette approche, le produit de la réaction enzymatique favorise l'expression du gène rapporteur (*LEU2*, gène requis pour la production de leucine) essentiel à la survie de l'espèce (auxotrophe pour la leucine). L'activation de ce gène est contrôlée par l'activateur de transcription **DBD-AD** composé de deux domaines : le **DBD** ("DNA Binding Domain") et le **AD** ("Activation Domain"). Chacun de ces deux domaines volontairement séparés est fusionné à un tri-complexe : Protéine-Ligand-Disaccharide donneur (pour **DBD**) ou accepteur (pour **AD**). Ces deux domaines ne seront réunis que sous l'action d'une glycosynthase active et pourront alors activer la production de leucine (**FIGURE 28**). Cette approche a été validée sur une banque de monovariants de la glycosynthase Cel 7B E197A (obtenus par mutagenèse à saturation sur la position 197). Elle a notamment permis d'isoler le variant E197S doté d'une efficacité catalytique accrue (facteur 5 par rapport au E197A). Depuis, cette stratégie a été optimisée pour diversifier les substrats dimériques³⁰² et être applicable aux glycosynthases de spécificité différente.

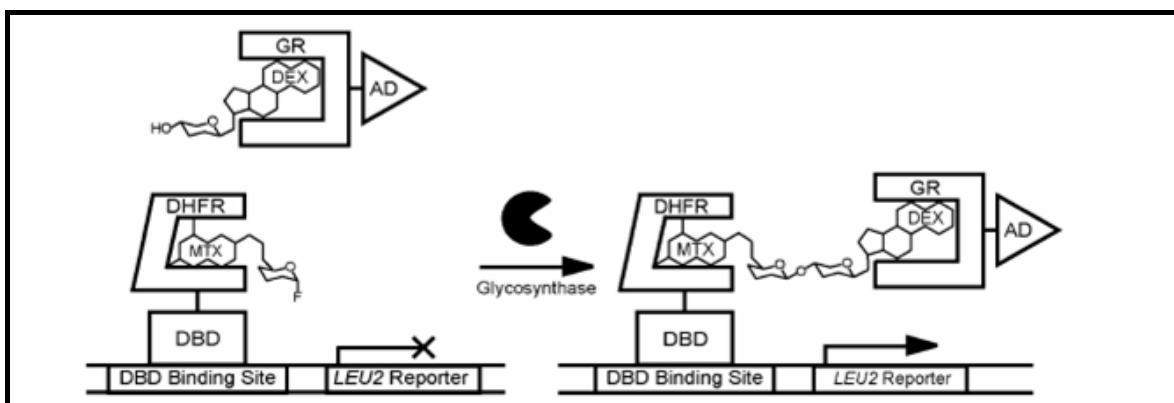


FIGURE 28 | Sélection *in vivo* par complémentation chimique.
(D'après Lin et coll.³⁰¹)

C.3.b. Des approches alternatives aux glycosynthèses

Les glycosynthèses constituent un nouvel outil de transglycosylation qui, cependant, ne s'utilise que très rarement en présence de substrats naturels. D'autres exemples d'ingénierie de glycosidases lèvent cette contrainte pour générer :

- **des β -transglycosydases plus performantes :**

En 2005, Feng et coll. améliorent le taux de transglycosylation (**T/H**= activité transglycosidase/activité hydrolase) d'une β -glycosidase de *Thermus thermophilus* via un cycle de mutagenèse aléatoire³⁰³. Leur stratégie de criblage sur support solide consiste à tester la banque sur un substrat devenant chromophore (X-Gal) s'il est clivé. Les colonies bleu clair (clones actifs ne présentant qu'une très faible activité hydrolase) sont supposées présenter un meilleur **T/H** et sont isolées. Cette équipe applique ce même principe à la conversion d'une α -fucosidase de *Thermotoga maritima* en α -transfucosidase³⁰⁴. Après un cycle de mutagenèse aléatoire et recombinaison des mutations bénéfiques, ils construisent un triple mutant présentant un **T/H** amélioré d'un facteur 32. Deux des 3 résidus mutés sont situés sur la seconde sphère entourant le site actif. Des travaux de modélisation moléculaire suggèrent que ces mutations induisent un réarrangement du site actif favorable à la fixation du substrat.

Zhang et coll. convertissent une β -galactosidase d'*Escherichia coli* en β -fucosidase en utilisant la technique du DNA shuffling (7 cycles). Le meilleur mutant présente un gain d' "activité β -fucosidase " de 10^{305} . Avec les mêmes objectifs, le groupe de Matsumura utilise une méthode semi-rationnelle sur cette enzyme en ciblant uniquement 3 positions (par mutagenèse à saturation). Au préalable, les 3 résidus correspondants, situés au niveau du site actif ont été identifiés comme cibles par docking moléculaire. Ils décrivent alors un mutant présentant un gain largement supérieur (180)³⁰⁶ et démontrent l'efficacité d'une méthode semi-rationnelle. Cette équipe a appliqué avec succès une stratégie similaire³⁰⁷ pour accroître l' "activité β -galactosidase " d'une β -glucuronidase d'*Escherichia coli*.

En mutant individuellement deux résidus impliqués dans la reconnaissance du substrat, Hancock et coll. adopte une stratégie rationnelle pour élargir la spécificité de synthèse³⁰⁸ d'une β -glycosidase de *Sulfolobus solfataricus*. Ils génèrent des glycosidases capables de former des osides de synthèse peu courants : β -xylosides et β -mannosides.

- des α -transglycosydases plus performantes :

De nombreux exemples existent dans la littérature d'ingénierie d' α -glycoside-hydrolases par approche rationnelle ou combinatoire visant à augmenter leur efficacité catalytique^{309, 310}, leur stabilité au pH³¹¹, à la température^{309, 312} et leur activité transglucosidase^{304, 313}.

Concernant l'amylosaccharase, parallèlement aux travaux d'ingénierie rationnelle décrits précédemment (cf. partie III.B.2.d.2., page 70), l'évolution dirigée de l'ASNP a abouti à la génération de mutants dont l'efficacité catalytique a été améliorée^{314, 315} et, plus récemment, à celle de variants thermostables³¹⁶. Emond et coll. décrivent ainsi un double mutant fonctionnel à 50°C³¹⁷ et de ce fait, plus adapté aux contraintes industrielles.

Pour les glucane-saccharases, différentes études se sont intéressés au facteur pouvant contrôler la régiospécificité de ces enzymes et se sont focalisés sur les résidus à proximité de la triade catalytique (FIGURE 29).

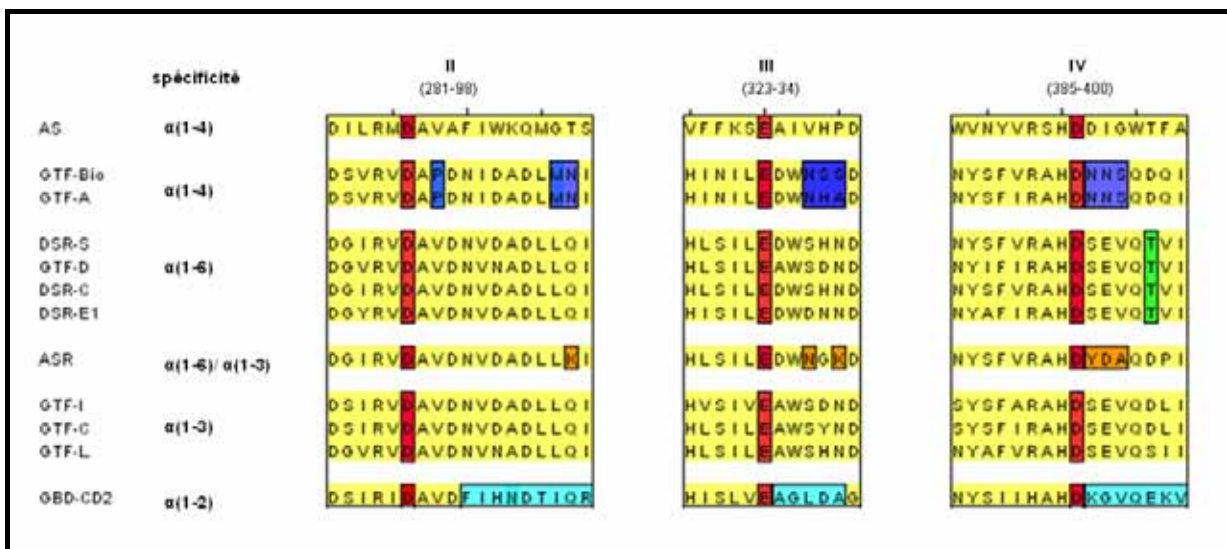


FIGURE 29| Alignement de séquences conservées des domaines catalytiques de glucane-saccharases de spécificités variées.

En rouge : acides aminés catalytiques. **En bleu foncé** : résidus supposés impliqués dans la liaison $\alpha(1-4)$. **En vert** : résidus supposés impliqués dans la liaison $\alpha(1-6)$. **En orange** : résidus responsables de la teneur en liaisons $\alpha-1,3/ \alpha-1,6$ chez les mutane- et dextrane-saccharases. **En bleu cyan** : résidus supposés impliqués dans la liaison $\alpha(1-2)$.

GH13: AS de *N. polysaccharea*

GH70: GTF-Bio de *Lb. reuteri* ATCC-55730. GTF-A de *Lb. reuteri* 121; DSR-S de *Ln. mesenteroides* NRRL B-512F; GTF-D de *S. mutans* ; DSR-C de *Ln. mesenteroides* NRRL B-1355; DSR-C, *Ln. mesenteroides* NRRL B-1355; DSR-E1, premier domaine catalytique de *Ln. mesenteroides* NRRL B-1299 ; ASR, alternane-saccharase de *Ln. mesenteroides* NRRL B-1355 GTF-I, *S. downei* ; GTF-C de *S. mutans* ; GTF-L de *S. salivarius* ATCC 25975; GBD-CD2 , deuxième domaine catalytique de *Ln. mesenteroides* NRRL B-1299.

La numérotation des acides aminés est donnée en référence à l'ASNP.

Notamment, en aval du 3^{ème} résidu catalytique, on peut observer des résidus ou zones qui semblent caractéristiques d'une spécificité donnée, en particulier :

- **de la teneur en liaison α -1,3/ α -1,6**: les enzymes qui catalysent la synthèse de polymère formé de liaisons osidiques α -1,6 (polymère soluble) possèdent une thréonine (**T**) à proximité du 3^{ème} résidu catalytique (acide aminé en vert, **FIGURE 29**) alors que celles produisant un glucane majoritairement composé de liaisons osidiques α -1,3 (polymère insoluble) présentent un acide aspartique (**D**) à cette même position. Ainsi, le changement de la thréonine de GTF-D par l'acide aspartique permet d'accroître de 85 % la synthèse de liaisons α -1,3¹⁵⁶. Ce type d'expérience de mutagenèse dirigée a été appliqué à d'autres enzymes³¹⁸. Par ailleurs, la mutation D569T sur la GTF-I de *S. downei* MFe28 conduit bien à la production d'une fraction de polymère soluble³¹⁹.
- **de la teneur en liaison α -1,4/ α -1,6**: Krajl et coll. ont montré qu'en modifiant le peptide NNS en SEV (acides aminés en bleus foncé, **FIGURE 29**), le variant de la GTF-A obtenu synthétisait un glucane composé à plus de 80 % de liaisons α -1,6 au détriment des liaisons α -1,4³²⁰. Les auteurs ont également muté les positions P1026V et I1029V en addition aux premières mutations NNS1134SEV, par analogie avec la séquence retrouvée dans la DSR-S. Le mutant obtenu synthétise alors un polymère composé à plus de 85 % de liaisons α -1,6, ce qui conduit les auteurs à supposer que la spécificité de liaisons de la GTF-A peut être contrôlée en jouant sur la nature des acides aminés présents en aval des deux acides aspartiques de la triade. Enfin, ils se sont plus particulièrement intéressés à la première asparagine (N1134), montrant ainsi son implication directe dans l'activité transglucosidase de l'enzyme³²¹. Toutefois, les paramètres cinétiques (activité initiale et K_M) de ces mutants sont largement affectés.
- **de la teneur en liaison α -1,2/ α -1,6** : Fabre et coll. ont également étudié par mutagenèse dirigée le rôle des résidus divergents présents dans le second domaine catalytique de DSR-E, responsable de la formation des liaisons α -1,2¹⁶⁸. Les zones présentes en aval des deux acides aspartiques de la triade (acides aminés en bleus cyan, **FIGURE 29**), fortement divergentes des motifs consensus, ont été remplacées dans le GBD-CD2 par les acides aminés présents dans le premier domaine catalytique, responsable de la synthèse des liaisons α -1,6. Or, si la mutation d'un seul peptide entraîne une perte complète de l'activité du catalyseur, le double mutant restaure une très faible activité détectable sur zymogramme alors qu'aucun produit n'est décelé après migration du GBD-CD2. Les auteurs suggèrent ainsi que la spécificité de liaison a été modifiée, et que le double mutant synthétise des traces d'oligosaccharides très probablement liés en α -1,6³²².

Il apparaît donc possible, par mutagenèse rationnelle, de favoriser la synthèse d'un lien osidique qui n'est produit qu'en minorité par la GS sauvage. Cependant, aucun **changement total de spécificité n'a encore été décrit.**

Finalement, nous constaterons que, de manière tout à fait surprenante, **très peu de travaux ont visé la modification de la spécificité de reconnaissance** de ces enzymes pour élargir leur aptitude à la glycosylation d'accepteurs non naturels. C'est le défi que nous entendons relever dans nos travaux et dans la stratégie que nous allons exposer plus en détail avant de nous consacrer aux résultats.

POSITIONNEMENT DU SUJET ET STRATEGIES

A. Stratégie de synthèse des haptènes 3a et 1b

A l'exception du sérotype 6, les antigènes O de *S. flexneri* sont tous construits à partir d'une structure commune : un tétrasaccharide linéaire [ABCD], composé de trois résidus rhamnoside, liés en 2 ou 3 et d'une unité D-glucoside N-acétylé liée en position 2. Cette structure non modifiée correspond à l'antigène natif de *S. flexneri*, donnant lieu au sérotype Y. La conversion antigénique en un autre sérotype résulte d'une addition de résidus glucosyl et/ou O-acétyl sur l'un des différents sucres de la structure tétrasaccharidique de base (FIGURE 12). Dans les 2 motifs choisis (3a et 1b), le tétrasaccharide est ramifié par un groupe O-acétyl au niveau du saccharide [C]. Ces deux motifs se distinguent uniquement par le greffage d'un groupement **glucosyle** [E] au niveau du saccharide [A] (pour 3a) ou [D] (pour 1b) (FIGURE 30).

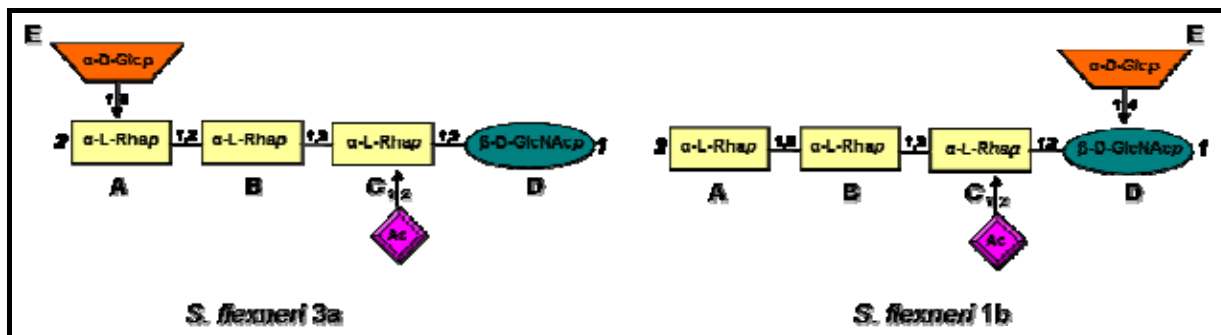


FIGURE 30| Unités répétitives de *S. flexneri* 3a et 1b.

La voie de synthèse proposée par les chimistes de l'Institut Pasteur pour ces 2 haptènes consiste à :

- i) transférer tout d'abord le groupement **glucosyle** [E] par voie enzymatique sur [A] (ou [D]) par formation d'une liaison osidique de type α -1,3 (ou α -1,4). Le développement de cette 1^{ère} étape fait l'objet de ce sujet de thèse.
- ii) greffer, par la suite, le motif [BCD] (ou [CBA]) sur le disaccharide [EA] (ou [ED]) par voie chimique. Cette 2^{ème} étape est réalisée au sein de l'Unité de Chimie des Biomolécules de l'Institut Pasteur dirigée par L. Mulard.

B. Stratégie de synthèse des disaccharides par voie enzymatique

B.1. Choix des couples glucane-saccharase/accepteur

B.1.a. Choix des glucane-saccharases

Dans notre étude, nous avons choisi de tester l'amylosaccharase de *N. polysaccharea* et trois glucane-saccharases de la famille GH70 du genre *L. mesenteroides*. Ces GS sont disponibles sous leur forme recombinante au Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP). Concernant les GS de la famille GH70 sélectionnées, chacune catalyse la synthèse de glucanes différents et originaux:

- la **dextrane-saccharase DSR-E** de *L. mesenteroides* NRRL B-1299 qui produit un dextrane ramifié à 28% par des liaisons α -1,2. Cette enzyme a fait l'objet d'une attention particulière depuis le clonage de son gène par Bozonnet et coll. en 2002. Il s'agit en effet à ce jour de la seule glucane-saccharase de structure primaire connue possédant deux domaines catalytiques¹⁴³.
- l'**alternane-saccharase ASR** de *L. mesenteroides* NRRL B-1355 qui synthétise un glucane formé d'un enchaînement de liaisons α -1,6 et α -1,3 alternées, et présente quelques ramifications liées en α -1,6 ou α -1,3.
- la **dextrane-saccharase DSR-S** de *L. mesenteroides* NRRL B-512F qui synthétise des dextrans composés à plus de **95 % de liaisons α -1,6**.

Ces 3 enzymes, utilisées d'abord dans leur forme entière, ont fait l'objet de travaux d'ingénierie qui ont abouti à la construction de formes tronquées plus spécifiques et/ou performantes :

- **GBD-CD2, forme tronquée de DSR-E** : par ingénierie rationnelle, Fabre et coll. ont montré que seul le deuxième domaine (CD2) catalytique était responsable de la synthèse des **liaisons α -1,2**, et ont construit une nouvelle transglucosidase (GBD-CD2) spécifique de la synthèse de ces liaisons rares dans la nature. Ce variant est incapable de synthétiser des polymères de glucose directement à partir de saccharose, mais réalise des branchements en α -1,2 par réactions d'accepteur sur des dextrans cliniques. Il s'agit donc d'une enzyme de branchement α -1,2 spécifique
- **ASR C-APY del, forme tronquée de l'ASR** : en délétant du gène *asr* un fragment codant pour une partie du domaine C-terminal, Joucla et coll. ont optimisé l'expression de cette enzyme en terme de production (facteur 120) et de pureté (disparition des formes dégradées observées avec la protéine entière)^{323,324}. En outre, cette délétion ne modifie pas le profil de produits obtenus avec l'ASR entière.
- **DSR-S vardel Δ 4N, forme tronquée de DSR-S** en parties N et C-Terminales. Tout comme pour l'ASR C-APY del, ce variant permet d'obtenir une enzyme hautement

purifiable à des taux de production élevées^{325, 326} et présente des profils de synthèse similaires à ceux de la forme entière. Signalons aussi, que d'autres variants, dépourvus d'activité polymérase ont été construits à partir de cette forme tronquée en mutant certains résidus situés juste en aval du troisième acide aspartique de la triade catalytique^{325, 326}.

En plus de l'ASNP, ce sont donc ces trois formes optimisées qui ont été utilisées lors ces travaux de thèse, permettant ainsi de disposer de GS possédant les 4 spécificités de liaison : α -1,2 ; α -1,3 ; α -1,4 et α -1,6. Le TABLEAU 7 compare les principales caractéristiques de production et biochimiques de ces 4 GS.

TABLEAU 7 | Comparaison des principales caractéristiques des 4GS.

| Enzymes recombinantes | GBD-CD2 | ASR-C-del-APY | DSR-S Δ 4N | ASNP |
|----------------------------|---------|---------------|-------------------|------|
| Niveau de production (U/L) | 500 | 5000 | 5600 | 300 |
| Activité spécifique (U/mg) | 24 | 160 | 424 | 4,9 |
| k_{cat} (s^{-1}) | 78 | 404 | 584 | 1 |

B.1.b Choix des accepteurs

Afin d'élaborer au mieux la stratégie de synthèse, les contraintes imposées par les voies chimique et enzymatique ont été prises en compte : la voie enzymatique doit permettre de fournir le disaccharide [EA] (ou [ED]) dans une forme compatible avec la suite de la synthèse. De ce fait, le résidu [A] (ou [D]) doit être au préalable protégé sur certaines de ses fonctions hydroxyle avant de subir la réaction enzymatique. En outre, la topologie du site actif des enzymes implique certaines restrictions au niveau du choix des groupes protecteurs. Notamment, des groupes volumineux sont à proscrire sur certaines positions du résidu de départ. Il convient donc d'ajuster au mieux cette double contrainte.

Dans un premier temps, les accepteurs [A] et [D] non protégés ont été utilisés pour évaluer la faisabilité de la réaction enzymatique. Dans un second temps, des dérivés de ces résidus modifiés exclusivement au niveau de la position anomérique ont été jugés envisageables. Finalement, les accepteurs [A'] et [D'] retenus pour la synthèse chimio-enzymatique des haptènes 3a et 1b sont respectivement le α -methyl L-rhamnose et l' α -allyl-N-Acétyle-D-glucosamine (FIGURE 31).

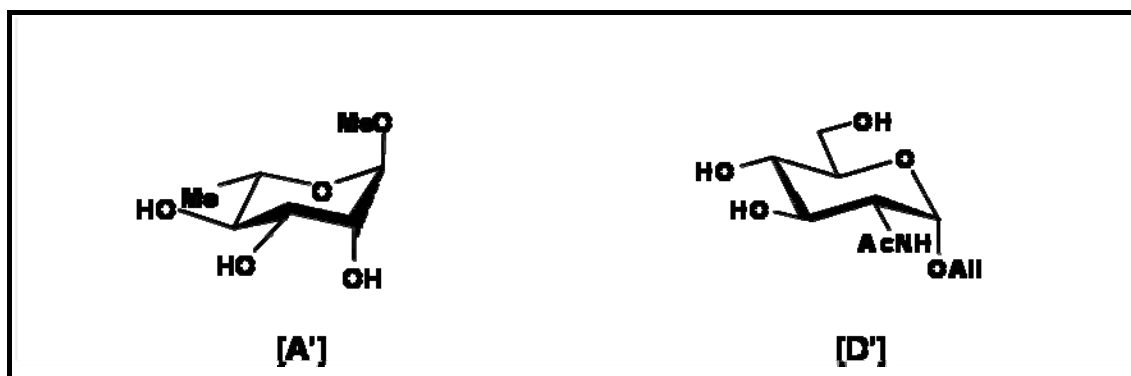


FIGURE 31| Accepteurs retenus ([A'] et [D'])
pour la synthèse chimio-enzymatique des haptènes de *S. flexneri* (3a et 1b)

B.2. Stratégie employée pour le développement de la voie enzymatique

Pour accéder aux disaccharides cibles, le développement de la voie enzymatique a été réalisé en deux phases :

- **Choix des couples glucane-saccharase/accepteur optimaux.** Le potentiel des quatre GS sauvages à glucosyler les deux types d'accepteurs [A] et [D] a été évalué. Les produits de glucosylation les plus abondants ont été caractérisés. Ces travaux, décrits au **chapitre II** de ce manuscrit, ont permis de définir les couples enzyme/accepteurs les plus prometteurs.
- **Ingénierie de l'amylosaccharase de *N. polysaccharea*.** Nous avons retenu finalement l'ASNP, comme enzyme candidate à l'optimisation de la glucosylation des deux types d'accepteurs. Une conception semi-rationnelle a alors été adoptée pour l'évolution de cette enzyme. Nous présenterons les résultats de ces travaux dans les **chapitres III et IV** de cette thèse.

Références bibliographiques

1. Verez-Bencomo, V.; Fernandez-Santana, V.; Hardy, E.; Toledo, M. E.; Rodriguez, M. C.; Heynngnezz, L.; Rodriguez, A.; Baly, A.; Herrera, L.; Izquierdo, M.; Villar, A.; Valdes, Y.; Cosme, K.; Deler, M. L.; Montane, M.; Garcia, E.; Ramos, A.; Aguilar, A.; Medina, E.; Torano, G.; Sosa, I.; Hernandez, I.; Martinez, R.; Muzachio, A.; Carmenates, A.; Costa, L.; Cardoso, F.; Campa, C.; Diaz, M.; Roy, R., A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b. *Science*, **2004**, 305, (5683), 522-5.
2. Phalipon, A.; Costachel, C.; Grandjean, C.; Thuizat, A.; Guerreiro, C.; Tanguy, M.; Nato, F.; Vulliez-Le Normand, B.; Belot, F.; Wright, K.; Marcel-Peyre, V.; Sansonetti, P. J.; Mulard, L. A., Characterization of functional oligosaccharide mimics of the *Shigella flexneri* serotype 2a O-antigen: implications for the development of a chemically defined glycoconjugate vaccine. *J Immunol*, **2006**, 176, (3), 1686-94.
3. Eloit, M., Vaccins traditionnels et recombinants. *INRA Prod. Anim.*, **1998**, 11, (1), 5-13.
4. Scarselli, M.; Giuliani, M. M.; Adu-Bobie, J.; Pizza, M.; Rappuoli, R., The impact of genomics on vaccine design. *Trends Biotechnol*, **2005**, 23, (2), 84-91.
5. Rappuoli, R., Reverse vaccinology. *Curr Opin Microbiol*, **2000**, 3, (5), 445-50.
6. Rappuoli, R., From Pasteur to genomics: progress and challenges in infectious diseases. *Nat Med*, **2004**, 10, (11), 1177-85.
7. WHO, Le point sur les vaccins et la vaccination dans le monde. **2002**.
8. Mulard, L.; Alonso, J. M.; Fournier, J. M., Polysaccharide vaccines *Annales de l'Institut Pasteur*, **2002**, 12, 37-54.
9. Ada, G.; Isaacs, D., Carbohydrate-protein conjugate vaccines. *Clin Microbiol Infect*, **2003**, 9, (2), 79-85.
10. Mulard, L., [Carbohydrates and vaccines: from purified polysaccharides to semi-synthetic glycoconjugate vaccines]. *Ann Pharm Fr*, **2007**, 65, (1), 14-32.
11. Vliegthart, J. F., Carbohydrate based vaccines. *FEBS Lett*, **2006**, 580, (12), 2945-50.
12. Seeberger, P. H.; Werz, D. B., Synthesis and medical applications of oligosaccharides. *Nature*, **2007**, 446, (7139), 1046-51.
13. Roy, R., New trends in carbohydrate-based vaccines. *Drug Discovery Today: Technologies* **2004**, 1, (3), 237-336.
14. Kamerling, J., Pneumococcal polysaccharides: a chemical view. *Tomasz (Ed.), Streptococcus pneumoniae, Larchmont: Mary Ann Liebert*, **2000**, 81-114.
15. Lindberg, B., Components of bacterial polysaccharides. *Adv Carbohydr Chem Biochem*, **1990**, 48, 279-318.
16. Wessels, M. R.; Paoletti, L. C.; Rodewald, A. K.; Michon, F.; DiFabio, J.; Jennings, H. J.; Kasper, D. L., Stimulation of protective antibodies against type Ia and Ib group B streptococci by a type Ia polysaccharide-tetanus toxoid conjugate vaccine. *Infect Immun*, **1993**, 61, (11), 4760-6.
17. Jones, C., Vaccines based on the cell surface carbohydrates of pathogenic bacteria. *Anais da Academia Brasileira de Ciências*, **2005**, 77, (2), 293-324.
18. Edstrom, R. D.; Heath, E. C., Isolation of colitose-containing oligosaccharides from the cell wall lipopolysaccharide of *Escherichia coli*. *Biochem Biophys Res Commun*, **1965**, 21, (6), 638-43.
19. Eklund, S. H.; Garegg, P. J.; Kenne, L.; Lindberg, A. A.; Lindberg, B., Structural studies on the *Escherichia coli* O111 lipopolysaccharide. *Abstracts of the IXth International Symposium of Carbohydrate Chemistry, London*, **1978**, Abstr 493.

20. Perry, M. B.; MacLean, L.; Griffith, D. W., Structure of the O-chain polysaccharide of the phenol-phase soluble lipopolysaccharide of *Escherichia coli* 0:157:H7. *Biochem Cell Biol*, **1986**, 64, (1), 21-8.
21. Moxon, E. R.; J. S. Kroll, J. S., The role of bacterial polysaccharide capsules as virulence factors. *Curr. Top. Microbiol. Immunol.*, **1990**, 150, 65-85.
22. Wessels, M. R.; Goldberg, J. B.; Moses, A. E.; DiCesare, T. J., Effects on virulence of mutations in a locus essential for hyaluronic acid capsule expression in group A streptococci. *Infect Immun*, **1994**, 62, (2), 433-41.
23. Rubens, C. E.; Wessels, M. R.; Heggen, L. M.; Kasper, D. L., Transposon mutagenesis of type III group B *Streptococcus*: correlation of capsule expression with virulence. *Proc Natl Acad Sci U S A*, **1987**, 84, (20), 7208-12.
24. Kim, K. S.; Itabashi, H.; Gemski, P.; Sadoff, J.; Warren, R. L.; Cross, A. S., The K1 capsule is the critical determinant in the development of *Escherichia coli* meningitis in the rat. *J Clin Invest*, **1992**, 90, (3), 897-905.
25. Zwahlen, A.; Kroll, J. S.; Rubin, L. G.; Moxon, E. R., The molecular basis of pathogenicity in *Haemophilus influenzae*: comparative virulence of genetically-related capsular transformants and correlation with changes at the capsulation locus cap. *Microb. Pathog*, **1989**, 7, 225-235.
26. West, N. P.; Sansonetti, P.; Mounier, J.; Exley, R. M.; Parsot, C.; Guadagnini, S.; Prevost, M. C.; Prochnicka-Chalufour, A.; Delepierre, M.; Tanguy, M.; Tang, C. M., Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science*, **2005**, 307, (5713), 1313-7.
27. Tanizaki, M. M.; Garcia, L. R.; Ramos, J. B.; Leite, L. C. C.; Hiss, H.; Furuta, J. A.; Cabrera-Crespo, J.; Raw, I., Purification of meningococcal group C polysaccharide by a procedure suitable for scale-up *Journal of Microbiological Methods*, **1996**, 27, (5), 19-23.
28. Pato, T. P.; Barbosa Ade, P.; da Silva Junior, J. G., Purification of capsular polysaccharide from *Neisseria meningitidis* serogroup C by liquid chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci*, **2006**, 832, (2), 262-7.
29. Mond, J. J.; Lees, A.; Snapper, C. M., T cell-independent antigens type 2. *Annu Rev Immunol*, **1995**, 13, 655-92.
30. Girard, M. P.; Preziosi, M. P.; Aguado, M. T.; Kieny, M. P., A review of vaccine research and development: meningococcal disease. *Vaccine*, **2006**, 24, (22), 4692-700.
31. Cadoz, M.; Armand, J.; Arminjon, F.; Gire, R.; Lafaix, C., Tetravalent (A, C, Y, W 135) meningococcal vaccine in children: immunogenicity and safety. *Vaccine*, **1985**, 3, (3), 340-2.
32. Avery, O. T.; Goebel, W. F., Chemo-Immunological Studies on Conjugated Carbohydrate-Proteins: V. The Immunological Specificity of an Antigen Prepared by Combining the Capsular Polysaccharide of Type III Pneumococcus with Foreign Protein *Journal of Experimental Medicine* **1931**, 54, 437-447.
33. Ada, G., Vaccines and vaccination. *N Engl J Med*, **2001**, 345, (14), 1042-53.
34. Shelly, M. A.; Jacoby, H.; Riley, G. J.; Graves, B. T.; Pichichero, M.; Treanor, J. J., Comparison of pneumococcal polysaccharide and CRM197-conjugated pneumococcal oligosaccharide vaccines in young and elderly adults. *Infect Immun*, **1997**, 65, (1), 242-7.
35. Vella, P. P.; Marburg, S.; Staub, J. M.; Kniskern, P. J.; Miller, W.; Hagopian, A.; Ip, C.; Tolman, R. L.; Rusk, C. M.; Chupak, L. S.; et al., Immunogenicity of conjugate vaccines consisting of pneumococcal capsular polysaccharide types 6B, 14, 19F, and 23F and a meningococcal outer membrane protein complex. *Infect Immun*, **1992**, 60, (12), 4977-83.
36. Dagan, R.; Eskola, J.; Leclerc, C.; Leroy, O., Reduced response to multiple vaccines sharing common protein epitopes that are administered simultaneously to infants. *Infect Immun*, **1998**, 66, (5), 2093-8.

37. Fattom, A.; Cho, Y. H.; Chu, C.; Fuller, S.; Fries, L.; Naso, R., Epitopic overload at the site of injection may result in suppression of the immune response to combined capsular polysaccharide conjugate vaccines. *Vaccine*, **1999**, 17, (2), 126-33.
38. Daum, R. S.; Zenko, C. E.; Given, G. Z.; Ballanco, G. A.; Parikh, H.; Germino, K., Magnitude of interference after diphtheria-tetanus toxoids-acellular pertussis/*Haemophilus influenzae* type b capsular polysaccharide-tetanus vaccination is related to the number of doses administered. *J Infect Dis*, **2001**, 184, (10), 1293-9.
39. Lin, F. Y.; Ho, V. A.; Khiem, H. B.; Trach, D. D.; Bay, P. V.; Thanh, T. C.; Kossaczka, Z.; Bryla, D. A.; Shiloach, J.; Robbins, J. B.; Schneerson, R.; Szu, S. C., The efficacy of a *Salmonella typhi* Vi conjugate vaccine in two-to-five-year-old children. *N Engl J Med*, **2001**, 344, (17), 1263-9.
40. Cohen, D.; Ashkenazi, S.; Green, M. S.; Gdalevich, M.; Robin, G.; Slepon, R.; Yavzori, M.; Orr, N.; Block, C.; Ashkenazi, I.; Shemer, J.; Taylor, D. N.; Hale, T. L.; Sadoff, J. C.; Pavliakova, D.; Schneerson, R.; Robbins, J. B., Double-blind vaccine-controlled randomised efficacy trial of an investigational *Shigella sonnei* conjugate vaccine in young adults. *Lancet*, **1997**, 349, (9046), 155-9.
41. Fattom, A. I.; Horwith, G.; Fuller, S.; Propst, M.; Naso, R., Development of StaphVAX, a polysaccharide conjugate vaccine against *S. aureus* infection: from the lab bench to phase III clinical trials. *Vaccine*, **2004**, 22, (7), 880-7.
42. Barington, T.; Gyhrs, A.; Kristensen, K.; Heilmann, C., Opposite effects of actively and passively acquired immunity to the carrier on responses of human infants to a *Haemophilus influenzae* type b conjugate vaccine. *Infect Immun*, **1994**, 62, (1), 9-14.
43. Falugi, F.; Petracca, R.; Mariani, M.; Luzzi, E.; Mancianti, S.; Carinci, V.; Melli, M. L.; Finco, O.; Wack, A.; Di Tommaso, A.; De Magistris, M. T.; Costantino, P.; Del Giudice, G.; Abrignani, S.; Rappuoli, R.; Grandi, G., Rationally designed strings of promiscuous CD4(+) T cell epitopes provide help to *Haemophilus influenzae* type b oligosaccharide: a model for new conjugate vaccines. *Eur J Immunol*, **2001**, 31, (12), 3816-24.
44. Baraldo, K.; Mori, E.; Bartoloni, A.; Petracca, R.; Giannozzi, A.; Norelli, F.; Rappuoli, R.; Grandi, G.; Del Giudice, G., N19 polyepitope as a carrier for enhanced immunogenicity and protective efficacy of meningococcal conjugate vaccines. *Infect Immun*, **2004**, 72, (8), 4884-7.
45. Baraldo, K.; Mori, E.; Bartoloni, A.; Norelli, F.; Grandi, G.; Rappuoli, R.; Finco, O.; Del Giudice, G., Combined conjugate vaccines: enhanced immunogenicity with the N19 polyepitope as a carrier protein. *Infect Immun*, **2005**, 73, (9), 5835-41.
46. Alexander, J.; Sidney, J.; Southwood, S.; Ruppert, J.; Oseroff, C.; Maewal, A.; Snoke, K.; Serra, H. M.; Kubo, R. T.; Sette, A.; et al., Development of high potency universal DR-restricted helper epitopes by modification of high affinity DR-blocking peptides. *Immunity*, **1994**, 1, (9), 751-61.
47. Wang, J. Y.; Chang, A. H.; Guttormsen, H. K.; Rosas, A. L.; Kasper, D. L., Construction of designer glycoconjugate vaccines with size-specific oligosaccharide antigens and site-controlled coupling. *Vaccine*, **2003**, 21, (11-12), 1112-7.
48. Anderson, P. W.; Pichichero, M. E.; Stein, E. C.; Porcelli, S.; Betts, R. F.; Connuck, D. M.; Korones, D.; Insel, R. A.; Zahradnik, J. M.; Eby, R., Effect of oligosaccharide chain length, exposed terminal group, and hapten loading on the antibody response of human adults and infants to vaccines consisting of *Haemophilus influenzae* type b capsular antigen unterminally coupled to the diphtheria protein CRM197. *J Immunol*, **1989**, 142, (7), 2464-8.
49. Paoletti, L. C.; Kasper, D. L.; Michon, F.; DiFabio, J.; Jennings, H. J.; Tosteson, T. D.; Wessels, M. R., Effects of chain length on the immunogenicity in rabbits of group B *Streptococcus* type III oligosaccharide-tetanus toxoid conjugates. *J Clin Invest*, **1992**, 89, (1), 203-9.

50. Pozsgay, V.; Chu, C.; Pannell, L.; Wolfe, J.; Robbins, J. B.; Schneerson, R., Protein conjugates of synthetic saccharides elicit higher levels of serum IgG lipopolysaccharide antibodies in mice than do those of the O-specific polysaccharide from *Shigella dysenteriae* type 1. *Proc Natl Acad Sci U S A*, **1999**, 96, (9), 5194-7.
51. Donnelly, J. J.; Deck, R. R.; Liu, M. A., Immunogenicity of a *Haemophilus influenzae* polysaccharide-*Neisseria meningitidis* outer membrane protein complex conjugate vaccine. *J Immunol*, **1990**, 145, (9), 3071-9.
52. Gupta, R. K.; Egan, W.; Bryla, D. A.; Robbins, J. B.; Szu, S. C., Comparative immunogenicity of conjugates composed of *Escherichia coli* O111 O-specific polysaccharide, prepared by treatment with acetic acid or hydrazine, bound to tetanus toxoid by two synthetic schemes. *Infect Immun*, **1995**, 63, (8), 2805-10.
53. Konadu, E.; Shiloach, J.; Bryla, D. A.; Robbins, J. B.; Szu, S. C., Synthesis, characterization, and immunological properties in mice of conjugates composed of detoxified lipopolysaccharide of *Salmonella paratyphi* A bound to tetanus toxoid with emphasis on the role of O acetyls. *Infect Immun*, **1996**, 64, (7), 2709-15.
54. Fries, L. F.; Montemarano, A. D.; Mallett, C. P.; Taylor, D. N.; Hale, T. L.; Lowell, G. H., Safety and immunogenicity of a proteosome-*Shigella flexneri* 2a lipopolysaccharide vaccine administered intranasally to healthy adults. *Infect Immun*, **2001**, 69, (7), 4545-53.
55. Wang, Z. G.; Zhang, X.; Visser, M.; Live, D.; Zatorski, A.; Iserloh, U.; Lloyd, K. O.; Danishefsky, S. J., Toward Fully Synthetic Homogeneous Glycoproteins: A High Mannose Core Containing Glycopeptide Carrying Full H-Type 2 Human Blood Group Specificity *Angew Chem Int Ed Engl*, **2001**, 40, (9), 1728-1732.
56. Seifert, J.; Lergenmuller, M.; Ito, Y., Synthesis of an alpha-(2,3)-Sialylated, Complex-Type Undecasaccharide. *Angew Chem Int Ed Engl*, **2000**, 39, (3), 531-534.
57. Plante, O. J.; Palmacci, E. R.; Seeberger, P. H., Automated solid-phase synthesis of oligosaccharides. *Science*, **2001**, 291, (5508), 1523-7.
58. Seeberger, P. H.; Werz, D. B., Automated synthesis of oligosaccharides as a basis for drug discovery. *Nat Rev Drug Discov*, **2005**, 4, (9), 751-63.
59. Wolfson, W., Ancora cooks with carbs synthesizing carbohydrate vaccines. *Chem Biol*, **2006**, 13, (7), 689-91.
60. Love, K. R.; Seeberger, P. H., Automated solid-phase synthesis of protected tumor-associated antigen and blood group determinant oligosaccharides. *Angew. Chem. Int. Ed.*, **2004**, 43, 602-605.
61. Zhang, Z.; Ollmann, I. R.; Ye, X. S.; Wischnat, R.; Baasov, T.; Wong, C. H., Programmable One-Pot Oligosaccharide Synthesis. *J. Am. Chem. Soc.*, **1999**, 121, (4), 734-53.
62. Koeller, K. M.; Wong, C. H., Synthesis of Complex Carbohydrates and Glycoconjugates: Enzyme-Based and Programmable One-Pot Strategies *Chem. Rev.*, **2000**, 100, (12), 4465-4494.
63. Benaissa-Trouw, B.; Lefeber, D. J.; Kamerling, J. P.; Vliegthart, J. F.; Kraaijeveld, K.; Snippe, H., Synthetic polysaccharide type 3-related di-, tri-, and tetrasaccharide-CRM(197) conjugates induce protection against *Streptococcus pneumoniae* type 3 in mice. *Infect Immun*, **2001**, 69, (7), 4698-701.
64. Jansen, W. T.; Hogenboom, S.; Thijssen, M. J.; Kamerling, J. P.; Vliegthart, J. F.; Verhoef, J.; Snippe, H.; Verheul, A. F., Synthetic 6B di-, tri-, and tetrasaccharide-protein conjugates contain pneumococcal type 6A and 6B common and 6B-specific epitopes that elicit protective antibodies in mice. *Infect Immun*, **2001**, 69, (2), 787-93.
65. Chong, P.; Chan, N.; Kandil, A.; Tripet, B.; James, O.; Yang, Y. P.; Shi, S. P.; Klein, M., A strategy for rational design of fully synthetic glycopeptide conjugate vaccines. *Infect Immun*, **1997**, 65, (12), 4918-25.

66. Wong, C. H.; Halcomb, R. L.; Ichikawa, Y.; Kajimoto, T., Enzymes in Organic Synthesis: Application to the Problems of Carbohydrate Recognition. *Angew. Chem. Int. Ed. Engl.*, **1995**, 34, (4), 412-432.
67. Hancock, S. M.; Vaughan, M. D.; Withers, S. G., Engineering of glycosidases and glycosyltransferases. *Curr Opin Chem Biol*, **2006**, 10, (5), 509-19.
68. Joosten, J. A.; Lazet, B. J.; Kamerling, J. P.; Vliegthart, J. F., Chemo-enzymatic synthesis of tetra-, penta-, and hexasaccharide fragments of the capsular polysaccharide of *Streptococcus pneumoniae* type 14. *Carbohydr Res*, **2003**, 338, (23), 2629-51.
69. Michalik, D.; Vliegthart, J. F. G.; Kamerling, J. P., Chemoenzymic synthesis of oligosaccharide fragments of the capsular polysaccharide of *Streptococcus pneumoniae* type 14. *J. Chem. Soc., Perkin Trans.1*, **2002**, 1973-1981.
70. Yan, F.; Wakarchuk, W. W.; Gilbert, M.; Richards, J. C.; Whitfield, D. M., Polymer-supported and chemoenzymatic synthesis of the *Neisseria meningitidis* pentasaccharide: a methodological comparison. *Carbohydr Res*, **2000**, 328, (1), 3-16.
71. Ruffing, A.; Chen, R. R., Metabolic engineering of microbes for oligosaccharide and polysaccharide synthesis. *Microb Cell Fact*, **2006**, 5, 25.
72. Koizumi, S.; Endo, T.; Tabata, K.; Ozaki, A., Large-scale production of UDP-galactose and globotriose by coupling metabolically engineered bacteria. *Nat Biotechnol*, **1998**, 16, (9), 847-50.
73. Jennison, A. V.; Verma, N. K., *Shigella flexneri* infection: pathogenesis and vaccine development. *FEMS Microbiol Rev*, **2004**, 28, (1), 43-58.
74. Phalipon, A.; Sansonetti, P. J., *Shigella's* ways of manipulating the host intestinal innate and adaptive immune system: a tool box for survival? *Immunol Cell Biol*, **2007**, 85, (2), 119-29.
75. Levine, M. M.; Kotloff, K. L.; Barry, E. M.; Pasetti, M. F.; Sztein, M. B., Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol*, **2007**, 5, (7), 540-53.
76. Lan, R.; Reeves, P. R., Escherichia coli in disguise: molecular origins of *Shigella*. *Microbes Infect*, **2002**, 4, (11), 1125-32.
77. Kotloff, K. L.; Winickoff, J. P.; Ivanoff, B.; Clemens, J. D.; Swerdlow, D. L.; Sansonetti, P. J.; Adak, G. K.; Levine, M. M., Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ*, **1999**, 77, (8), 651-66.
78. DuPont, H. L.; Levine, M. M.; Hornick, R. B.; Formal, S. B., Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis*, **1989**, 159, (6), 1126-8.
79. Lindberg, A. A.; Karnell, A.; Weintraub, A., The lipopolysaccharide of *Shigella bacteria* as a virulence factor. *Rev Infect Dis*, **1991**, 13 Suppl 4, S279-84.
80. Tran Van Nhieu, G.; Parsot, C.; Sansonetti, P., Les protéines Ipa et la réorganisation du cytosquelette au moment de l'entrée de *Shigella flexneri* dans les cellules épithéliales *Annales de l'Institut Pasteur*, **1997**, 8, (2), 139-146.
81. Hueck, C. J., Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev*, **1998**, 62, (2), 379-433.
82. Galan, J. E.; Collmer, A., Type III secretion machines: bacterial devices for protein delivery into host cells. *Science*, **1999**, 284, (5418), 1322-8.
83. Lasa, I.; Egile, C.; Cossart, P.; Sansonetti, P., Motilité intracellulaire et polymérisation de l'actine par *Listeria monocytogenes* et *Shigella flexneri* *Annales de l'Institut Pasteur*, **1997**, 8, (2), 163-172.

84. Sansonetti, P. J.; Phalipon, A.; Arondel, J.; Thirumalai, K.; Banerjee, S.; Akira, S.; Takeda, K.; Zychlinsky, A., Caspase-1 activation of IL-1beta and IL-18 are essential for *Shigella flexneri*-induced inflammation. *Immunity*, **2000**, 12, (5), 581-90.
85. Philpott, D. J.; Yamaoka, S.; Israel, A.; Sansonetti, P. J., Invasive *Shigella flexneri* activates NF-kappa B through a lipopolysaccharide-dependent innate intracellular response and leads to IL-8 expression in epithelial cells. *J Immunol*, **2000**, 165, (2), 903-14.
86. Kim, D. W.; Lenzen, G.; Page, A. L.; Legrain, P.; Sansonetti, P. J.; Parsot, C., The *Shigella flexneri* effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proc Natl Acad Sci U S A*, **2005**, 102, (39), 14046-51.
87. Arbibe, L.; Kim, D. W.; Batsche, E.; Pedron, T.; Mateescu, B.; Muchardt, C.; Parsot, C.; Sansonetti, P. J., An injected bacterial effector targets chromatin access for transcription factor NF-kappaB to alter transcription of host genes involved in immune responses. *Nat Immunol*, **2007**, 8, (1), 47-56.
88. Islam, D.; Bardhan, P. K.; Lindberg, A. A.; Christensson, B., *Shigella* infection induces cellular activation of T and B cells and distinct species-related changes in peripheral blood lymphocyte subsets during the course of the disease. *Infect Immun*, **1995**, 63, (8), 2941-9.
89. Islam, D.; Veress, B.; Bardhan, P. K.; Lindberg, A. A.; Christensson, B., In situ characterization of inflammatory responses in the rectal mucosae of patients with shigellosis. *Infect Immun*, **1997**, 65, (2), 739-49.
90. Way, S. S.; Borczuk, A. C.; Goldberg, M. B., Adaptive immune response to *Shigella flexneri* 2a cydC in immunocompetent mice and mice lacking immunoglobulin A. *Infect Immun*, **1999**, 67, (4), 2001-4.
91. Mel, D.; Papo, R.; Terzin, A.; Vuksic, L., Studies on vaccination against bacillary dysentery. *Bull World Health Organ*, **1965**, 32, 647-677.
92. Phalipon, A.; Kaufmann, M.; Michetti, P.; Cavailon, J. M.; Huerre, M.; Sansonetti, P.; Kraehenbuhl, J. P., Monoclonal immunoglobulin A antibody directed against serotype-specific epitope of *Shigella flexneri* lipopolysaccharide protects against murine experimental shigellosis. *J Exp Med*, **1995**, 182, (3), 769-78.
93. WHO, *Weekly epidemiological Record*, **1997**, 72, 73-80.
94. Phalipon, A.; Sansonetti, P., Live attenuated *Shigella flexneri* mutants as vaccine candidates against shigellosis and vectors for antigen delivery. *Biologicals*, **1995**, 23, (2), 125-34.
95. WHO, *Weekly epidemiological Record*, **2006**, 81, 49-60.
96. Coster, T. S.; Hoge, C. W.; VanDeVerg, L. L.; Hartman, A. B.; Oaks, E. V.; Venkatesan, M. M.; Cohen, D.; Robin, G.; Fontaine-Thompson, A.; Sansonetti, P. J.; Hale, T. L., Vaccination against shigellosis with attenuated *Shigella flexneri* 2a strain SC602. *Infect Immun*, **1999**, 67, (7), 3437-43.
97. Noriega, F. R.; Losonsky, G.; Wang, J. Y.; Formal, S. B.; Levine, M. M., Further characterization of delta aroA delta virG *Shigella flexneri* 2a strain CVD 1203 as a mucosal *Shigella* vaccine and as a live-vector vaccine for delivering antigens of enterotoxigenic *Escherichia coli*. *Infect Immun*, **1996**, 64, (1), 23-7.
98. Kotloff, K. L.; Pasetti, M. F.; Barry, E. M.; Nataro, J. P.; Wasserman, S. S.; Sztein, M. B.; Picking, W. D.; Levine, M. M., Deletion in the *Shigella* enterotoxin genes further attenuates *Shigella flexneri* 2a bearing guanine auxotrophy in a phase 1 trial of CVD 1204 and CVD 1208. *J Infect Dis*, **2004**, 190, (10), 1745-54.
99. Kotloff, K. L.; Noriega, F. R.; Samandari, T.; Sztein, M. B.; Losonsky, G. A.; Nataro, J. P.; Picking, W. D.; Barry, E. M.; Levine, M. M., *Shigella flexneri* 2a strain CVD 1207, with specific deletions in virG, sen, set, and guaBA, is highly attenuated in humans. *Infect Immun*, **2000**, 68, (3), 1034-9.

- 100.** Katz, D. E.; Coster, T. S.; Wolf, M. K.; Trespalacios, F. C.; Cohen, D.; Robins, G.; Hartman, A. B.; Venkatesan, M. M.; Taylor, D. N.; Hale, T. L., Two studies evaluating the safety and immunogenicity of a live, attenuated *Shigella flexneri* 2a vaccine (SC602) and excretion of vaccine organisms in North American volunteers. *Infect Immun*, **2004**, 72, (2), 923-30.
- 101.** Hartman, A. B.; Venkatesan, M. M., Construction of a stable attenuated *Shigella sonnei* DeltavirG vaccine strain, WRSS1, and protective efficacy and immunogenicity in the guinea pig keratoconjunctivitis model. *Infect Immun*, **1998**, 66, (9), 4572-6.
- 102.** Orr, N.; Katz, D. E.; Atsmon, J.; Radu, P.; Yavzori, M.; Halperin, T.; Sela, T.; Kayouf, R.; Klein, Z.; Ambar, R.; Cohen, D.; Wolf, M. K.; Venkatesan, M. M.; Hale, T. L., Community-based safety, immunogenicity, and transmissibility study of the *Shigella sonnei* WRSS1 vaccine in Israeli volunteers. *Infect Immun*, **2005**, 73, (12), 8027-32.
- 103.** Pavliakova, D.; Chu, C.; Bystricky, S.; Tolson, N. W.; Shiloach, J.; Kaufman, J. B.; Bryla, D. A.; Robbins, J. B.; Schneerson, R., Treatment with succinic anhydride improves the immunogenicity of *Shigella flexneri* type 2a O-specific polysaccharide-protein conjugates in mice. *Infect Immun*, **1999**, 67, (10), 5526-9.
- 104.** Passwell, J. H.; Ashkenazi, S.; Harlev, E.; Miron, D.; Ramon, R.; Farzam, N.; Lerner-Geva, L.; Levi, Y.; Chu, C.; Shiloach, J.; Robbins, J. B.; Schneerson, R., Safety and immunogenicity of *Shigella sonnei*-CRM9 and *Shigella flexneri* type 2a-rEPASucc conjugate vaccines in one- to four-year-old children. *Pediatr Infect Dis J*, **2003**, 22, (8), 701-6.
- 105.** Orr, N.; Robin, G.; Cohen, D.; Arnon, R.; Lowell, G. H., Immunogenicity and efficacy of oral or intranasal *Shigella flexneri* 2a and *Shigella sonnei* proteosome-lipopolysaccharide vaccines in animal models. *Infect Immun*, **1993**, 61, (6), 2390-5.
- 106.** Levenson, V. I.; Egorova, T. P.; Belkin, Z. P.; Fedosova, V. G.; Subbotina, J. L.; Rukhadze, E. Z.; Dzhikidze, E. K.; Stassilevich, Z. K., Protective ribosomal preparation from *Shigella sonnei* as a parenteral candidate vaccine. *Infect Immun*, **1991**, 59, (10), 3610-8.
- 107.** Shim, D. H.; Chang, S. Y.; Park, S. M.; Jang, H.; Carbis, R.; Czerkinsky, C.; Uematsu, S.; Akira, S.; Kweon, M. N., Immunogenicity and protective efficacy offered by a ribosomal-based vaccine from *Shigella flexneri* 2a. *Vaccine*, **2007**, 25, (25), 4828-36.
- 108.** Dmitriev, B. A.; Knirel, Y. A.; Kochetkov, N. K., Somatic antigens of shigella. Structural investigation on the O-specific polysaccharide chain of *Shigella dysenteriae* type 1 lipopolysaccharide. *Eur J Biochem*, **1976**, 66, (3), 559-66.
- 109.** Pozsgay, V., Synthesis of Glycoconjugate Vaccines against *Shigella dysenteriae* Type 1. *J Org Chem*, **1998**, 63, (17), 5983-5999.
- 110.** Pozsgay, V.; Kubler-Kielb, J.; Schneerson, R.; Robbins, J. B., Effect of the nonreducing end of *Shigella dysenteriae* type 1 O-specific oligosaccharides on their immunogenicity as conjugates in mice. *Proc Natl Acad Sci U S A*, **2007**, 104, (36), 14478-82.
- 111.** Wright, K.; Guerreiro, C.; Laurent, I.; Baleux, F.; Mulard, L. A., Preparation of synthetic glycoconjugates as potential vaccines against *Shigella flexneri* serotype 2a disease. *Org Biomol Chem*, **2004**, 2, (10), 1518-27.
- 112.** Belot, F.; Guerreiro, C.; Baleux, F.; Mulard, L. A., Synthesis of two linear PADRE conjugates bearing a deca- or pentadecasaccharide B epitope as potential synthetic vaccines against *Shigella flexneri* serotype 2a infection. *Chemistry*, **2005**, 11, (5), 1625-35.
- 113.** Henrissat, B., A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, **1991**, 280 (Pt 2), 309-16.
- 114.** Davies, G.; Henrissat, B., Structures and mechanisms of glycosyl hydrolases. *Structure*, **1995**, 3, (9), 853-9.
- 115.** Henrissat, B.; Bairoch, A., New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, **1993**, 293 (Pt 3), 781-8.

116. Henrissat, B.; Bairoch, A., Updating the sequence-based classification of glycosyl hydrolases. *Biochem J*, **1996**, 316 (Pt 2), 695-6.
117. Campbell, J. A.; Davies, G. J.; Bulone, V.; Henrissat, B., A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J*, **1997**, 326 (Pt 3), 929-39.
118. Coutinho, P. M.; Henrissat, B., Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering*, The Royal Society of Chemistry Cambridge, **1999**, 3.
119. Allison, G. E.; Verma, N. K., Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol*, **2000**, 8, (1), 17-23.
120. Guan, S.; Bastin, D. A.; Verma, N. K., Functional analysis of the O antigen glucosylation gene cluster of *Shigella flexneri* bacteriophage SfX. *Microbiology*, **1999**, 145 (Pt 5), 1263-73.
121. Korres, H.; Mavris, M.; Morona, R.; Manning, P. A.; Verma, N. K., Topological analysis of GtrA and GtrB proteins encoded by the serotype-converting cassette of *Shigella flexneri*. *Biochem Biophys Res Commun*, **2005**, 328, (4), 1252-60.
122. Lehane, A. M.; Korres, H.; Verma, N. K., Bacteriophage-encoded glucosyltransferase GtrII of *Shigella flexneri*: membrane topology and identification of critical residues. *Biochem J*, **2005**, 389, (Pt 1), 137-43.
123. Korres, H.; Verma, N. K., Identification of essential loops and residues of glucosyltransferase V (GtrV) of *Shigella flexneri*. *Mol Membr Biol*, **2006**, 23, (5), 407-19.
124. Korres, H.; Verma, N. K., Topological analysis of glucosyltransferase GtrV of *Shigella flexneri* by a dual reporter system and identification of a unique reentrant loop. *J Biol Chem*, **2004**, 279, (21), 22469-76.
125. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B., An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol*, **2003**, 328, (2), 307-17.
126. Davies, G. J.; Gloster, T. M.; Henrissat, B., Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr Opin Struct Biol*, **2005**, 15, (6), 637-45.
127. Segawa, H.; Miyamoto, K.; Ogura, Y.; Haga, H.; Morita, K.; Katai, K.; Tatsumi, S.; Nii, T.; Taketani, Y.; Takeda, E., Cloning, functional expression and dietary regulation of the mouse neutral and basic amino acid transporter (NBAT). *Biochem J*, **1997**, 328 (Pt 2), 657-64.
128. Durand, A.; Hughes, R.; Roussel, A.; Flatman, R.; Henrissat, B.; Juge, N., Emergence of a subfamily of xylanase inhibitors within glycoside hydrolase family 18. *Febs J*, **2005**, 272, (7), 1745-55.
129. Houston, D. R.; Recklies, A. D.; Krupa, J. C.; van Aalten, D. M., Structure and ligand-induced conformational change of the 39-kDa glycoprotein from human articular chondrocytes. *J Biol Chem*, **2003**, 278, (32), 30206-12.
130. Kitahara, S.; Murakami, H.; Sone, Y.; Misaki, A., The acceptor specificity of amyloamylase from *Eschenchia coli* IFO 3806. *Agric. Biol. Chem.*, **1989**, 53, (10), 2661-2666.
131. Haworth, W. N.; Hirst, E. L.; Reynolds, R. J. W., Maltal and 4- α -glucosidomannose. *J Chem Soc*, **1934**, 302.
132. Aisaka, K.; Masuda-Kato, T.; Chikamune, T.; Kamitori, K.; Uosaki, Y.; Saito, Y., Enzymatic synthesis of novel disaccharides using disaccharide phosphorylases. *J Biosci Bioeng*, **2000**, 90, (2), 208-13.
133. Seymour, F. R.; Knapp, R. D.; Chen, E. C. M.; Bishop, H.; Jeanes, A., Structural analysis of *Leuconostoc* dextrans containing 3-O- α -D-glucosyl residues in both linear-chain and branch-point positions, or only in branch-point positions, by methylation and by ¹³C-N.M.R. spectroscopy. *Carbohydr Res*, **1979**, 74, 41-62.

134. Jeanes, A.; Haynes, W.; Williams, C.; Rankin, J.; Melvin, E.; Austin, M.; Clusray, J.; Fisher, B.; Tsuchiya, A.; Rist, C., Characterisation and classification of dextrans from ninety six strains of bacteria. *J Am Chem Soc*, **1954**, (76), 5041-5052.
135. Cote, G. L.; Robyt, J. F., Isolation and partial characterization of an extracellular glucansucrase from *Leuconostoc mesenteroides* NRRL B-1355 that synthesizes an alternating (1 goes to 6), (1 goes to 3)-alpha-D-glucan. *Carbohydr Res*, **1982**, 101, (1), 57-74.
136. Sidebotham, R. L., Dextrans. *Adv Carbohydr Chem Biochem*, **1974**, 30, 371-444.
137. Kralj, S.; van Geel-Schutten, G. H.; Rahaoui, H.; Leer, R. J.; Faber, E. J.; van der Maarel, M. J.; Dijkhuizen, L., Molecular characterization of a novel glucosyltransferase from *Lactobacillus reuteri* strain 121 synthesizing a unique, highly branched glucan with α -(1 \rightarrow 4) and α -(1 \rightarrow 6) glucosidic bonds. *Appl Environ Microbiol*, **2002**, 68, (9), 4283-91.
138. Kralj, S.; Stripling, E.; Sanders, P.; van Geel-Schutten, G. H.; Dijkhuizen, L., Highly hydrolytic reuteransucrase from probiotic *Lactobacillus reuteri* strain ATCC 55730. *Appl Environ Microbiol*, **2005**, 71, (7), 3942-50.
139. Monsan, P.; Bozonnet, S.; Albenne, C.; Joucla, G.; Willemot, R. M.; Remaud-Simeon, M., Homopolysaccharides from lactic acid bacteria. *International Dairy Journal*, **2001**, 11, (9), 675-685.
140. Sato, S.; Koga, T.; Inoue, M., Isolation and some properties of extracellular -glucosyltransferases and -fructosyltransferases from *Streptococcus mutans* serotypes c, e, and f. *Carbohydr Res*, **1984**, 68, 113-121.
141. Côté, G. L.; Robyt, J. F., The formation of α -(1,3) branch linkages by an exocellular glucansucrase from *Leuconostoc mesenteroides* NRRL B-742. *Carbohydr Res*, **1983**, 119, 141-156.
142. Smith, M. R.; Zahnley, J. C.; Wong, R. Y.; Lundin, R. E.; Ahlgren, J. A., A mutant strain of *Leuconostoc mesenteroides* B-1355 producing a glucosyltransferase synthesizing α (1-2) glucosidic linkages. *Journal of industrial microbiology & biotechnology*, **1998**, 21, 37-45.
143. Bozonnet, S.; Dols-Laffargue, M.; Fabre, E.; Pizzut, S.; Remaud-Simeon, M.; Monsan, P.; Willemot, R. M., Molecular characterization of DSR-E, an α (1,2) linkage-synthesizing dextransucrase with two catalytic domains. *J Bacteriol*, **2002**, 184, (20), 5753-61.
144. Monchois, V.; Willemot, R. M.; Remaud-Simeon, M.; Croux, C.; Monsan, P., Cloning and sequencing of a gene coding for a novel dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only α (1-6) and α (1-3) linkages. *Gene*, **1996**, 182, (1-2), 23-32.
145. Kralj, S.; van Geel-Schutten, G. H.; Dondorff, M. M.; Kirsanovs, S.; van der Maarel, M. J.; Dijkhuizen, L., Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. *Microbiology*, **2004**, 150, (Pt 11), 3681-90.
146. De Montalk, G. P.; Remaud-Simeon, M.; Willemot, R. M.; Planchot, V.; Monsan, P., Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J Bacteriol*, **1999**, 181, (2), 375-81.
147. Pizzut-Serin, S.; Potocki-Veronese, G.; van der Veen, B. A.; Albenne, C.; Monsan, P.; Remaud-Simeon, M., Characterisation of a novel amylosucrase from *Deinococcus radiodurans*. *FEBS Lett*, **2005**, 579, (6), 1405-10.
148. Emond, S., Characterisation of a novel amylosucrase from *Deinococcus geothermalis*. *unpublished results*.
149. Monchois, V.; Remaud-Simeon, M.; Russell, R. R.; Monsan, P.; Willemot, R. M., 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*, **1997**, 48, (4), 465-72.

150. Funane, K.; Mizuno, K.; Takahara, H.; Kobayashi, M., Gene encoding a dextransucrase-like protein in *Leuconostoc mesenteroides* NRRL B-512F. *Biosci.Biotechnol.Biochem.*, **2000**, 64, (1), 29-38.
151. Dols, M.; Remaud-Simeon, M.; Vignon, M.; Monsan, P., Characterisation of dextransucrases from *Leuconostoc mesenteroides* NRRL B-1299. *App Biochem Biotechnol*, **1997**, 62, 47-59.
152. Monchois, V.; Willemot, R. M.; Reyn, A.; Croux, C.; Monsan, P., Cloning and sequencing of a gene coding for a novel dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299 synthesising only α -1,6 and α -1,3 linkages. *Gene*, **1996**, 182, 23-32.
153. Monchois, V.; Willemot, R. M.; Remaud-Simeon, M.; Croux, C.; Soucaille, P.; Monsan, P., Primary structure of dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299. Comparison with dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *Ann. N.Y. Acad. Sci.*, **1996**, 799, 151-156.
154. Arguello-Morales, M. A.; Remaud-Simeon, M.; Pizzut, S.; Sarcabal, P.; Willemot, R.; Monsan, P., Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. *FEMS Microbiol Lett*, **2000**, 182, (1), 81-5.
155. Neubauer, H.; Bauche, A.; Mollet, B., Molecular characterization and expression analysis of the dextransucrase DsrD of *Leuconostoc mesenteroides* Lcc4 in homologous and heterologous *Lactococcus lactis* cultures. *Microbiology*, **2003**, 149, (Pt 4), 973-82.
156. Shimamura, A.; Nakano, Y. J.; Mukasa, H.; Kuramitsu, H. K., Identification of amino acid residues in *Streptococcus mutans* glucosyltransferases influencing the structure of the glucan product. *J Bacteriol*, **1994**, 176, (16), 4845-50.
157. Fujiwara, T.; Hoshino, T.; Ooshima, T.; Sobue, S.; Hamada, S., Purification, characterization, and molecular analysis of the gene encoding glucosyltransferase from *Streptococcus oralis*. *Infect Immun*, **2000**, 68, (5), 2475-83.
158. Seibel, J.; Hellmuth, H.; Hofer, B.; Kicinska, A. M.; Schmalbruch, B., Identification of new acceptor specificities of glycosyltransferase R with the aid of substrate microarrays. *Chembiochem*, **2006**, 7, (2), 310-20.
159. Gilmore, K. S.; Russell, R. R.; Ferretti, J. J., Analysis of the *Streptococcus downei* gtfS gene, which specifies a glucosyltransferase that synthesizes soluble glucans. *Infect Immun*, **1990**, 58, (8), 2452-8.
160. Hanada, N.; Isobe, Y.; Aizawa, Y.; Katayama, T.; Sato, S.; Inoue, M., Nucleotide sequence analysis of the gtfT gene from *Streptococcus sobrinus* OMZ176. *Infect Immun*, **1993**, 61, (5), 2096-103.
161. Hanada, N.; Fukushima, K.; Nomura, Y.; Senpuku, H.; Hayakawa, M.; Mukasa, H.; Shiroza, T.; Abiko, Y., Cloning and nucleotide sequence analysis of the *Streptococcus sobrinus* gtfU gene that produces a highly branched water-soluble glucan. *Biochim Biophys Acta*, **2002**, 1570, (1), 75-9.
162. Simpson, C. L.; Cheetham, N. W.; Giffard, P. M.; Jacques, N. A., Four glucosyltransferases, GtfJ, GtfK, GtfL and GtfM, from *Streptococcus salivarius* ATCC 25975. *Microbiology*, **1995**, 141 (Pt 6), 1451-60.
163. Vickerman, M. M.; Sulavik, M. C.; Nowak, J. D.; Gardner, N. M.; Jones, G. W.; Clewell, D. B., Nucleotide sequence analysis of the *Streptococcus gordonii* glucosyltransferase gene, gtfG. *DNA Seq*, **1997**, 7, (2), 83-95.
164. Shiroza, T.; Ueda, S.; Kuramitsu, H. K., Sequence analysis of the gtfB gene from *Streptococcus mutans*. *J Bacteriol*, **1987**, 169, (9), 4263-70.
165. Ueda, S.; Shiroza, T.; Kuramitsu, H. K., Sequence analysis of the gtfC gene from *Streptococcus mutans* GS-5. *Gene*, **1988**, 69, (1), 101-9.

166. Russell, R. R. B.; Gilpin, M. L.; Mukasa, H.; Dougan, G., Characterization of glucosyltransferase expressed from a *Streptococcus sobrinus* gene cloned in *Escherichia coli*. *J Gen Microbiol*, **1987**, 133, (4), 935-44.
167. Konishi, N.; Torii, Y.; Yamamoto, T.; Ohta, H.; Fukui, K.; Ohta, H.; Hanamoto, S.; Matsuno, H.; Komatsu, H.; Kodama, T.; Katayama, E., Structure and enzymatic properties of genetically truncated forms of the water-insoluble glucan-synthesizing glucosyltransferase from *Streptococcus sobrinus* *Journal of Biochemistry (Tokyo)*, **1999**, 126, 287.
168. Fabre, E.; Bozonnet, S.; Arcache, A.; Willemot, R. M.; Vignon, M.; Monsan, P.; Remaud-Simeon, M., 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*, **2005**, 187, (1), 296-303.
169. Monchois, V.; Willemot, R. M.; Monsan, P., Glucansucrases: mechanism of action and structure-function relationships. *FEMS Microbiol Rev*, **1999**, 23, (2), 131-51.
170. Pijning, T., *unpublished results*.
171. Mooser, G.; Hefta, S. A.; Paxton, R. J.; Shively, J. E.; Lee, T. D., Isolation and sequence of an active-site peptide containing a catalytic aspartic acid from two *Streptococcus sobrinus* alpha-glucosyltransferases. *J Biol Chem*, **1991**, 266, (14), 8916-22.
172. Devulapalle, K. S.; Goodman, S. D.; Gao, Q.; Hemsley, A.; Mooser, G., Knowledge-based model of a glucosyltransferase from the oral bacterial group of mutans streptococci. *Protein Sci*, **1997**, 6, (12), 2489-93.
173. MacGregor, E. A.; Jespersen, H. M.; Svensson, B., A circularly permuted alpha-amylase-type α/β -barrel structure in glucan-synthesizing glucosyltransferases. *FEBS Lett*, **1996**, 378, (3), 263-6.
174. Tsai, Y. W.; Chia, J. S.; Shiau, Y. Y.; Chou, H. C.; Liaw, Y. C.; Lou, K. L., Three-dimensional modelling of the catalytic domain of *Streptococcus mutans* glucosyltransferase GtfB. *FEMS Microbiol Lett*, **2000**, 188, (1), 75-9.
175. Kato, C.; Nakano, Y.; Lis, M.; Kuramitsu, H. K., Molecular genetic analysis of the catalytic site of *Streptococcus mutans* glucosyltransferases. *Biochem Biophys Res Commun*, **1992**, 189, (2), 1184-8.
176. Monchois, V.; Vignon, M.; Escalier, P. C.; Svensson, B.; Russell, R. R., Involvement of Gln937 of *Streptococcus downei* GTF-I glucansucrase in transition-state stabilization. *Eur J Biochem*, **2000**, 267, (13), 4127-36.
177. Tsumori, H.; Minami, T.; Kuramitsu, H. K., Identification of essential amino acids in the *Streptococcus mutans* glucosyltransferases. *J Bacteriol*, **1997**, 179, (11), 3391-6.
178. Mooser, G.; Iwaoka, K. R., Sucrose 6- α -D-glucosyltransferase from *Streptococcus sobrinus*: characterization of a glucosyl-enzyme complex. *Biochemistry*, **1989**, 28, (2), 443-9.
179. Skov, L. K.; Mirza, O.; Henriksen, A.; De Montalk, G. P.; Remaud-Simeon, M.; Sarcabal, P.; Willemot, R. M.; Monsan, P.; Gajhede, M., Amylosucrase, a glucan-synthesizing enzyme from the α -amylase family. *J Biol Chem*, **2001**, 276, (27), 25273-8.
180. Sinnott, M. L., Catalytic mechanisms of glycosyl transfer. *Chem. Rev.*, **1990**, 90, 1171-1202.
181. McCarter, J. D.; Withers, S. G., Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol*, **1994**, 4, (6), 885-92.
182. Mosi, R.; He, S.; Uitdehaag, J.; Dijkstra, B. W.; Withers, S. G., Trapping and characterization of the reaction intermediate in cyclodextrin glucosyltransferase by use of activated substrates and a mutant enzyme. *Biochemistry*, **1997**, 36, (32), 9927-34.
183. Uitdehaag, J. C.; Mosi, R.; Kalk, K. H.; van der Veen, B. A.; Dijkhuizen, L.; Withers, S. G.; Dijkstra, B. W., X-ray structures along the reaction pathway of cyclodextrin glucosyltransferase elucidate catalysis in the α -amylase family. *Nat Struct Biol*, **1999**, 6, (5), 432-6.

184. Yoshioka, Y.; Hasegawa, K.; Matsuura, Y.; Katsube, Y.; Kubota, M., Crystal structures of a mutant maltotetraose-forming exo-amylase cocrystallized with maltopentaose. *J Mol Biol*, **1997**, 271, (4), 619-28.
185. Jencks, W. P., Binding energy, specificity, and enzymic catalysis: the circe effect. *Adv Enzymol Relat Areas Mol Biol*, **1975**, 43, 219-410.
186. Jensen, M. H.; Mirza, O.; Albenne, C.; Remaud-Simeon, M.; Monsan, P.; Gajhede, M.; Skov, L. K., Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. *Biochemistry*, **2004**, 43, (11), 3104-10.
187. Koshland, D. E., Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev. Camb. Philos. Soc.*, **1953**, 28, 416-436.
188. Mayer, R. M., Dextranucrase: a glucosyltransferase from *Streptococcus sanguis*. *Methods Enzymol*, **1987**, 138, 649-61.
189. Luzio, G. A.; Parnaik, V. K.; Mayer, R. M., A D-glucosylated form of dextranucrase: demonstration of partial reactions. *Carbohydr Res*, **1983**, 121, 269-78.
190. Koepsell, H. J.; Tsuchiya, H. M.; Hellman, N. N.; Kazenko, A.; Hoffman, C. A.; Sharpe, E. S.; Jackson, R. W., Enzymatic synthesis of dextran; acceptor specificity and chain initiation. *J Biol Chem*, **1953**, 200, (2), 793-801.
191. 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., Molecular basis of the amylose-like polymer formation catalyzed by *Neisseria polysaccharea* amylosucrase. *J Biol Chem*, **2004**, 279, (1), 726-34.
192. Moulis, C.; Joucla, G.; Harrison, D.; Fabre, E.; Potocki-Veronese, G.; Monsan, P.; Remaud-Simeon, M., Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J Biol Chem*, **2006**, 281, (42), 31254-67.
193. Binder, T. P.; Robyt, J. F., p-Nitrophenyl α -D-glucopyranoside, a new substrate for glucansucrases. *Carbohydr Res*, **1983**, 124, (2), 287-99.
194. Binder, T. P.; Robyt, J. F., Inhibition of *Streptococcus mutans* 6715 glucosyltransferases by sucrose analogs modified at positions 6 and 6'. *Carbohydr Res*, **1985**, 140, (1), 9-20.
195. Robyt, J. F.; Walseth, T. F., The mechanism of acceptor reactions of *Leuconostoc mesenteroides* B-512F dextranucrase. *Carbohydr Res*, **1978**, 61, 433-45.
196. Potocki de Montalk, G.; Remaud-Simeon, M.; Willemot, R. M.; Monsan, P., Characterisation of the activator effect of glycogen on amylosucrase from *Neisseria polysaccharea*. *FEMS Microbiol Lett*, **2000**, 186, (1), 103-8.
197. Robyt, J. F.; Eklund, S. H., Relative, quantitative effects of acceptors in the reaction of *Leuconostoc mesenteroides* B-512F dextranucrase. *Carbohydr Res*, **1983**, 121, 279-86.
198. Cote, G. L.; Dunlap, C. A., Alternansucrase acceptor reactions with methyl hexopyranosides. *Carbohydr Res*, **2003**, 338, (19), 1961-7.
199. Cote, G. L.; Holt, S. M.; Miller-Fosmore, C. M., Prebiotic oligosaccharides via alternansucrase acceptor reactions. *American Chemical Society*, **2003**, 849, 76-89.
200. Paul, F.; Oriol, E.; Auriol, D.; Monsan, P., Acceptor reactions of a highly purified dextranucrase with maltose and oligosaccharides. Application to the synthesis of controlled molecular weight dextrans. *Carbohydr Res*, **1986**, 149, 433-441.
201. Dols-Lafargue, M.; Willemot, R. M.; Monsan, P. F.; Remaud-Simeon, M., Factors affecting α -1,2 glucooligosaccharide synthesis by *Leuconostoc mesenteroides* NRRL B-1299 dextranucrase. *Biotechnol Bioeng*, **2001**, 74, (6), 498-504.
202. Richard, G.; Yu, S.; Monsan, P.; Remaud-Simeon, M.; Morel, S., A novel family of glucosyl 1,5-anhydro-D-fructose derivatives synthesised by transglucosylation with dextranucrase from *Leuconostoc mesenteroides* NRRL B-512F. *Carbohydr Res*, **2005**, 340, (3), 395-401.

203. Devulapalle, K. S.; Mooser, G., Subsite specificity of the active site of glucosyltransferases from *Streptococcus sobrinus*. *J Biol Chem*, **1994**, 269, (16), 11967-71.
204. Champion, E., *unpublished results*.
205. Hellmuth, H.; Hillringhaus, L.; Hobbel, S.; Kralj, S.; Dijkhuizen, L.; Seibel, J., Highly efficient chemoenzymatic synthesis of novel branched thiooligosaccharides by substrate direction with glucansucrases. *Chembiochem*, **2007**, 8, (3), 273-6.
206. Cote, G. L.; Dunlap, C. A.; Appell, M.; Momany, F. A., Alternansucrase acceptor reactions with D-tagatose and L-glucose. *Carbohydr Res*, **2005**, 340, (2), 257-62.
207. Arguello Morales, M. A.; Remaud-Simeon, M.; Willemot, R. M.; Vignon, M. R.; Monsan, P., Novel oligosaccharides synthesized from sucrose donor and cellobiose acceptor by alternansucrase. *Carbohydr Res*, **2001**, 331, (4), 403-11.
208. Yoon, S. H.; Robyt, J. F., Synthesis of acarbose analogues by transglycosylation reactions of *Leuconostoc mesenteroides* B-512FMC and B-742CB dextransucrases. *Carbohydr Res*, **2002**, 337, (24), 2427-35.
209. Rolland-Sabaté, A.; Colonna, P.; Potocki-Véronèse, G.; Monsan, P.; Planchot, V., Elongation and insolubilisation of α -glucans by the action of *Neisseria polysaccharea* amylosucrase. *Journal of Cereal Science*, **2004**, 40, (1), 17-30.
210. Demuth, K.; Jordening, H. J.; Buchholz, K., Oligosaccharide synthesis by dextransucrase: new unconventional acceptors. *Carbohydr Res*, **2002**, 337, (20), 1811-20.
211. Richard, G.; Morel, S.; Willemot, R. M.; Monsan, P.; Remaud-Simeon, M., Glucosylation of α -butyl- and α -octyl-D-glucopyranosides by dextransucrase and alternansucrase from *Leuconostoc mesenteroides*. *Carbohydr Res*, **2003**, 338, (9), 855-64.
212. Nakahara, K.; Kontani, M.; Ono, H.; Kodama, T.; Tanaka, T.; Ooshima, T.; Hamada, S., Glucosyltransferase from *Streptococcus sobrinus* Catalyzes Glucosylation of Catechin. *Appl Environ Microbiol*, **1995**, 61, (7), 2768-2770.
213. Meulenbeld, G. H.; Zuilhof, H.; van Veldhuizen, A.; van den Heuvel, R. H.; Hartmans, S., Enhanced (+)-catechin transglucosylating activity of *Streptococcus mutans* GS-5 glucosyltransferase-D due to fructose removal. *Appl Environ Microbiol*, **1999**, 65, (9), 4141-7.
214. Seo, E. S.; Lee, J. H.; Park, J. Y.; Kim, D.; Han, H. J.; Robyt, J. F., Enzymatic synthesis and anti-coagulant effect of salicin analogs by using the *Leuconostoc mesenteroides* glucansucrase acceptor reaction. *J Biotechnol*, **2005**, 117, (1), 31-8.
215. Bertrand, A.; Morel, S.; Lefoulon, F.; Rolland, Y.; Monsan, P.; Remaud-Simeon, M., *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr Res*, **2006**, 341, (7), 855-63.
216. Bounaix, M. S., *unpublished results*.
217. Nugent, A. P., Health properties of resistant starch. *Nutrition Bulletin*, **2005**, 30, (1), 27-54.
218. Kochetkov, N. K.; Dmitriev, B. A.; Chernyak, A.; Levinsky, A. B., A new type of carbohydrate-containing synthetic antigen: synthesis of carbohydrate-containing polysaccharide copolymers with the specificity of o:3 and o:4 factors of *Salmonella*. *Carbohydr Res*, **1982**, 110, (2), C16-20.
219. Richard, G. Synthèse de dérivés glucosylés du 1,5-anhydro-D-fructose à l'aide de glucane-saccharides natives et recombinantes. Thèse de doctorat en biocatalyse, Institut National des Sciences Appliquées (INSA), Toulouse, **2004**.
220. Meulenbeld, G. H.; Hartmans, S., Transglycosylation by *Streptococcus mutans* GS-5 glucosyltransferase-D: acceptor specificity and engineering of reaction conditions. *Biotechnol Bioeng*, **2000**, 70, (4), 363-9.
221. Côté, G. L.; Holt, S. M. Prebiotic oligosaccharides via alternansucrase acceptor reactions. *US 7182954*, **2003**.

222. Stam, M. R.; Danchin, E. G.; Rancurel, C.; Coutinho, P. M.; Henrissat, B., Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of α -amylase-related proteins. *Protein Eng Des Sel*, **2006**, 19, (12), 555-62.
223. MacKenzie, C. R.; McDonald, I. J.; Johnson, K. G., Glycogen metabolism in the genus *Neisseria*: synthesis from sucrose by amylosucrase. *Can J Microbiol*, **1978**, 24, (4), 357-62.
224. Loesche, W. J., Bacterial succession in dental plaque: role in dental disease. In *Microbiology*, Ed. SCHLESSINGER D., A. S. M., Washington DC., Ed. 1975; pp 132-136.
225. Riou, J. Y.; Guibourdenche, M.; Popoff, M. Y., A new taxon in the genus *Neisseria*. *Ann Microbiol (Paris)*, **1983**, 134B, (2), 257-67.
226. Kim, H. S.; Park, H. J.; Heu, S.; Jung, J., Molecular and functional characterization of a unique sucrose hydrolase from *Xanthomonas axonopodis* pv. *glycines*. *J Bacteriol*, **2004**, 186, (2), 411-8.
227. Champion, E., Characterisation of a sucrose-hydrolase from *Xanthomonas campestris* pv. *campestris*. *unpublished results*.
228. Janecek, S., How many conserved sequence regions are there in the α -amylase family? *Biologia*, **2002**, 57, 29-41.
229. Nakajima, R.; Imanaka, T.; Aiba, S., Comparison of amino acid sequences of eleven different α -amylases. *Applied Microbiology and Biotechnology*, **1986**, 23, (5), 355-360.
230. Takata, H.; Kuriki, T.; Okada, S.; Takesada, Y.; Iizuka, M.; Minamiura, N.; Imanaka, T., Action of neopullulanase. Neopullulanase catalyzes both hydrolysis and transglycosylation at α -(1-4)- and α -(1-6)-glucosidic linkages. *J Biol Chem*, **1992**, 267, (26), 18447-52.
231. Boel, E.; Brady, L.; Brzozowski, A. M.; Derewenda, Z.; Dodson, G. G.; Jensen, V. J.; Petersen, S. B.; Swift, H.; Thim, L.; Woldike, H. F., Calcium binding in α -amylases: an X-ray diffraction study at 2.1-Å resolution of two enzymes from *Aspergillus*. *Biochemistry*, **1990**, 29, (26), 6244-9.
232. Machius, M.; Declerck, N.; Huber, R.; Wiegand, G., Activation of *Bacillus licheniformis* α -amylase through a disorder-->order transition of the substrate-binding site mediated by a calcium-sodium-calcium metal triad. *Structure*, **1998**, 6, (3), 281-92.
233. Janecek, S., Close evolutionary relatedness among functionally distantly related members of the (α/β)₈-barrel glycosyl hydrolases suggested by the similarity of their fifth conserved sequence region. *FEBS Lett*, **1995**, 377, (1), 6-8.
234. Albenne, C. Ingénierie rationnelle de l'amylosaccharase de *Neisseria polysaccharea*: de la structure tri-dimensionnelle aux bases moléculaires de la catalyse. Thèse de doctorat en biocatalyse, Institut National des Sciences Appliquées, Toulouse, **2002**.
235. Potocki de Montalk, G.; Remaud-Simeon, M.; Willemot, R. M.; Sarcabal, P.; Planchot, V.; Monsan, P., Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties. *FEBS Lett*, **2000**, 471, (2-3), 219-23.
236. Skov, L. K., *unpublished results*.
237. Quijoch, F. A., Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu Rev Biochem*, **1986**, 55, 287-315.
238. Thoma, J. A.; Rao, G. V.; Brothers, C.; Spradlin, J.; Li, L. H., Subsite mapping of enzymes. Correlation of product patterns with Michaelis parameters and substrate-induced strain. *J Biol Chem*, **1971**, 246, (18), 5621-35.
239. Suganuma, T.; Matsuno, R.; Ohnishi, M.; Hiromi, K., A study of the mechanism of action of Taka-amylase A1 on linear oligosaccharides by product analysis and computer simulation. *J Biochem (Tokyo)*, **1978**, 84, (2), 293-316.
240. Suganuma, T.; Ohnishi, M.; Hiromi, K.; Nagahama, T., Elucidation of the subsite structure of bacterial saccharifying alpha-amylase and its mode of degradation of maltose. *Carbohydr Res*, **1996**, 282, (1), 171-80.

241. Davies, G. J.; Wilson, K. S.; Henrissat, B., Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*, **1997**, 321 (Pt 2), 557-9.
242. Mirza, O.; Skov, L. K.; Remaud-Simeon, M.; Potocki de Montalk, G.; Albenne, C.; Monsan, P.; Gajhede, M., 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*, **2001**, 40, (30), 9032-9.
243. Sarcabal, P.; Remaud-Simeon, M.; Willemot, R.; Potocki de Montalk, G.; Svensson, B.; Monsan, P., Identification of key amino acid residues in *Neisseria polysaccharea* amylosucrase. *FEBS Lett*, **2000**, 474, (1), 33-7.
244. Albenne, C.; de Montalk, G. P.; Monsan, P.; Skov, L. K.; Mirza, O.; Gajhede, M.; Remaud-Simeon, M., Site-directed mutagenesis of key amino acids in the active site of amylosucrase from *Neisseria polysaccharea*. *Biologia, Bratislava*, **2002**, 57, 119-128.
245. Skov, L. K.; Mirza, O.; Sprogoe, D.; Dar, I.; Remaud-Simeon, M.; Albenne, C.; Monsan, P.; Gajhede, M., Oligosaccharide and sucrose complexes of amylosucrase. Structural implications for the polymerase activity. *J Biol Chem*, **2002**, 277, (49), 47741-7.
246. Cunningham, B. C.; Wells, J. A., High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*, **1989**, 244, (4908), 1081-5.
247. Singer, B.; Kusmierk, J. T., Chemical mutagenesis. *Annu Rev Biochem*, **1982**, 51, 655-93.
248. Peng, W.; Shaw, B. R., Accelerated deamination of cytosine residues in UV-induced cyclobutane pyrimidine dimers leads to CC-->TT transitions. *Biochemistry*, **1996**, 35, (31), 10172-81.
249. Pfeifer, G. P.; You, Y. H.; Besaratinia, A., Mutations induced by ultraviolet light. *Mutat Res*, **2005**, 571, (1-2), 19-31.
250. Nguyen, A. W.; Daugherty, P. S., Production of randomly mutated plasmid libraries using mutator strains. *Methods Mol Biol*, **2003**, 231, 39-44.
251. Eckert, K. A.; Kunkel, T. A., High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res*, **1990**, 18, (13), 3739-44.
252. Cline, J.; Braman, J. C.; Hogrefe, H. H., PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res*, **1996**, 24, (18), 3546-51.
253. Tindall, K. R.; Kunkel, T. A., Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry*, **1988**, 27, (16), 6008-13.
254. Lin-Goerke, J. L.; Robbins, D. J.; Burczak, J. D., PCR-based random mutagenesis using manganese and reduced dNTP concentration. *Biotechniques*, **1997**, 23, (3), 409-12.
255. Cadwell, R. C.; Joyce, G. F., Randomization of genes by PCR mutagenesis. *PCR Methods Appl*, **1992**, 2, (1), 28-33.
256. Vartanian, J. P.; Henry, M.; Wain-Hobson, S., Hypermutagenic PCR involving all four transitions and a sizeable proportion of transversions. *Nucleic Acids Res*, **1996**, 24, (14), 2627-31.
257. Zaccolo, M.; Williams, D. M.; Brown, D. M.; Gherardi, E., An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J Mol Biol*, **1996**, 255, (4), 589-603.
258. Stemmer, W. P., Rapid evolution of a protein in vitro by DNA shuffling. *Nature*, **1994**, 370, (6488), 389-91.
259. Cramer, A.; Raillard, S. A.; Bermudez, E.; Stemmer, W. P., DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, **1998**, 391, (6664), 288-91.
260. Aharoni, A.; Thieme, K.; Chiu, C. P.; Buchini, S.; Lairson, L. L.; Chen, H.; Strynadka, N. C.; Wakarchuk, W. W.; Withers, S. G., High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat Methods*, **2006**, 3, (8), 609-14.

261. Oliphant, A. R.; Nussbaum, A. L.; Struhl, K., Cloning of random-sequence oligodeoxynucleotides. *Gene*, **1986**, 44, (2-3), 177-83.
262. Airaksinen, A.; Hovi, T., Modified base compositions at degenerate positions of a mutagenic oligonucleotide enhance randomness in site-saturation mutagenesis. *Nucleic Acids Res*, **1998**, 26, (2), 576-81.
263. Short, J. M., Saturation mutagenesis in directed evolution *Patent US6171820*, **2001**.
264. Delcourt, M.; Blesa, S. Method for massive directed mutagenesis (Biomethodes). *US7202086*, **2001**.
265. Ness, J. E.; Kim, S.; Gottman, A.; Pak, R.; Krebber, A.; Borchert, T. V.; Govindarajan, S.; Mundorff, E. C.; Minshull, J., Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat Biotechnol*, **2002**, 20, (12), 1251-5.
266. Gibbs, M. D.; Nevalainen, K. M.; Bergquist, P. L., Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling. *Gene*, **2001**, 271, (1), 13-20.
267. Zha, D.; Eipper, A.; Reetz, M. T., Assembly of designed oligonucleotides as an efficient method for gene recombination: a new tool in directed evolution. *Chembiochem*, **2003**, 4, (1), 34-9.
268. Fox, R. J.; Davis, S. C.; Mundorff, E. C.; Newman, L. M.; Gavrilovic, V.; Ma, S. K.; Chung, L. M.; Ching, C.; Tam, S.; Muley, S.; Grate, J.; Gruber, J.; Whitman, J. C.; Sheldon, R. A.; Huisman, G. W., Improving catalytic function by ProSAR-driven enzyme evolution. *Nat Biotechnol*, **2007**, 25, (3), 338-44.
269. Yoshikuni, Y.; Ferrin, T. E.; Keasling, J. D., Designed divergent evolution of enzyme function. *Nature*, **2006**, 440, (7087), 1078-82.
270. MacKenzie, C. R.; Wang, Q. P.; Warren, R. A. J.; Withers, S. G., Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J Am Chem Soc*, **1998**, 120, 5583-5584.
271. Okuyama, M.; Mori, H.; Watanabe, K.; Kimura, A.; Chiba, S., α -glucosidase mutant catalyzes " α -glycosynthase"-type reaction. *Biosci Biotechnol Biochem*, **2002**, 66, (4), 928-33.
272. Honda, Y.; Kitaoka, M., The first glycosynthase derived from an inverting glycoside hydrolase. *J Biol Chem*, **2006**, 281, (3), 1426-31.
273. Mayer, C.; Zechel, D. L.; Reid, S. P.; Warren, R. A.; Withers, S. G., The E358S mutant of *Agrobacterium* sp. β -glucosidase is a greatly improved glycosynthase. *FEBS Lett*, **2000**, 466, (1), 40-4.
274. Fairweather, J. K.; Fajjes, M.; Driguez, H.; Planas, A., Specificity studies of bacillus 1,3-1,4-beta-glucanases and application to glycosynthase-catalyzed transglycosylation. *Chembiochem*, **2002**, 3, (9), 866-73.
275. Nashiru, O.; Zechel, D. L.; Stoll, D.; Mohammadzadeh, T.; Warren, R. A.; Withers, S. G., β -Mannosynthase: Synthesis of β -Mannosides with a Mutant β -Mannosidase. *Angew Chem Int Ed Engl*, **2001**, 40, (2), 417-420.
276. Jackeman, D. L.; Withers, S. G., On expanding the repertoire of glycosynthases: mutant β -galactosidases forming β -(1,6)-linkages. *Can j Chem*, **2002**, 80, 866-870.
277. Drone, J.; Feng, H. Y.; Tellier, C.; Hoffman, L.; Tran, V.; Rabiller, C.; Dion, M., *Thermus thermophilus* glycosynthases for the efficient synthesis of galactosyl and glucosyl β -1,3-glycosides. *Eur J Org Chem*, **2005**, 1977-1983.
278. Mullegger, J.; Chen, H. M.; Chan, W. Y.; Reid, S. P.; Jahn, M.; Warren, R. A.; Salleh, H. M.; Withers, S. G., Thermostable glycosynthases and thioglycoligases derived from *Thermotoga maritima* β -glucuronidase. *Chembiochem*, **2006**, 7, (7), 1028-30.
279. Jahn, M.; Marles, J.; Warren, R. A.; Withers, S. G., Thioglycoligases: mutant glycosidases for thioglycoside synthesis. *Angew Chem Int Ed Engl*, **2003**, 42, (3), 352-4.

- 280.** Jahn, M.; Chen, H.; Mullegger, J.; Marles, J.; Warren, R. A.; Withers, S. G., Thioglycosynthases: double mutant glycosidases that serve as scaffolds for thioglycoside synthesis. *Chem Commun (Camb)*, **2004**, (3), 274-5.
- 281.** Moracci, M.; Trincone, A.; Perugino, G.; Ciaramella, M.; Rossi, M., Restoration of the activity of active-site mutants of the hyperthermophilic β -glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles. *Biochemistry*, **1998**, 37, (49), 17262-70.
- 282.** Trincone, A.; Perugino, G.; Rossi, M.; Moracci, M., A novel thermophilic glycosynthase that effects branching glycosylation. *Bioorg Med Chem Lett*, **2000**, 10, (4), 365-8.
- 283.** Perugino, G.; Trincone, A.; Giordano, A.; van der Oost, J.; Kaper, T.; Rossi, M.; Moracci, M., Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic pH. *Biochemistry*, **2003**, 42, (28), 8484-93.
- 284.** Fajjes, M.; Saura-Valls, M.; Perez, X.; Conti, M.; Planas, A., Acceptor-dependent regioselectivity of glycosynthase reactions by *Streptomyces* E383A beta-glucosidase. *Carbohydr Res*, **2006**, 341, (12), 2055-65.
- 285.** Hommalai, G.; Withers, S. G.; Chuenchor, W.; Ketudat Cairns, J. R.; Svasti, J., Enzymatic synthesis of cello-oligosaccharides by rice BGlu1 β -glucosidase glycosynthase mutants. *Glycobiology*, **2007**, 17, (7), 744-753.
- 286.** Vaughan, M. D.; Johnson, K.; DeFrees, S.; Tang, X.; Warren, R. A.; Withers, S. G., Glycosynthase-mediated synthesis of glycosphingolipids. *J Am Chem Soc*, **2006**, 128, (19), 6300-1.
- 287.** Fort, S.; Boyer, V.; Greffe, L.; Davies, G. J.; Moroz, O.; Christiansen, L.; Schülein, M.; Cottaz, S.; Driguez, H., Highly Efficient Synthesis of (1-4)-Oligo- and -Polysaccharides Using a Mutant Cellulase *J. Am. Chem. Soc*, **2000**, 122, (23), 5429-5437.
- 288.** Blanchard, S.; Armand, S.; Couthino, P.; Patkar, S.; Vind, J.; Samain, E.; Driguez, H.; Cottaz, S., Unexpected regioselectivity of *Humicola insolens* Cel7B glycosynthase mutants. *Carbohydr Res*, **2007**, 342, (5), 710-6.
- 289.** Kim, Y. W.; Fox, D. T.; Hekmat, O.; Kantner, T.; McIntosh, L. P.; Warren, R. A.; Withers, S. G., Glycosynthase-based synthesis of xylo-oligosaccharides using an engineered retaining xylanase from *Cellulomonas fimi*. *Org Biomol Chem*, **2006**, 4, 2025-2032.
- 290.** Sugimura, M.; Nishimoto, M.; Kitaoka, M., Characterization of glycosynthase mutants derived from glycoside hydrolase family 10 xylanases. *Biosci Biotechnol Biochem*, **2006**, 70, (5), 1210-7.
- 291.** Malet, C.; Planas, A., From β -glucanase to β -glucansynthase: glycosyl transfer to α -glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. *FEBS Lett*, **1998**, 440, (1-2), 208-12.
- 292.** van Lieshout, J.; Fajjes, M.; Nieto, J.; van der Oost, J.; Planas, A., Hydrolase and glycosynthase activity of endo-1,3- β -glucanase from the thermophile *Pyrococcus furiosus*. *Archaea*, **2004**, 1, (4), 285-92.
- 293.** Piens, K.; Henriksson, A. M.; Gullfot, F.; Lopez, M.; Faure, R.; Ibatullin, F. M.; Teeri, T. T.; Driguez, H.; Brumer, H., Glycosynthase activity of hybrid aspen xyloglucan endo-transglycosylase PttXET16-34 nucleophile mutants. *Org Biomol Chem*, **2007**, 5, (24), 3971-8.
- 294.** Hrmova, M.; Imai, T.; Rutten, S. J.; Fairweather, J. K.; Pelosi, L.; Bulone, V.; Driguez, H.; Fincher, G. B., Mutated barley (1,3)- β -D-glucan endohydrolases synthesize crystalline (1,3)- β -D-glucans. *J Biol Chem*, **2002**, 277, (33), 30102-11.
- 295.** Fairweather, J. K.; Hrmova, M.; Rutten, S. J.; Fincher, G. B.; Driguez, H., Synthesis of complex oligosaccharides by using a mutated (1,3)- β -D-glucan endohydrolase from barley. *Chemistry*, **2003**, 9, (11), 2603-10.

- 296.** Jahn, M.; Stoll, D.; Warren, R. A.; Szabo, L.; Singh, P.; Gilbert, H. J.; Ducros, V. M.; Davies, G. J.; Withers, S. G., Expansion of the glycosynthase repertoire to produce defined manno-oligosaccharides. *Chem Commun (Camb)*, **2003**, (12), 1327-9.
- 297.** Kim, Y. W.; Lovering, A. L.; Chen, H.; Kantner, T.; McIntosh, L. P.; Strynadka, N. C.; Withers, S. G., Expanding the thioglycoligase strategy to the synthesis of α -linked thioglycosides allows structural investigation of the parent enzyme/substrate complex. *J Am Chem Soc*, **2006**, 128, (7), 2202-3.
- 298.** Ben-David, A.; Bravman, T.; Balazs, Y. S.; Czjzek, M.; Schomburg, D.; Shoham, G.; Shoham, Y., Glycosynthase activity of *Geobacillus stearothermophilus* GH52 β -xylosidase: efficient synthesis of xylooligosaccharides from α -D-xylopyranosyl fluoride through a conjugated reaction. *Chembiochem*, **2007**, 8, (17), 2145-51.
- 299.** Kim, Y. W.; Lee, S. S.; Warren, R. A.; Withers, S. G., Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire. *J Biol Chem*, **2004**, 279, (41), 42787-93.
- 300.** Mayer, C.; Jakeman, D. L.; Mah, M.; Karjala, G.; Gal, L.; Warren, R. A.; Withers, S. G., Directed evolution of new glycosynthases from *Agrobacterium* β -glucosidase: a general screen to detect enzymes for oligosaccharide synthesis. *Chem Biol*, **2001**, 8, (5), 437-43.
- 301.** Lin, H.; Tao, H.; Cornish, V. W., Directed evolution of a glycosynthase via chemical complementation. *J Am Chem Soc*, **2004**, 126, (46), 15051-9.
- 302.** Tao, H.; Peralta-Yahya, P.; Lin, H.; Cornish, V. W., Optimized design and synthesis of chemical dimerizer substrates for detection of glycosynthase activity via chemical complementation. *Bioorg Med Chem*, **2006**, 14, (20), 6940-53.
- 303.** Feng, H. Y.; Drone, J.; Hoffmann, L.; Tran, V.; Tellier, C.; Rabiller, C.; Dion, M., Converting a β -glycosidase into a β -transglycosidase by directed evolution. *J Biol Chem*, **2005**, 280, (44), 37088-97.
- 304.** Osanjo, G.; Dion, M.; Drone, J.; Solleux, C.; Tran, V.; Rabiller, C.; Tellier, C., Directed evolution of the α -L-fucosidase from *Thermotoga maritima* into an α -L-transfucosidase. *Biochemistry*, **2007**, 46, (4), 1022-33.
- 305.** Zhang, J. H.; Dawes, G.; Stemmer, W. P., Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci U S A*, **1997**, 94, (9), 4504-9.
- 306.** Parikh, M. R.; Matsumura, I., Site-saturation mutagenesis is more efficient than DNA shuffling for the directed evolution of β -fucosidase from β -galactosidase. *J Mol Biol*, **2005**, 352, (3), 621-8.
- 307.** Geddie, M. L.; Matsumura, I., Rapid evolution of β -glucuronidase specificity by saturation mutagenesis of an active site loop. *J Biol Chem*, **2004**, 279, (25), 26462-8.
- 308.** Hancock, S. M.; Corbett, K.; Fordham-Skelton, A. P.; Gatehouse, J. A.; Davis, B. G., Developing promiscuous glycosidases for glycoside synthesis: residues W433 and E432 in *Sulfolobus solfataricus* β -glycosidase are important glucoside- and galactoside-specificity determinants. *Chembiochem*, **2005**, 6, (5), 866-75.
- 309.** Lee, S.; Mouri, Y.; Minoda, M.; Oneda, H.; Inouye, K., Comparison of the wild-type alpha-amylase and its variant enzymes in *Bacillus amyloliquefaciens* in activity and thermal stability, and insights into engineering the thermal stability of bacillus alpha-amylase. *J Biochem*, **2006**, 139, (6), 1007-15.
- 310.** Wong, D. W.; Batt, S. B.; Lee, C. C.; Robertson, G. H., High-activity barley α -amylase by directed evolution. *Protein J*, **2004**, 23, (7), 453-60.
- 311.** Bessler, C.; Schmitt, J.; Maurer, K. H.; Schmid, R. D., Directed evolution of a bacterial α -amylase: toward enhanced pH-performance and higher specific activity. *Protein Sci*, **2003**, 12, (10), 2141-9.

- 312.** Kim, Y. W.; Choi, J. H.; Kim, J. W.; Park, C.; Kim, J. W.; Cha, H.; Lee, S. B.; Oh, B. H.; Moon, T. W.; Park, K. H., Directed evolution of *Thermus* maltogenic amylase toward enhanced thermal resistance. *Appl Environ Microbiol*, **2003**, 69, (8), 4866-74.
- 313.** Fujii, K.; Minagawa, H.; Terada, Y.; Takaha, T.; Kuriki, T.; Shimada, J.; Kaneko, H., Use of random and saturation mutageneses to improve the properties of *Thermus aquaticus* amyloamylase for efficient production of cycloamyloses. *Appl Environ Microbiol*, **2005**, 71, (10), 5823-7.
- 314.** van der Veen, B. A.; Potocki-Veronese, G.; Albenne, C.; Joucla, G.; Monsan, P.; Remaud-Simeon, M., Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett*, **2004**, 560, (1-3), 91-7.
- 315.** van der Veen, B. A.; Skov, L. K.; Potocki-Veronese, G.; Gajhede, M.; Monsan, P.; Remaud-Simeon, M., Increased amylosucrase activity and specificity, and identification of regions important for activity, specificity and stability through molecular evolution. *Febs J*, **2006**, 273, (4), 673-81.
- 316.** Emond, S. Evolution dirigée de l'amylosaccharase par la technique MUTAGEN pour l'obtention de variants thermostables. Thèse de doctorat en biocatalyse, Institut National des Sciences Appliquées (INSA), Toulouse, **2006**.
- 317.** Emond, S.; André, I.; Jaziri, K.; Potocki-Véonèse, G.; Mondon, P.; Bouayadi, K.; Kharrat, H.; Monsan, P.; Remaud-Siméon, M., Combinatorial engineering to enhance amylosucrase thermostability. *unpublished results*.
- 318.** Remaud-Simeon, M.; Willemot, R. M.; Sarcabal, P.; Potocki de Montalk, G.; Monsan, P., Glucansucrases: molecular engineering and oligosaccharide synthesis. *J.Mol.Catal.B. Enzym.*, **2000**, 10, 117-128.
- 319.** Monchois, V.; Vignon, M.; Russell, R. R., Mutagenesis of asp-569 of glucosyltransferase I glucansucrase modulates glucan and oligosaccharide synthesis. *Appl Environ Microbiol*, **2000**, 66, (5), 1923-7.
- 320.** Kralj, S.; van Geel-Schutten, I. G.; Faber, E. J.; van der Maarel, M. J.; Dijkhuizen, L., Rational transformation of *Lactobacillus reuteri* 121 reuteransucrase into a dextransucrase. *Biochemistry*, **2005**, 44, (25), 9206-16.
- 321.** Kralj, S.; Eeuwema, W.; Eckhardt, T. H.; Dijkhuizen, L., Role of asparagine 1134 in glucosidic bond and transglycosylation specificity of reuteransucrase from *Lactobacillus reuteri* 121. *Febs J*, **2006**, 273, (16), 3735-42.
- 322.** Fabre, E. Caractérisation de la dextrane-saccharase DSR-E de *Leuconostoc mesenteroides* NRRL-B 1299 et application à la synthèse de composés prébiotiques. Thèse de doctorat en biocatalyse, Institut National des Sciences Appliquées (INSA), Toulouse, **2004**.
- 323.** Joucla, G. Caractérisation de l'alternane-saccharase de *Leuconostoc mesenteroides* NRRL B-1355 : Approche rationnelle et aléatoire pour la conception de nouvelles glucane-saccharases. Thèse de doctorat en biocatalyse, Institut National des Sciences Appliquées (INSA), Toulouse, **2003**.
- 324.** Joucla, G.; Pizzut, S.; Monsan, P.; Remaud-Simeon, M., Construction of a fully active truncated alternansucrase partially deleted of its carboxy-terminal domain. *FEBS Lett*, **2006**, 580, (3), 763-8.
- 325.** Moulis, C. Ingénierie rationnelle de la dextrane-saccharase DSR-S: Compréhension du mécanisme de polymérisation pour la synthèse de dextranses de taille contrôlée. Thèse de doctorat en biocatalyse, Institut National des Sciences Appliquées (INSA), Toulouse, **2006**.
- 326.** Moulis, C.; Arcache, A.; Escalier, P. C.; Rinaudo, M.; Monsan, P.; Remaud-Simeon, M.; Potocki-Veronese, G., High-level production and purification of a fully active

recombinant dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *FEMS Microbiol Lett*, **2006**, 261, (2), 203-10.

**Chapitre II : Synthèse enzymatique de
disaccharides intervenant dans la
composition de polysaccharides
bactériens**

Le premier objectif de cette thèse a été d'évaluer le potentiel de quatre glucane-saccharases recombinantes de spécificités différentes à glucosyler les accepteurs cibles. Ces travaux préliminaires étaient destinés à sélectionner les enzymes les plus prometteuses pour la synthèse des disaccharides cibles, en vue de les optimiser par ingénierie moléculaire.

Les quatre glucane-saccharases, utilisées sous leur forme recombinante dans cette étude préliminaire, ont été décrites en page 89. Rappelons, cependant, leurs noms et spécificités :

- **GBD-CD2, forme tronquée de DSR-E** (liaison α , 1-2)
- **ASR C-APY del, forme tronquée de l'ASR** (liaisons α , 1-3 et α , 1-6 alternées)
- **DSR-S vardel Δ 4N, forme tronquée de DSR-S** (liaison α , 1-6)
- **ASNP, amylosaccharase** (liaison α , 1-4)

Pour compléter cette étude, il nous a également semblé judicieux de tester deux mutants de **DSR-S vardel Δ 4N**, dépourvus d'activité polymérase :

- **DSR-S vardel Δ 4N SEV663YDA**
- **DSR-S vardel Δ 4N SEV663NNS**

Les trois accepteurs testés dans cette étude sont :

- **la N-acétyl-D-glucosamine** (accepteur non protégé)
- **le L-rhamnose** (accepteur non protégé)
- **l' α -methyl- L-rhamnose** (accepteur protégé en position anomérique)

Il s'agissait plus particulièrement de déterminer :

- l'aptitude de ces enzymes à reconnaître les accepteurs cibles (taux de conversion de l'accepteur)
- leur capacité à former un produit mono-glucosylé,
- la liaison osidique formée entre l'accepteur et l'unité D-glucosyle. Pour ce faire, les produits de glucosylation les plus abondants ont été purifiés et caractérisés par spectroscopie RMN et spectrométrie de masse.

Ces travaux font l'objet de la première publication présentée dans ce chapitre.

ENZYMATIC SYNTHESIS OF DISACCHARIDE MOTIVES ENTERING IN THE COMPOSITION OF BACTERIAL POLYSACCHARIDES BY USE OF GLUCANSUCRASES

Elise Champion^{a,b,c}, Sandrine Morel^{a,b,c}, Laurence Mulard^e, Pierre Monsan^{a,b,c,d}, Magali Remaud-Siméon^{a,b,c}, and Isabelle André^{a,b,c}.

^a Université de Toulouse; INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

^b CNRS, UMR5504, F-31400 Toulouse, France

^c INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

^d Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris

^e Unité de Chimie des Biomolécules, URA CNRS 2128, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris cedex 15, France

A. Summary

Transglucosylation reactions using sucrose as glucosyl donor and either *N*-acetyl-D-glucosamine, L-rhamnose or methyl α -L-rhamnopyranoside as acceptors were carried out with recombinant glucansucrases from family 70 or 13 of glycoside-hydrolases. Depending on the enzyme specificity, various carbohydrate structures were synthesized and characterized including α -D-glucopyranosyl-(1 \rightarrow 6)-*N*-acetyl-D-glucosamine, α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acetyl-D-glucosamine, α -D-glucopyranosyl-(1 \rightarrow 1)- β -L-rhamnopyranoside, α -D-glucopyranosyl-(1 \rightarrow 4)- α -D-glucopyranosyl-(1 \rightarrow 1)- β -L-rhamnopyranoside, methyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranoside and methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside. Several of the obtained disaccharides enter in the composition of surface polysaccharides of pathogenic bacteria, among which *Shigella flexneri*, *Escherichia coli*, *Streptococcus pneumoniae*, *Yersinia kristensenii* or *Halomonas pantelleriensis*. Our results outline the potential of glucansucrases in the chemo-enzymatic synthesis of complex carbohydrates of biological interest, whose chemical synthesis may be seen as a limitation.

B. Introduction

Two categories of natural enzymes catalyze glycosyl transfer onto sugar acceptors: **glycosyltransferases** of the Leloir pathway, which use expensive nucleotide activated sugars as glycosyl donor, and **transglycosidases**. For both types, reactivity towards acceptors is strictly controlled by the enzyme selectivity and specificity. Among transglycosidases, glucansucrases (GS, EC 2.4.1), which are found in glycoside hydrolases (GH) family 13 and 70 according to the CAZY classification¹, are very attractive to synthesize oligosaccharides. They naturally catalyze **glucan synthesis** and the concomitant release of fructose from sucrose donor, a largely available and low cost substrate. The type, number and organization of the glucosidic linkages displayed in the glucan polymer is strongly dependant on glucansucrase specificities²⁻⁷. An interesting particularity of glucansucrases is their ability to catalyze glycosylation reactions of non-natural acceptors introduced in the reaction medium to give rise to gluco-derivatives (Figure 1). Glucansucrases are indeed able to transfer glucosyl residues onto a large variety of acceptors, including sugar⁸⁻¹⁵ and non-sugar^{8, 13, 16-21} molecules. In this process, acceptor recognition, glycosylation rate and resulting products are also strongly dependant on glucansucrase stereo- and regio-specificities.

In the present study, we have investigated the glycosylation of commercially available *N*-acetyl-D-glucosamine, and L-rhamnose, as well as of methyl α -L rhamnopyranoside^{22, 23} prepared in one step from the later, by use of several recombinant or native glucansucrases, displaying distinct specificities (**TABLE 1**). The objective is to challenge the synthesis of glycosyl derivatives of these molecules that could enter into the chemo-enzymatic synthesis of oligosaccharide motives mimicking key elements of surface polysaccharides of pathogenic bacteria, such as *Shigella flexneri*, *Streptococcus pneumoniae*, *Escherichia coli*, *Yersinia kristensenii* and *Halomonas pantelleriensis*. In addition to some fundamental interest in the role that bacterial surface polysaccharide play in pathogenicity, a number of recent studies have shown that chemically and immunologically defined carbohydrate haptens could be part of semi-synthetic glycoconjugate immunogens, also termed third generation polysaccharide vaccines.²⁴ Interestingly, the potency of this alternative to conventional polysaccharide-based vaccines has recently been demonstrated in human^{25, 26}. However, chemical synthesis often remains tedious and introducing an enzymatic step may greatly improve the process. The acceptor reaction products synthesized with the selected enzymes were isolated and characterized. Their potential interest for the synthesis of disaccharide motives which could serve as starting building blocks for the synthesis of antigenic haptens is discussed.

C. Results and Discussion

Glucansucrases are α -retaining transglucosidases that follow a double displacement retaining mechanism, in which a β -D-glucosyl enzyme covalent intermediate is first formed from sucrose substrate²⁷⁻²⁹. In a second step, the glucosyl moiety is transferred to an acceptor which, depending on the conditions of reaction, can be: (1) either water leading to the production of glucose, (2) fructose to form sucrose isomers, (3) the glucose moiety released from sucrose hydrolysis to form soluble oligosaccharides and α -glucan polymers, or (4) an exogenous hydroxylated acceptor, when the latter is well recognized. Depending on the enzyme regiospecificity, distinct types of glucosidic linkage are found in the polymer formed or in the non-natural glucosyl derivative (FIGURE 32).

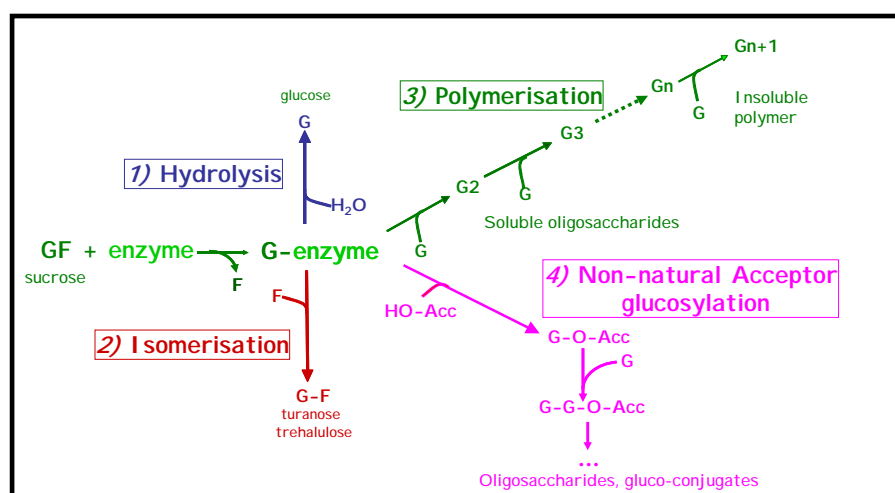


FIGURE 32| Reactions catalyzed by glucansucrases.

G: glucosyl; F: fructosyl or fructose; AccOH: non-natural acceptor

In order to take advantage of the ability of glucansucrases to glucosylate non-natural acceptors in a regio- and stereo-specific manner, six recombinant glucansucrases were tested for the glucosylation of *N*-acetyl-D-glucosamine (D-GlcpNAc), L-rhamnose (L-Rhap) and methyl α -L-rhamnopyranoside (α -L-RhapOMe). These six enzymes are namely: amylosucrase (AS)⁷, an engineered alternansucrase (ASR-C-APY del)³⁰, three engineered dextransucrases (a truncated form of DSR-S dextransucrase³¹ and two mutants of this form³²) and one α -1,2 branching enzyme, engineered from DSR-E dextransucrase³³ (TABLE 1). Notably, mutant enzymes DSR-S vardel Δ 4N SEV663YDA and DSR-S vardel Δ 4N SEV663NNS carry mutations in the region downstream the second aspartic acid of the catalytic triad (known to participate in the distortion of the glucosyl-enzyme and in the positioning of the acceptor), which could potentially alter the enzyme stereo- and regio-specificity³². As shown in TABLE 1, these enzymes catalyze the synthesis of diverse types of glucosidic linkage and were chosen on purpose to attempt the synthesis of structurally distinct disaccharides.

TABLE 1| Origin and specificity of the recombinant glucansucrases selected for the glucosylation of the target acceptors: D-Glc_pNAc, L-Rhap and α-L-RhapOMe.

| Recombinant Enzyme | Origin | EC | GH Number | Type of glucosidic linkages | Ref. |
|--|--|-----------|-----------|--------------------------------------|------|
| DSR-S vardel Δ4N | Truncated form of DSR-S dextranucrase from <i>Leuconostoc. mesenteroides</i> NRRL B-512F | 2.4.1.5 | GH70 | α-1,6 | 31 |
| DSR-S vardel Δ4N SEV663YDA | Mutant of DSR-S vardel Δ4N | 2.4.1.5 | GH70 | α-1,6 | 32 |
| DSR-S vardel Δ4N SEV663NNS | Mutant of DSR-S vardel Δ4N | 2.4.1.5 | GH70 | α-1,6 | 32 |
| ASR-C-APY del | Truncated form of ASR alternansucrase from <i>L. mesenteroides</i> NRRL B-1355 | 2.4.1.140 | GH70 | α-1,6 and α-1,3 (mainly alternating) | 30 |
| DSR-E GBD-CD2 (α-1,2 branching enzyme) | Truncated form of DSR-E dextranucrase from <i>L. mesenteroides</i> NRRL B-1299 | 2.4.1-XX | GH70 | α-1,2 | 33 |
| AS | Recombinant form of <i>Neisseria polysaccharea</i> amylosucrase | 2.4.1.4 | GH13 | α-1,4 | 7 |

C.1. Glucosylation of N-acetyl-D-glucosamine

Enzyme screening. Acceptor reactions were first carried out using sucrose as donor and D-Glc_pNAc as acceptor in a molar ratio of 1. As shown in **TABLE 2**, only traces of acceptor products were synthesized using GBD-CD2. For the five remaining enzymes which were tested, D-Glc_pNAc conversion rates were comprised between 6 and 64 %. Using four different HPLC column systems, the chromatographic profiles of the acceptor reaction products synthesized by DSR-S vardel Δ4N, its variants (Δ4N SEV663YDA and Δ4N SEV663NNS), ASR and GBD-CD2 were all identical (**FIGURE 33**). It can thus be assumed that all tested GH70 glucansucrases produced the same compound, herein called **P1**. Interestingly, the main acceptor reaction product obtained by action of AS, product **P2**, displayed a retention time differing from that of **P1**, indicating that a structurally distinct product was synthesized by transglucosylation (**FIGURE 34**).

TABLE 2| Transglucosylation of D-Glc_pNAc using different recombinant glucansucrases at Acceptor:Sucrose ratio of 1:1 (146 mM).

The conversion rate of the acceptor was calculated using the formula: $([\text{Acceptor}]_{\text{initial}} - [\text{Acceptor}]_{t_{24h}}) / [\text{Acceptor}]_{\text{initial}} \times 100$.
The percentage of Monoglucosylated Acceptor was determined from the ratio: $[\text{Glc-Acceptor}]_{t_{24h}} / [\text{Sucrose consumed}]$.
At final time, sucrose is fully consumed.

| | Recombinant enzyme | | | | | |
|---|--------------------|---------------|---------------|-----------|---------|----|
| | Δ4N | Δ4N SEV663YDA | Δ4N SEV663NNS | ASR | GBD-CD2 | AS |
| Conversion Rate (%) | 26% | 54% | 64% | 14% | <1% | 6% |
| % Monoglucosylated D-Glc _p NAc | 24% | 54% | 61% | 13% | <1% | 4% |
| Monoglucosylated D-Glc _p NAc product | P1 | | | P2 | | |

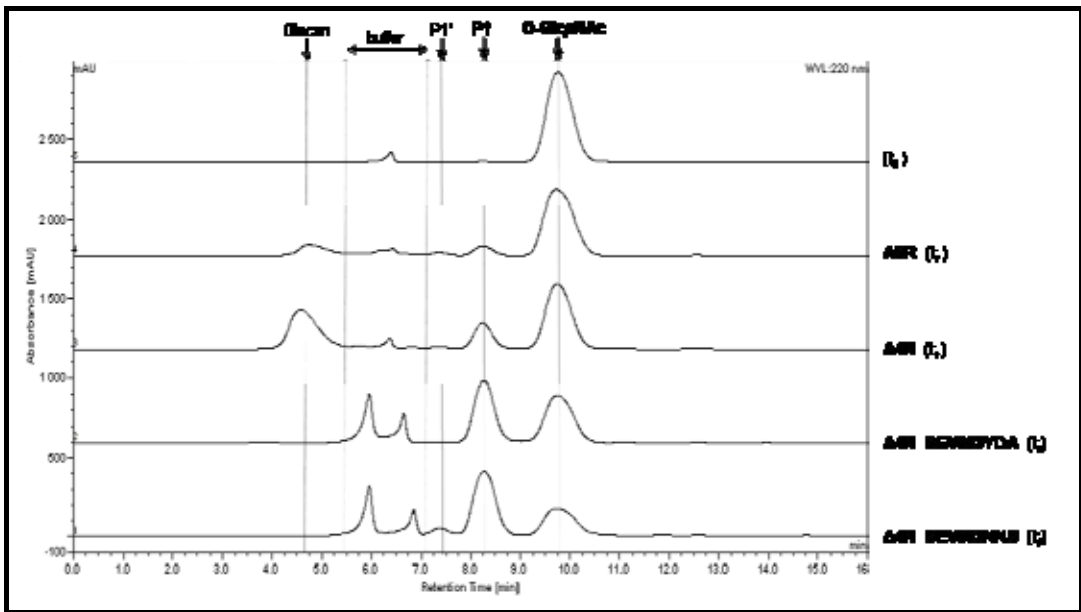


FIGURE 33| Analysis of D-GlcpNAc glucosylation by 4 different GH70 glucansucrases. Comparison of HPLC chromatograms (with $UV_{\lambda=220nm}$ detection) before and after total sucrose consumption. (t_0): initial time; (t_i): final time=24h; P1': Uncharacterized product.

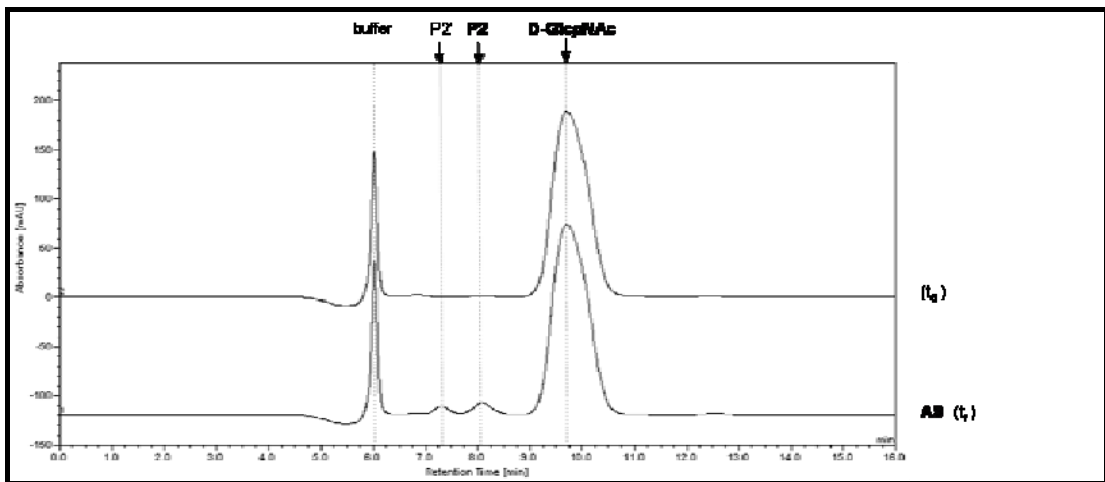


FIGURE 34| Analysis of D-GlcpNAc glucosylation by GH13 glucansucrase, amylosucrase from *N. polysaccharia*.

Comparison of HPLC chromatograms (with $UV_{\lambda=220nm}$ detection) after total sucrose consumption. (t_0): initial time; (t_i): final time=24h; P2': Uncharacterized product.

Structural characterization. The structure of **P1**, which was synthesized with $\Delta 4N$ SEV663YDA mutant was analyzed by HRMS and NMR (details given in Experimental section).

HRMS analysis of **P1** indicated a molecular weight of 383 g/mol was found, corresponding to a monoglucosylated form of D-GlcpNAc. The 1H NMR spectrum of **P1** shows two doublets at 4.87 and 4.88 ppm, assigned to the anomeric proton of the D-glucosyl residue, and two doublets at 4.66 and 5.13 ppm, which by comparison with the D-GlcpNAc spectrum were assigned to the anomeric proton of β - and α -D-GlcpNAc, respectively. Signal integration indicates a disaccharide structure in good agreement with MS data. Moreover, both $H-1_{Glc}$ doublets exhibit weak J_{1-2} coupling

constants (3.4 and 3.5 Hz) which are characteristic of an α -D-glucosidic linkage. Besides, H-1_{Glc}-C-6_{GlcNAc} and H-6_{GlcNAc}-C-1_{Glc} cross peaks in the HMBC spectrum ascertain the presence of an α -(1→6) linkage. All ¹H and ¹³C signal assignments are listed in TABLE 3.

TABLE 3] ¹H and ¹³C chemical shifts (ppm, D₂O) and ³J coupling constants (Hz) of P1 obtained by action of Δ 4N SEV663YDA on D-GlcpNAc.

| Compound | Residue | δ (ppm) | Carbohydrate Ring | | | | | | N-Acetyl group | |
|---------------------------------|---------------------|-----------------|------------------------------|-------|-------|-------|-------|-----------|----------------|-----------------|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | C | CH ₃ |
| P1 (β anomer) | α -D-Glcp | ¹ H | 4.87 | 3.46 | 3.67 | 3.32 | 3.65 | 3.67 | | |
| | | ¹³ C | 97.97 | 71.45 | 73.03 | 69.47 | 71.78 | 60.43 | | |
| | β -D-GlcpNAc | ¹ H | 4.66 | 3.61 | 3.49 | 3.60 | 3.58 | 3.92-3.66 | | 1.96 |
| | | | (J _{1,2} : 11.2 Hz) | | | | | | | |
| | | ¹³ C | 95.05 | 56.58 | 74.08 | 69.97 | 74.30 | 65.65 | 174.47 | 22.16 |
| | | | | | | | | | | |
| P1 (α anomer) | α -D-Glcp | ¹ H | 4.88 | 3.46 | 3.67 | 3.32 | 3.65 | 3.67 | | |
| | | ¹³ C | 97.91 | 71.49 | 73.03 | 69.47 | 71.78 | 60.43 | | |
| | α -D-GlcpNAc | ¹ H | 5.13 | 3.82 | 3.48 | 3.92 | 3.67 | 3.92-3.66 | | 1.96 |
| | | | (J _{1,2} : 4.8 Hz) | | | | | | | |
| | | ¹³ C | 90.91 | 53.98 | 69.64 | 70.15 | 70.96 | 65.73 | 174.47 | 21.88 |
| | | | | | | | | | | |

From these assignments, it can be concluded that **P1** is a new disaccharide corresponding to α -D-glucopyranosyl-(1→6)-*N*-acetyl-D-glucosamine. (**FIGURE 35**)

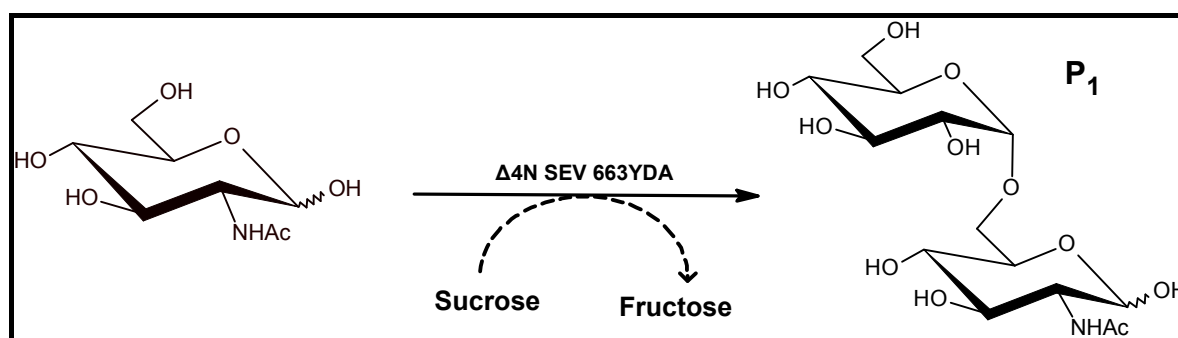


FIGURE 35] Structure of disaccharide P1 obtained by the Δ 4N SEV663YDA-mediated glycosylation of *N*-acetyl-D-glucosamine.

Analogously, with a molecular weight of 383 g/mol determined by HRMS analysis, **P2** is also a disaccharide. Moreover, NMR analysis (**TABLE 4**) indicates two major differences in the ¹³C

spectra of **P2** (TABLE 4) in comparison to that of **P1** (TABLE 3). Differences involve δ_{C-6} , lowered by 4.5 ppm, and δ_{C-6} which is increased by at least 7.2 ppm. Taking into account the H-1_{Glc}-C-4_{Glc_pNAc} and H-4_{Glc_pNAc}-C-1_{Glc} cross peaks in the HMBC spectrum, all data converge to suggest that **P2** corresponds to α -D-glucopyranosyl-(1→4)-*N*-acetyl-D-glucosamine (FIGURE 36).

TABLE 4| ¹H and ¹³C chemical shifts (ppm, D₂O) and ³J coupling constants (Hz) of **P2** resulting from AS-mediated glucosylation of D-Glc_pNAc

| Compounds | Residue | δ (ppm) | Carbohydrate Ring | | | | | | N-Acetyl group | |
|---------------------------------|---------------------|-------------------|-------------------|-------|-------|-------|-------|-------|----------------|-------|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | C | CH3 |
| P2 (β anomer) | α -D-Glcp | ¹ H | 5.39 | 3.56 | 3.66 | 3.40 | 3.70 | 3.82 | | |
| | | ¹³ C | 99.87 | 72.06 | 73.27 | 69.78 | 73.14 | 60.97 | | |
| | β -D-GlcpNAc | ¹ H | 4.70 | 3.68 | 3.79 | 3.69 | 3.56 | 3.82 | | 2.02 |
| | | ¹³ C | 95.25 | 57.09 | 74.79 | 77.22 | 75.06 | 61.00 | 175.22 | 22.62 |
| P2 (α anomer) | α -D-Glcp | ¹ H | 5.40 | 3.56 | 3.66 | 3.40 | 3.70 | 3.82 | | |
| | | ¹³ C | 100.15 | 72.17 | 73.30 | 69.78 | 73.14 | 60.97 | | |
| | α -D-GlcpNAc | ¹ H | 5.18 | 3.89 | 4.00 | 3.66 | 3.94 | 3.82 | | 2.02 |
| | | ¹³ C | 91.13 | 54.36 | 71.62 | 77.94 | 70.57 | 61.00 | 174.98 | 22.34 |

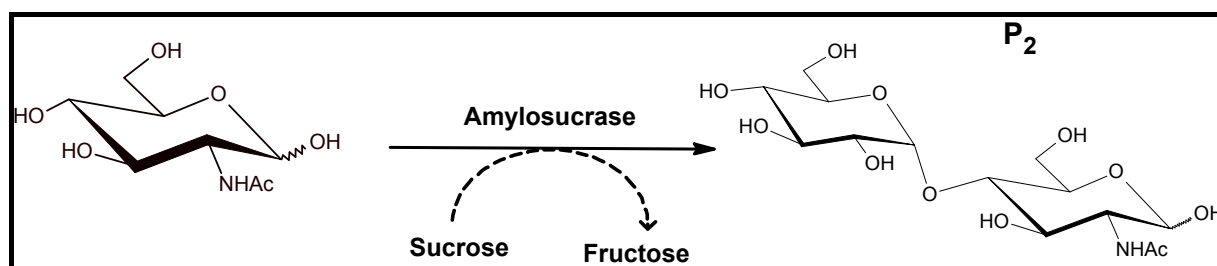


FIGURE 36| Structure of disaccharide **P2** obtained by AS-mediated glucosylation of D-Glc_pNAc

D-Glc_pNAc glucosylation by the native alternansucrase from *L. mesenteroides* NRRL B-1355 was previously reported.⁸ The reaction was very low yielding and the product structures were not determined. We showed here that all tested GH family 70 enzymes produced the same disaccharide **P1** upon D-Glc_pNAc glucosylation. These enzymes thus conserve their same natural regiospecificity towards D-Glc_pNAc. DSR-S vardel Δ 4N and its variants were the most efficient enzymes for the synthesis of **P1**. Mutations introduced downstream the second aspartic acids of the catalytic triad of the enzymes did not alter the enzyme regiospecificity, but had a beneficial

effect on the conversion rate which increased by more than 2 fold for both mutants ($\Delta 4N$ SEV663YDA and $\Delta 4N$ SEV663NNS) compared to $\Delta 4N$. Located close to the catalytic residues, the mutations introduced may favour binding of the non natural acceptor, indicating that the engineering of glucansucrases can greatly enhance the ability of these enzymes for the synthesis of novel oligosaccharides.

C.2. Glucosylation of L-rhamnose

Enzyme screening. Acceptor reactions were carried using a L-rhamnose: sucrose ratio of 1 (TABLE 5). The best conversion rate of L-rhamnose was obtained with AS, resulting in two reaction products (P3 and P4). Family 70 glucansucrases also converted L-rhamnose. Interestingly, they synthesized several products, distinct from those obtained with AS. The yield of each independant produced oligosaccharide was very low, thus preventing further characterization. Therefore, we focused on the characterization of P3 and P4.

TABLE 5| Transglucosylation of L-rhamnose using different recombinant glucansucrases at Acceptor:Sucrose ratio of 1:1 (146 mM).

The conversion rate of the acceptor was calculated using the formula $([Acceptor]_{initial} - [Acceptor]_{t24h})/[Acceptor]_{initial} \times 100$. The percentage of Monoglucosylated Acceptor was determined by the ratio $[Glc-Acceptor]_{t24h} / [sucrose\ consumed]$

| | <i>Recombinant enzyme</i> | | | | | |
|-------------------------|---------------------------|--------------------------|--------------------------|-----|---------|-----------|
| | $\Delta 4N$ | $\Delta 4N$ SEV663YDA | $\Delta 4N$ SEV663NNS | ASR | GBD-CD2 | AS |
| Conversion Rate (%) | <1% | <10% | 1% | 6% | <1% | 12% |
| Novel observed products | 0 | 5 | 0 | 4 | 0 | 2 (P3+P4) |

Optimization of P3 and P4 synthesis by AS. To improve P3 and P4 synthesis, the influence of Acceptor:Sucrose ratio on the conversion rate was first studied at two sucrose concentrations (146 mM and 292 mM) (TABLE 6). In both cases, increasing the acceptor concentration favours its glucosylation at the cost of polymer formation, and yields higher amount of glucosylation products. However, L-rhamnose conversion rate decreases concomitantly. Maintaining similar Acceptor:Sucrose ratio but varying the initial sucrose concentration did not improve significantly the percentage of glucosyl transfer. Therefore, in search for the best compromise between P3 yield and L-rhamnose conversion rate, we retained a Sucrose concentration of 292 mM and an Acceptor:Sucrose ratio of 2.5 for the synthesis.

TABLE 6| Effect of the Acceptor/Sucrose molar ratio (A/S) at sucrose concentrations of 146 and 292 mM. Percentages of Glcp and Rhap transferred were determined at final tinal (24h). At final time, sucrose is fully consumed .

| | A/S Ratio ([S]) = 146 mM | | | | A/S Ratio ([S]) = 292 mM | | | |
|--|--------------------------|------|-------|-------|--------------------------|------|-------------|-------|
| | 1 | 2.5 | 5 | 10 | 0.5 | 1.25 | 2.5 | 5 |
| L-Rhamnose (mM) | 146 | 365 | 730 | 1460 | 146 | 365 | 730 | 1460 |
| P3 (mM) | 6.4 | 12.3 | 18.4 | 25.5 | 7.8 | 13.2 | 26 | 36 |
| % Glcp transferred onto L-Rhamnose to form P3 | 4.4% | 8.4% | 12.6% | 17.5% | 2.7% | 4.5% | 8.9% | 12.3% |
| % Rhap converted into P3 | 4.4% | 3.4% | 2.5% | 1.2% | 2.7% | 3.6% | 3.6% | 2.5% |

Structural characterization of P3 and P4. **P3** and **P4** were separated by preparative HPLC and analyzed by NMR. The 1D and 2D L-Rhamnose spectra were used as reference (data not shown). L- Rhamnose is found in a $^1\text{C}_4$ pyranose ring conformation, as revealed by $J_{1,2}$ coupling constants ($J_{1,2}(\beta)\sim 1\text{ Hz}$) ($^1\text{C}_4$ conformation: $J_{1,2}(\beta)\sim 3.5\text{ Hz}$). Signals at 4.76 ppm and 5 ppm were assigned to the anomeric proton in β and α form respectively.

HRMS data of **P3** indicates a molecular weight of 326 g/mol corresponding to the mass of a glucosylated rhamnose. Two anomeric signals of equal intensity, only, are observed in the ^1H NMR spectrum. Aided by the chemical shifts and $^1J_{\text{C,H}}$ values of the anomeric carbons (**TABLE 7**), signals at 5.19 ppm ($^1J_{\text{C,H}} = 173.9\text{ Hz}$) and 4.78 ppm ($^1J_{\text{C,H}} = 161.3\text{ Hz}$) were assigned to the anomeric protons of a D-glucosyl and a L-rhamnosyl residues in α and β configuration, respectively. No signal corresponding to the α/β -anomerization of a reducing disaccharide could be found, suggesting that the two sugar moieties are linked by their anomeric carbons through an ($\alpha 1\rightarrow\beta 1$) linkage. The glycosidic linkage of the disaccharide was confirmed by HMBC analysis showing two inter-residual scalar couplings, $\text{C-1}_{\text{Rha}}\text{-H-1}_{\text{Glc}}$ and $\text{C-1}_{\text{Glc}}\text{-H-1}_{\text{Rha}}$, respectively.

In agreement with MS data, which strongly suggest that **P4** is a trisaccharide (MW: 488 g/mol) made of two hexoses and a deoxyhexose, the ^1H NMR spectrum of **P4** indicates three doublets of equal intensity in the anomeric region (**TABLE 7**). In support to this assumption, the ^{13}C NMR spectrum showed a single signal in the C-6_{Rha} region (16-17 ppm), and two signals in the C-6_{Glc} region (60-61 ppm), indicating the presence of one L-rhamnosyl ring and two D-glucosyl rings. Besides, the HMBC spectrum showed strong cross-peaks between an H-1_{Glc} and a C-4_{Glc} , and between a H-4_{Glc} and a C-1_{Glc} , respectively, indicating the presence of an α -(1 \rightarrow 4)-linkage between the two glucosyl residues. As for **P3**, NMR analysis revealed that the glucosyl moiety was linked to a L-rhamnosyl residue through an ($\alpha 1\rightarrow\beta 1$) linkage. The whole sets of **P3** and **P4** ^1H and ^{13}C chemical shifts are given in **TABLE 7**.

TABLE 7 | ^1H and ^{13}C chemical shifts (ppm, D_2O) and ^3J coupling constants of the AS-mediated transglucosylation products (P3 and P4) of L-rhamnose

| Compounds | Residue | δ (ppm) | Carbohydrate Ring | | | | | |
|-----------------|------------------------------------|-----------------------------|-----------------------------|-------|-------|-------|-------|------|
| | | | 1 | 2 | 3 | 4 | 5 | 6 |
| P3 | $\alpha\text{-D-Glcp}$ | ^1H | 5.19 | 3.49 | 3.64 | 3.35 | 3.55 | 3.72 |
| | | | (J _{1,2} : 3.7 Hz) | | | | | |
| | ^{13}C | 95.23 | 71.13 | 72.93 | 69.55 | 72.73 | 60.63 | |
| | $\beta\text{-L-Rhap}$ | ^1H | 4.78 | 3.92 | 3.58 | 3.33 | 3.32 | 1.23 |
| | | (J _{1,2} : 1.2 Hz) | | | | | | |
| P4 | $\alpha\text{-D-Glcp}^{\text{II}}$ | ^1H | 5.31 | 3.47 | 3.93 | 3.60 | 3.61 | 3.74 |
| | | | (J _{1,2} : 3.6 Hz) | | | | | |
| | ^{13}C | 99.83 | 71.86 | 73.42 | 69.47 | 73.02 | 60.64 | |
| | $\alpha\text{-D-Glcp}^{\text{I}}$ | ^1H | 5.20 | 3.51 | 3.55 | 3.32 | 3.70 | 3.74 |
| | | | (J _{1,2} : 3.6 Hz) | | | | | |
| | ^{13}C | 95.05 | 71.00 | 72.87 | 76.78 | 71.12 | 60.64 | |
| | $\beta\text{-L-Rhap}$ | ^1H | 4.79 | 3.93 | 3.545 | 3.32 | 3.31 | 1.22 |
| | | | (J _{1,2} : 1.1 Hz) | | | | | |
| ^{13}C | 94.96 | 70.97 | 72.72 | 72.15 | 72.69 | 16.77 | | |

Finally, **P3** and **P4** are new oligosaccharides identified as the $\alpha\text{-D-glucopyranosyl-(1}\rightarrow\text{1)}\beta\text{-L-rhamnopyranoside}$ and $\alpha\text{-D-glucopyranosyl-(1}\rightarrow\text{4)}\text{-}\alpha\text{-D-glucopyranosyl-(1}\rightarrow\text{1)}\text{-}\beta\text{-L-rhamnopyranoside}$, respectively (**FIGURE 37**).

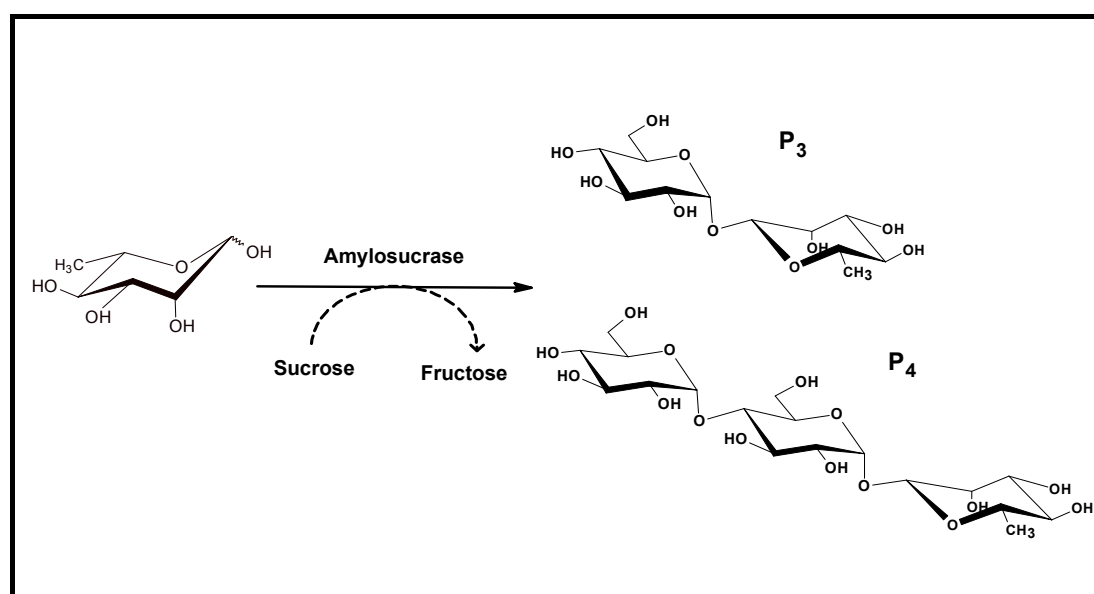


FIGURE 37 | Structures of disaccharide P3 and trisaccharide P4

obtained by AS-mediated glucosylation of L-Rhamnose

C.3 Glucosylation of methyl α -L-Rhamnopyranoside

The acceptor reaction was carried out using an Acceptor:Sucrose ratio of 1 and a sucrose concentration of 146 mM. The acceptor reaction products were analyzed by HPLC on C30 and C18 RP-Fusion columns. The retention times of the products were compared to those of chemically synthesized methyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranoside³⁴ and methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside³⁵ used as reference compounds. Using the family 70 glucansucrases, methyl α -L-rhamnopyranoside conversion rates ranging from 3 to 17% were obtained, indicating poor recognition of this acceptor. The best conversion rate was obtained for ASR, forming preferentially methyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranoside and to a lesser extent, methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside to a lesser extent. Two additional peaks present on the chromatogram, were tentatively assigned to the homologous trisaccharides but were not further characterized. Noteworthy, AS was unable to use methyl α -L-rhamnopyranoside as an acceptor. Indeed, the anomeric hydroxyl group, which was the preferred AS-mediated glucosylation site of L-rhamnose is blocked in methyl α -L-rhamnopyranoside, thus preventing the enzyme from transferring the glucosyl residue onto that or any other hydroxyl groups of α -L-RhapOMe (TABLE 8).

TABLE 8| Transglucosylation of α -L-RhapOMe using different recombinant glucansucrases at Acceptor:Sucrose ratio of 1 (146 mM).

The conversion rate of the acceptor was calculated according to the formula $([\text{Acceptor}]_{\text{initial}} - [\text{Acceptor}]_{t_{24h}})/[\text{Acceptor}]_{\text{initial}} \times 100$.
The percentage of monoglucosylated acceptor was determined by the ratio $[\text{Glc-Acceptor}]_{t_{24h}}/[\text{sucrose consumed}]$.
At final time, fully sucrose was consumed.

| | <i>Recombinant enzyme</i> | | | | | |
|--|---------------------------|--------------------------|--------------------------|-----|---------|-----|
| | Δ 4N | Δ 4N SEV663YDA | Δ 4N SEV663NNS | ASR | GBD-CD2 | AS |
| Conversion Rate (%) | 3% | 5% | 6% | 17% | 5% | <1% |
| % of α -D-Glcp-(1 \rightarrow 3)- α -L-RhapOMe | 0% | 4% | 3% | 3% | 0% | 0% |
| % of α -D-Glcp-(1 \rightarrow 4)- α -L-RhapOMe | <1% | <1% | 3% | 10% | 5% | 0% |

Within the past decades, semi-synthetic glycoconjugate immunogens were proposed as third generation polysaccharide vaccines.³⁶ This alternative to conventional polysaccharide-protein conjugate vaccines involves synthetic carbohydrate haptens that mimic the immunodominant segment of the natural surface polysaccharide of selected pathogens. Among others, oligosaccharides mimicking *S. flexneri* 2a O-antigen have already been synthesized chemically,³⁷

and found immunogenic in an animal model upon conjugation to a protein carrier.²⁴ As an alternative to the chemical route, chemo-enzymatic pathways could be foreseen to access specific carbohydrate structures. In particular, α -D-glucopyranosyl-(1 \rightarrow 6)-*N*-acetyl-D-glucosamine, α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acetyl-D-glucosamine, methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside and methyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranoside are parts of known bacterial polysaccharides (for examples, see **TABLE 9**). Thus, if enzymatically produced, those disaccharides could serve as building blocks in the chemo-enzymatic synthesis of polysaccharide fragments of biological interest. Noteworthy, providing an enzymatic access to these disaccharide motives could, and help to overcome a well known limiting chemical step in the chemical synthesis of carbohydrates, that is the introduction of a 1,2-*cis* glucosidic linkage, which is common to all listed examples.

TABLE 9| Examples of repeating units and/or cores of bacterial surface polysaccharides containing the disaccharide motives synthesized by glucansucrases as demonstrated in this work.

| Target disaccharide | Organism | Structure | Refs |
|--|--------------------------------|--|--------|
| α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc | <i>S. flexneri</i> type 1a | \rightarrow 2)- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 4)]- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>S. flexneri</i> type 1b | \rightarrow 2)- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)-[2Ac]- α -L-Rhap-(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 4)]- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>E. coli</i> O18B1 | \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-[β -D-Glcp-(1 \rightarrow 3)]- α -D-Galp(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 4)]- α -D-GlcpNAc-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow | 39 |
| | <i>Y. kristensenii</i> O:25.35 | \rightarrow 6)- β -D-Glcp-(1 \rightarrow 4)-[α -D-Glcp-(1 \rightarrow 6)- α -D-Galp(1 \rightarrow 3)]- α -L-FucpNAc-(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 4)]- β -D-GlcpNAc-(1 \rightarrow 2)-Gro-(1-P \rightarrow | 40 |
| | <i>H. pantelleriensis</i> | \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- α -D-GlcpNAc-(1 \rightarrow as core fragment | 41 |
| α -D-Glcp-(1 \rightarrow 6)-D-GlcpNAc | <i>S. flexneri</i> type 4a | \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 6)]- β -D-GlcpNAc-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow | 38, 42 |
| | <i>S. flexneri</i> type 4b | \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 6)]- β -D-GlcpNAc-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)-[2Ac]- α -L-Rhap-(1 \rightarrow | 38, 42 |
| | <i>E. coli</i> O18A1 | \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-[β -D-GlcpNAc-(1 \rightarrow 3)]- α -D-Galp(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 6)]- α -D-GlcpNAc-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow | 39 |
| α -D-Glcp-(1 \rightarrow 3)-L-Rhap | <i>S. flexneri</i> type 2b | \rightarrow 2)-[α -D-Glcp-(1 \rightarrow 3)]- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 4)]- α -L-Rhap-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>S. flexneri</i> type 3a | \rightarrow 2)-[α -D-Glcp-(1 \rightarrow 3)]- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)-[2Ac]- α -L-Rhap-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>S. flexneri</i> type 5a | \rightarrow 2)- α -L-Rhap-(1 \rightarrow 2)-[α -D-Glcp-(1 \rightarrow 3)]- α -L-Rhap-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>S. flexneri</i> type 5b | \rightarrow 2)-[α -D-Glcp-(1 \rightarrow 3)]- α -L-Rhap-(1 \rightarrow 2)-[α -D-Glcp-(1 \rightarrow 3)]- α -L-Rhap-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>S. flexneri</i> type X | \rightarrow 2)-[α -D-Glcp-(1 \rightarrow 3)]- α -L-Rhap-(1 \rightarrow 2)- (1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow | 38 |
| | <i>S. pneumoniae</i> type 6B | \rightarrow 2)- α -D-Galp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 4)-D-Rib-ol-(5 \rightarrow P \rightarrow | 43 |
| | <i>S. pneumoniae</i> type 18C | \rightarrow 4)- β -D-Glcp-(1 \rightarrow 4)-[α -D-Glcp(6OAc)-(1 \rightarrow 2)]-[Gro-(1 \rightarrow P \rightarrow 3)- β -D-Galp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)]- β -L-Rhap-(1 \rightarrow P \rightarrow | 44 |
| | <i>S. pneumoniae</i> type 19A | \rightarrow 4)- β -D-ManpNAc-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow P \rightarrow | 45 |
| α -D-Glcp-(1 \rightarrow 4)-L-Rhap | <i>S. flexneri</i> type 2a | \rightarrow 2)- α -L-Rhap-(1 \rightarrow 2)- α -L-Rhap-(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 4)]- α -L-Rhap-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow | 38, 46 |

D. Conclusion

Glucosylation of *N*-acetyl-D-glucosamine using DSR-S vardel Δ4N led to the synthesis of α-D-glucopyranosyl-(1→6)-*N*-acetyl-D-glucosamine with a conversion rate of 24% yield. Notably, the use of a triple mutant of DSR-S vardel Δ4N, mutated in the region close to the catalytic residues (Δ4N SEV663NNS), increased the conversion rate up to 64 %, demonstrating that engineering of glucansucrases can significantly improve their transglucosylation activities toward poorly recognized non-natural acceptors. Recombinant AS is also shown to produce α-D-glucopyranosyl-(1→4)-*N*-acetyl-D-glucosamine. However, the glucosylation yield is very low, and rational engineering of the acceptor binding site should be considered to improve the enzyme.

Enzymatic glucosylation of L-rhamnose using glucansucrases appeared also doable. However, this monosaccharide was not an efficient acceptor substrate of glucansucrases. Multiple products were obtained in low yield with family 70 glucansucrases. Interestingly, a yet unreported (1→1)-linked disaccharide was synthesized using AS. Alternatively, blocking the anomeric position of L-rhamnose in methyl α-L-rhamnopyranoside enabled the formation of methyl α-D-glucopyranosyl-(1→3)-α-L-rhamnopyranoside and methyl α-D-glucopyranosyl-(1→4)-α-L-rhamnopyranoside with family 70 glycoenzymes. Methyl α-D-glucopyranosyl-(1→3)-α-L-rhamnopyranoside might be further considered for the chemo-enzymatic synthesis of a number of *S. flexneri* O-antigen fragments. Unfortunately, poor yields were obtained and GH70 glucansucrases turned out to lack regiospecificity. Besides, AS from family 13 did not recognize methyl α-L-rhamnopyranoside as an acceptor. This clearly demonstrates that, although they appear as promising tools for the stereoselective glucosylation of non natural targets of interest, glucansucrases have to be improved for efficient use with L-rhamnose or L-rhamnose derivatives, and more generally, in order to produce appropriate building blocks in amounts compatible with their use in multi-step synthesis. Semi-rational engineering will now be investigated to reach this purpose.

E. Experimental Section

E.1. Materials

E.1.a. Choice of recombinant glucansucrases

For glucansucrases (DSR-S, ASR), the enzyme constructs chosen for this study are respectively DSR-S vardel Δ4N and ASR C-del APY. Both enzymes are truncated variants of from *L. mesenteroides* NRRL B-512F dextranucrase and from *L. mesenteroides* NRRL B-1355 alternansucrase. They were constructed to reduce the glucansucrase degradation occurring during

heterologous enzyme expression by *E. coli* and have previously been shown to display the same behaviour than the wild-type enzyme in terms of specificity and synthesized products^{30,31}.

DSR-S varde1 $\Delta 4N$ variants ($\Delta 4N$ SEV663YDA, DSR-S SEV663NNS³²) are mutated in the active site downstream the catalytic aspartic acid D662. These mutants were designed by sequence alignment analyses with glucansucrases of various linkage specificities, and described to have a particular influence on acceptor binding, as they preferentially catalyze the formation of a disaccharide, to the detriment of dextran formation

Recombinant protein DSR-E from *L. mesenteroides* NRRL B-1299 synthesizes both α -1,6 and α -1,2 glucosidic linkages. In this work, the truncated form GBD-CD2, which is an engineered transglucosidase specific for the formation of α -1,2 linkages,³³ was used.

All recombinant enzymes were produced in *E.coli* as reported elsewhere^{7, 30, 31, 33}. Non-purified DSR-S and ASR (sonication supernatants) stored at -20°C were used for enzymatic assays. Purified GBD-CD2 and AS, conserved at -20°C or -80°C, served for enzymatic reactions

E.1.b. Chemical material

Sucrose, L-rhamnose and *N*-acetyl-D-glucosamine were purchased from Sigma-Aldrich. Methyl α -L-rhamnopyranoside,^{22, 23} and reference compounds methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside³⁵ and methyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranoside³⁴ were synthesised chemically at UCB.

E.2. Acceptor reaction assay

The glucosylation reaction was performed in the enzyme optimal buffer: in Tris-HCl (50 mM, pH = 7.5) for AS assay, in sodium acetate buffer (AcONa) (20 mM, pH = 5.4) for ASR assays, in AcONa (50 mM, [CaCl₂]= 0.05 g/L, pH = 5.2) for DSR-S assays and in AcONa (20 mM, [CaCl₂]= 0.5 g/L, pH = 5.4) for GBD-CD2 assays. The reaction mixture was carried out at 30 °C with sucrose and acceptor in equimolar ratio (146 mM). Enzymes were used at 1 U/mL. For glucansucrases (DSR-S, ASR and GBD-CD2), one unit is defined as the amount of enzyme that catalyzes the formation of 1 μ mol of fructose/min at 30°C, in enzyme buffer and 292 mM sucrose. For AS, activity was determined under the same conditions, except that sucrose was used at a concentration of 146 mM. The reaction was stopped by heating at 95°C for 5 min. The final mixture was centrifuged at 18 000 g for 10 min and filtered on a 0.22 μ m membrane before HPLC analysis.

E.3. Glucosyl acceptor production

In order to characterize glucosylated products of L-rhamnose and *N*-acetyl-D-glucosamine, acceptor reactions were conducted at preparative scale.

P1 from *N*-acetyl-D-glucosamine glucosylation by $\Delta 4N$ SEV667YDA (0.8 U/mL) was synthesized in 50 mL mixture reaction (500 mM in sucrose, 100 mM in acceptor); **P2** from *N*-acetyl-D-glucosamine glucosylation by purified AS (1 U/mL) was produced in 100 mL mixture reaction (292 mM in sucrose, 730 mM in acceptor). **P3** and **P4** from L-rhamnose glucosylation by purified AS (1 U/mL) were produced in 60 mL mixture reaction (292 mM in sucrose, 730 mM in acceptor). After a 24 h reaction time at 30°C, the media were centrifugated at 4800 rpm, for 20 min at 4°C to remove proteins and filtered for a better clarification. The purification of the glucosylated products was performed on a preparative octadecyl reverse-phase chromatography column (C18 column) (Bischoff Chromatography). Ultra pure water was used as eluent at a constant flow rate of 50 mL/min. Glucosyl detection was carried out with a refractometer, and each peak was collected separately, concentrated and reinjected into an analytical HPLC system to check the purity of the compounds.

E.4. Analytical methods

E.4.a. High Performance Liquid Chromatography (HPLC)

HPLC analysis device consisted in a Dionex P 680 series pump, a Shodex RI 101 series refractometer, a Dionex UVD 340 UV/Vis detector and an autosampler HTC PAL. Five columns were employed to separate the acceptor reaction products and to determine the acceptor conversion rate and product yields (i) a Biorad HPLC Carbohydrate Analysis columns: AMINEX HPX-87C at 80°C (elution with ultra-pure water at 0.6 mL/min) (ii) HPX-87K columns (300 x 7.8 mm) at 65°C (elution with ultra-pure water at 0.6 mL/min) (iii) C18 column Bischoff Prontosil Eurobond, 5 μ m (elution with ultra pure water at room temperature and 1mL/min) (iv) C30: Bischoff Prontosil Eurobond, 5 μ m, 250 x 4.0 mm (elution with ultra pure water at room temperature and 1mL/min) (v) C18RP: Sinergi Fusion RP Phenomenex, 4 μ m, 250 x 4.6 mm (elution with ultra pure water at room temperature and 1 mL/min).

E.4.b. High Resolution Mass Spectrometry (HRMS) and Nuclear Magnetic Resonance (NMR)

Accurate mass determination was carried out using an Autospec mass spectrometer arranged in an EBE geometry (Micromass, Manchester, UK). The instrument was operated at 8 kV accelerating voltage in positive mode. The caesium gun was set to 35 keV energy and 1 μ L of sample was mixed in the tip of the probe with a glycerol or dithiothreitol/dithioerythritol matrix.

NMR analyses: ^1H (400.130 MHz), ^{13}C (100.612 MHz), HSQC and HMBC were registered on a Bruker-ARX 400 spectrometer equipped with an ultrashim system. Samples were dissolved in deuterium oxide at c.a. 80 g/L and experiments were performed at 300K.

E.4.b.1. α -D-glucopyranosyl-(1 \rightarrow 6)-N-acetyl-D-glucosamine (P1)

HRMS (FAB): Anal. Calcd for C₁₄H₂₆NO₁₁: 384.1506 [MH⁺], Found: 384.1518; ¹H and ¹³C listed in TABLE 3.

E.4.b.2. α -D-glucopyranosyl-(1 \rightarrow 4)-N-acetyl-D-glucosamine (P2)

HRMS (FAB): Anal. Calcd for C₁₄H₂₅NO₁₁Na: 406.1325 [MNa⁺], Found: 406.1356; ¹H and ¹³C listed in TABLE 4.

E.4.b.3. α -D-glucopyranosyl-(1 \rightarrow 1)- β -L-rhamnopyranoside (P3)

HRMS (FAB): Anal. Calcd for C₁₂H₂₂O₁₀K: 365,0850 [MK⁺], Found: 365,0861; ¹H and ¹³C listed in TABLE 7.

E.4.b.4. α -D-glucopyranosyl-(1 \rightarrow 4)- α -D-glucopyranosyl-(1 \rightarrow 1)- β -L-rhamnopyranoside (P4)

HRMS (FAB): Anal. Calcd for C₁₈H₃₂O₁₅K: 527,1378 [MK⁺], Found: 527,1366; ¹H and ¹³C listed in TABLE 7.

F. References

1. Henrissat, B., A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, **1991**, 280 (Pt 2), 309-16.
2. Robyt, J. F.; Walseth, T. F., The mechanism of acceptor reactions of *Leuconostoc mesenteroides* B-512F dextranucrase. *Carbohydr Res*, **1978**, 61, 433-45.
3. Monchois, V.; Willemot, R. M.; Monsan, P., Glucansucrases: mechanism of action and structure-function relationships. *FEMS Microbiol Rev*, **1999**, 23, (2), 131-51.
4. Dols, M.; Remaud-Simeon, M.; Willemot, R. M.; Vignon, M.; Monsan, P., Characterization of the Different Dextranucrase Activities Excreted in Glucose, Fructose, or Sucrose Medium by *Leuconostoc mesenteroides* NRRL B-1299. *Appl Environ Microbiol*, **1998**, 64, (4), 1298-1302.
5. Bourne, E. J.; Sidebotham, R. L.; Weigel, H., Studies on dextrans and dextranases. XI. The structure of a dextran elaborated by *Leuconostoc mesenteroides* NRRL B-1299. *Carbohydr Res*, **1974**, 34, (2), 279-88.
6. Kobayashi, M.; Mitsuishi, Y.; Takagi, S.; Matsuda, K., Enzymic degradation of water-soluble dextran from *Leuconostoc mesenteroides* NRRL B-1299 *Carbohydr Res*, **1984**, 127, 305-317.
7. De Montalk, G. P.; Remaud-Simeon, M.; Willemot, R. M.; Planchot, V.; Monsan, P., Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J Bacteriol*, **1999**, 181, (2), 375-81.
8. Cote, G. L.; Holt, S. M.; Miller-Fosmore, C. M., Prebiotic oligosaccharides via alternansucrase acceptor reactions. *American Chemical Society*, **2003**, 849, 76-89.
9. Arguello Morales, M. A.; Remaud-Simeon, M.; Willemot, R. M.; Vignon, M. R.; Monsan, P., Novel oligosaccharides synthesized from sucrose donor and cellobiose acceptor by alternansucrase. *Carbohydr Res*, **2001**, 331, (4), 403-11.
10. Yoon, S. H.; Robyt, J. F., Synthesis of acarbose analogues by transglycosylation reactions of *Leuconostoc mesenteroides* B-512FMC and B-742CB dextranucrases. *Carbohydr Res*, **2002**, 337, (24), 2427-35.
11. Cote, G. L.; Dunlap, C. A.; Appell, M.; Momany, F. A., Alternansucrase acceptor reactions with D-tagatose and L-glucose. *Carbohydr Res*, **2005**, 340, (2), 257-62.
12. Richard, G.; Yu, S.; Monsan, P.; Remaud-Simeon, M.; Morel, S., A novel family of glucosyl 1,5-anhydro-D-fructose derivatives synthesised by transglucosylation with dextranucrase from *Leuconostoc mesenteroides* NRRL B-512F. *Carbohydr Res*, **2005**, 340, (3), 395-401.
13. Demuth, K.; Jordening, H. J.; Buchholz, K., Oligosaccharide synthesis by dextranucrase: new unconventional acceptors. *Carbohydr Res*, **2002**, 337, (20), 1811-20.
14. Cote, G. L.; Dunlap, C. A., Alternansucrase acceptor reactions with methyl hexopyranosides. *Carbohydr Res*, **2003**, 338, (19), 1961-7.
15. Richard, G.; Morel, S.; Willemot, R. M.; Monsan, P.; Remaud-Simeon, M., Glucosylation of α -butyl- and α -octyl-D-glucopyranosides by dextranucrase and alternansucrase from *Leuconostoc mesenteroides*. *Carbohydr Res*, **2003**, 338, (9), 855-64.
16. Seo, E. S.; Lee, J. H.; Park, J. Y.; Kim, D.; Han, H. J.; Robyt, J. F., Enzymatic synthesis and anti-coagulant effect of salicin analogs by using the *Leuconostoc mesenteroides* glucansucrase acceptor reaction. *J Biotechnol*, **2005**, 117, (1), 31-8.
17. Yoon, S. H.; Bruce Fulton, D.; Robyt, J. F., Enzymatic synthesis of two salicin analogues by reaction of salicyl alcohol with *Bacillus macerans* cyclomaltodextrin glucanyltransferase and *Leuconostoc mesenteroides* B-742CB dextranucrase. *Carbohydr Res*, **2004**, 339, (8), 1517-29.

18. Meulenbeld, G. H.; Hartmans, S., Transglycosylation by *Streptococcus mutans* GS-5 glucosyltransferase-D: acceptor specificity and engineering of reaction conditions. *Biotechnol Bioeng*, **2000**, 70, (4), 363-9.
19. Meulenbeld, G. H.; Zuilhof, H.; van Veldhuizen, A.; van den Heuvel, R. H.; Hartmans, S., Enhanced (+)-catechin transglucosylating activity of *Streptococcus mutans* GS-5 glucosyltransferase-D due to fructose removal. *Appl Environ Microbiol*, **1999**, 65, (9), 4141-7.
20. Bertrand, A.; Morel, S.; Lefoulon, F.; Rolland, Y.; Monsan, P.; Remaud-Simeon, M., *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr Res*, **2006**, 341, (7), 855-63.
21. Moon, Y. H.; Lee, J. H.; Ahn, J. S.; Nam, S. H.; Oh, D. K.; Park, D. H.; Chung, H. J.; Kang, S.; Day, D. F.; Kim, D., Synthesis, structure analyses, and characterization of novel epigallocatechin gallate (EGCG) glycosides using the glucansucrase from *Leuconostoc mesenteroides* B-1299CB. *J Agric Food Chem.*, **2006**, 54, (4), 1230-1237.
22. Rainer, H.; Scharf, H.-D.; Runsink, J., Reactions of Partially Acylated Aldohexopyranosides, VIII. A New Synthesis of the Tetradeoxydisaccharide in Avermectins. *Liebigs Ann. Chem.*, **1992**, (2), 103-107.
23. Binkley, R. W.; Goewey, G. S.; Johnston, J. C., Regioselective ring opening of selected benzylidene acetals: a photochemically initiated reaction for partial deprotection of carbohydrates. *J.Org.Chem.*, **1984**, 49, 992-996.
24. Phalipon, A.; Costachel, C.; Grandjean, C.; Thuizat, A.; Guerreiro, C.; Tanguy, M.; Nato, F.; Vulliez-Le Normand, B.; Belot, F.; Wright, K.; Marcel-Peyre, V.; Sansonetti, P. J.; Mulard, L. A., Characterization of functional oligosaccharide mimics of the *Shigella flexneri* serotype 2a O-antigen: implications for the development of a chemically defined glycoconjugate vaccine. *J Immunol*, **2006**, 176, (3), 1686-94.
25. Verez-Bencomo, V.; Fernandez-Santana, V.; Hardy, E.; Toledo, M. E.; Rodriguez, M. C.; Heynngnezz, L.; Rodriguez, A.; Baly, A.; Herrera, L.; Izquierdo, M.; Villar, A.; Valdes, Y.; Cosme, K.; Deler, M. L.; Montane, M.; Garcia, E.; Ramos, A.; Aguilar, A.; Medina, E.; Torano, G.; Sosa, I.; Hernandez, I.; Martinez, R.; Muzachio, A.; Carmenates, A.; Costa, L.; Cardoso, F.; Campa, C.; Diaz, M.; Roy, R., A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b. *Science*, **2004**, 305, (5683), 522-5.
26. Torano, G.; Toledo, M. E.; Baly, A.; Fernandez-Santana, V.; Rodriguez, F.; Alvarez, Y.; Serrano, T.; Musachio, A.; Hernandez, I.; Hardy, E.; Rodriguez, A.; Hernandez, H.; Aguilar, A.; Sanchez, R.; Diaz, M.; Muzio, V.; Dfana, J.; Rodriguez, M. C.; Heynngnezz, L.; Verez-Bencomo, V., Phase I clinical evaluation of a synthetic oligosaccharide-protein conjugate vaccine against *Haemophilus influenzae* type b in human adult volunteers. *Clin Vaccine Immunol*, **2006**, 13, (9), 1052-6.
27. Mooser, G.; Hefta, S. A.; Paxton, R. J.; Shively, J. E.; Lee, T. D., Isolation and sequence of an active-site peptide containing a catalytic aspartic acid from two *Streptococcus sobrinus* α -glucosyltransferases. *J Biol Chem*, **1991**, 266, (14), 8916-22.
28. Mooser, G.; Iwaoka, K. R., Sucrose 6- α -D-glucosyltransferase from *Streptococcus sobrinus*: characterization of a glucosyl-enzyme complex. *Biochemistry*, **1989**, 28, (2), 443-9.
29. Jensen, M. H.; Mirza, O.; Albenne, C.; Remaud-Simeon, M.; Monsan, P.; Gajhede, M.; Skov, L. K., Crystal structure of the covalent intermediate of amylosucrase from *Neisseria polysaccharea*. *Biochemistry*, **2004**, 43, (11), 3104-10.
30. Joucla, G.; Pizzut, S.; Monsan, P.; Remaud-Simeon, M., Construction of a fully active truncated alternansucrase partially deleted of its carboxy-terminal domain. *FEBS Lett*, **2006**, 580, (3), 763-8.
31. Moulis, C.; Arcache, A.; Escalier, P. C.; Rinaudo, M.; Monsan, P.; Remaud-Simeon, M.; Potocki-Veronese, G., High-level production and purification of a fully active

recombinant dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. *FEMS Microbiol Lett*, **2006**, 261, (2), 203-10.

32. Moulis, C.; Joucla, G.; Harrison, D.; Fabre, E.; Potocki-Veronese, G.; Monsan, P.; Remaud-Simeon, M., Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J Biol Chem*, **2006**, 281, (42), 31254-67.

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

34. Mulard, L. A.; Costachel, C.; Sansonetti, P. J., Synthesis of the methyl glycosides of a di- and two trisaccharide fragments specific for the *Shigella flexneri* serotype 2a O-antigen. *J. Carbohydr. Chem.*, **2000**, 19, (7), 849-877.

35. Bakinovskii, L. V.; Gomtsian, A. R.; Bairamova, N. E.; Kochetkov, N. K., [Synthesis of oligosaccharide fragments of O-specific *Shigella flexneri* polysaccharides. II. Synthesis of trisaccharide α Glc(1-3) α Rha-(1-2) α RhaOMe and tetrasaccharide β GlcNAc(1-2) α Glc(1-3) α Rha(1-2) α RhaOMe]. *Bioorg Khim*, **1985**, 11, (2), 254-63.

36. Pozsgay, V., Oligosaccharide-protein conjugates as vaccine candidates against bacteria. *Adv Carbohydr Chem Biochem*, **2000**, 56, 153-99.

37. Belot, F.; Guerreiro, C.; Baleux, F.; Mulard, L. A., Synthesis of two linear PADRE conjugates bearing a deca- or pentadecasaccharide B epitope as potential synthetic vaccines against *Shigella flexneri* serotype 2a infection. *Chemistry*, **2005**, 11, (5), 1625-35.

38. Simmons, D. A., Immunochemistry of *Shigella flexneri* O-antigens: a study of structural and genetic aspects of the biosynthesis of cell-surface antigens. *Bacteriol Rev*, **1971**, 35, (2), 117-48.

39. Jann, B.; Shashkov, A. S.; Gupta, D. S.; Jann, K., The O18 antigens (lipopolysaccharides) of *Escherichia coli*. Structural characterization of the O18A, O18A1, O18B and O18B1-specific polysaccharides. *Eur J Biochem*, **1992**, 210, (1), 241-8.

40. Gorshkova, R. P.; Isakov, V. V.; Nazarenko, E. L.; Ovodov, Y. S.; Guryanova, S. V.; Dmitriev, B. A., Structure of the O-specific polysaccharide of the lipopolysaccharide from *Yersinia kristensenii* O:25.35. *Carbohydr Res*, **1993**, 241, 201-8.

41. Pieretti, G.; Corsaro, M. M.; Lanzetta, R.; Parrilli, M.; Nicolaus, B.; Gambacorta, A.; Lindner, B.; Holst, O., Structural Characterization of the Core Region of the Lipopolysaccharide from the *Haloalkaliphilic Halomonas pantelleriensis*: Identification of the Biological O-Antigen Repeating Unit *Eur. J. Org. Chem.*, **2008**, (4), 721-728.

42. Adams, M. M.; Allison, G. E.; Verma, N. K., Type IV O antigen modification genes in the genome of *Shigella flexneri* NCTC 8296. *Microbiology*, **2001**, 147, (Pt 4), 851-60.

43. Kenne, L.; Lindberg, B.; Madden, J. K., Structural studies of the capsular antigen from *Streptococcus pneumoniae* type 26. *Carbohydr Res*, **1979**, 73, 175-82.

44. Jones, C., NMR assays for carbohydrate-based vaccines. *J Pharm Biomed Anal*, **2005**, 38, (5), 840-50.

45. Katzenellenbogen, E.; Jennings, H. J., Structural determination of the capsular polysaccharide of *Streptococcus pneumoniae* type 19A (57). *Carbohydr Res*, **1983**, 124, (2), 235-45.

46. Lindberg, A. A.; Karnell, A.; Weintraub, A., The lipopolysaccharide of *Shigella* bacteria as a virulence factor. *Rev Infect Dis*, **1991**, 13 Suppl 4, S279-84.

**CHAPITRE III : Conception à façon de
transglucosidases pour la synthèse
chimio-enzymatique d'antigènes
osidiques**

L'étude précédente nous a permis de sélectionner les deux glucane-saccharases recombinantes, les plus prometteuses pour la glucosylation de la *N*-acétyl-D-glucosamine :

- l'amylosaccharase, seule enzyme capable de produire le disaccharide d'intérêt mais en très faible quantité (4%).
- DSR-S vardel $\Delta 4N$, et plus particulièrement ses variants, avec lesquels on obtient majoritairement un accepteur mono-glucosylé (54% et 61%) mais ne présentant toutefois pas la bonne spécificité de liaison.

Concernant la glucosylation du L-rhamnose et de ses dérivés, les rendements en disaccharide obtenus avec les glucane-saccharases recombinantes sont faibles. Seuls les variants de DSR-S vardel $\Delta 4N$ et l'ASR forment le disaccharide ciblé mais en quantité très faible (3-4%).

De plus, un autre élément a été déterminant dans le choix des enzymes à optimiser : l'amylosaccharase est la seule enzyme pour laquelle nous disposons de la structure tridimensionnelle. Il s'agit ainsi d'un atout considérable pour identifier de façon rationnelle des cibles de mutagenèse visant à modifier le site de reconnaissance de l'accepteur.

Nous avons finalement retenu, comme enzymes candidates à l'optimisation, DSR-S vardel $\Delta 4N$ pour la glucosylation des dérivés de la *N*-acétyl-D-glucosamine et l'amylosaccharase pour la glucosylation des deux types d'accepteurs.

Optimisation de DSR-S vardel $\Delta 4N$:

Au vu des résultats obtenus à l'aide de l'enzyme recombinante, notre objectif était de modifier par ingénierie de l'enzyme la spécificité de liaison (α , 1-6) obtenue avec la *N*-acétyl-D-glucosamine pour former une nouvelle liaison de type (α , 1-4). En l'absence de structure 3-D, nous nous sommes appuyés sur des alignements de séquences réalisés sur des glucane-saccharases de spécificités variées, notamment les GTF-A et GTF-Bio, qui sont les 2 seules enzymes à présenter une spécificité de liaison (α , 1-4). Cinq zones, dont quatre sont identifiées en page 85 et situées en aval des trois résidus catalytiques, sont supposées être impliquées dans la spécificité de cette liaison (α , 1-4). La mutation de la 4^{ème} zone correspond à la construction du mutant, DSR-S vardel $\Delta 4N$ SEV663NNS.

Nous avons réalisé 7 autres constructions (non détaillées dans ce manuscrit), à partir du mutant DSR-S vardel $\Delta 4N$ SEV663NNS, pour introduire individuellement ou simultanément les résidus de la GTF-A supposés être impliqués dans la spécificité de cette liaison (α , 1-4). De telles mutations ont conduit à une perte totale de l'activité catalytique de l'enzyme.

Ces travaux illustrent, une fois de plus, la difficulté à changer la spécificité d'une enzyme, en particulier en l'absence d'information sur la structure tridimensionnelle de l'enzyme.

Optimisation de l' amylosaccharase de *Neisseria polysaccharea* :

Nos efforts se sont, donc, davantage concentrés sur l'exploitation de la structure 3-D de l'amylosaccharase afin de remodeler le site de fixation de l'accepteur (sous-site (+1)) et concevoir de nouvelles protéines offrant une meilleure complémentarité topographique et électrostatique vis-à-vis des accepteurs considérés. Des travaux de modélisation moléculaire ont été entrepris sur les 2 motifs disaccharidiques ciblés : α -D-Glcp-(1→3)- α -L-Rhap (sérotipe 3a) et α -D-Glcp-(1→4)-D-GlcNAcp (sérotipe 1b). Sur la base de ces résultats, 7 positions ont été identifiées comme prometteuses. Une banque de 133 mono-mutants a ainsi été construite en remplaçant individuellement les 7 résidus sélectionnés par les 19 autres acides aminés possibles. Cette banque a ensuite été criblée sur les deux accepteurs cibles.

Dans la deuxième publication, nous décrivons plus amplement la stratégie employée, l'identification de 43 mono-mutants d'intérêt, la caractérisation des 2 variants les plus améliorés et leur utilisation dans la validation de l'approche chimio-enzymatique.

TAILORED TRANSGLUCOSIDASES FOR INNOVATIVE CHEMO-ENZYMATIC SYNTHESIS OF ANTIGENIC OLIGOSACCHARIDES

Elise Champion^{a,b,c}, Isabelle André^{a,b,c}, Claire Moulis^{a,b,c}, Julien Boutet^e, Karine Descroix^e,
Sandrine Morel^{a,b,c}, Pierre Monsan^{a,b,c,d}, Laurence Mulard^e and Magali Remaud-Siméon^{a,b,c}.

^a Université de Toulouse; INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

^b CNRS, UMR5504, F-31400 Toulouse, France

^c INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

^d Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris

^e Unité de Chimie des Biomolécules, URA CNRS 2128, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris cedex 15, France

A. Summary

The difficult access to complex carbohydrates and glycoconjugates using a chemical synthesis impairs their development on a large scale. The use of biocatalysts holds great potential to circumvent the boundaries defined by chemistry, yet poorly exploited in spite of the fast-growing development of engineering technologies allowing the search and the construction of better performing enzymes, displaying novel specificities. So far, the challenging design of an appropriate enzymatic glycosylation tool that would allow an optimal combination of the chemical and enzymatic steps involved in the synthesis of complex oligosaccharides has never been attempted, although it would be of major interest to develop new synthetic pathways.

We describe promising *in-vitro* chemo-enzymatic synthetic methodologies to produce complex bacterial carbohydrates, that demonstrate the applicability of glycoenzymes engineering, purposely designed, to realize an enzymatic glycosylation step at an early stage in the synthesis, compatible with the requirements defined to ensure the correct stereochemical outcome of the glucosylation step for subsequent chemical elongation. A semi-rational engineering approach was used to re-design the enzyme active site and adapt it to the glucosylation of non-natural acceptors to produce the serotype-specific branched α -D-glucopyranosyl linkages of *Shigella flexneri* O-antigen oligosaccharides, which are of prime interest for the synthesis of third generation vaccines against shigellose. This innovative chemo-enzymatic process may be seen as an alternative for the synthesis of *S. flexneri* O-antigen fragments, but more generally the concept validation disclosed herein offers new opportunities in the synthesis of complex oligosaccharides providing that an

appropriate combination of engineered transglycosidase/donor/acceptor system is made available based on compatibility with chemical chain elongation.

B. Introduction

Recent developments in glycochemistry, providing more straightforward access to well-defined carbohydrates, have opened the way to significant progress in the fields of glycobiology and glycotherapeutics¹⁻³. Efficient and elegant synthetic methods such as one-pot oligosaccharide synthesis⁴, or automated synthetic tools⁵, have been developed to provide more straightforward access to well-defined carbohydrates. The use of lightly protected precursors and the regioselective one-pot protection of monosaccharides were recently emphasized⁶. Nevertheless, despite accomplished advancements, chemical approaches towards specific usable microbial oligosaccharides still need considerable effort. They often remain highly dependent on the design of appropriate combinations of multiple protection, deprotection, and efficient glycosylation steps⁷.

Avoiding the need for protecting groups, but remaining to date limited to the synthesis of short oligosaccharides which can passively cross the cell membrane, the use of metabolically engineered cell factories is found attractive^{8, 9, 10}. Alternatively, enzymes have emerged as practical alternatives to a number of limitations encountered in chemical synthesis^{11, 12}. Noteworthy, the combination of multiple overexpressed native glycosyltransferases was highly rewarding for the synthesis of the upstream pentasaccharide terminus of the *Neisseria meningitidis* lipooligosaccharide¹³, but examples of *in vitro* enzymatic synthesis of complex carbohydrate remain scarce¹⁴. Despite the increasing number of available carbohydrate-active enzymes, in particular glycosyltransferases and transglycosidases, the lack of appropriate enzymatic tools with requisite substrate specificity has prevented extensive exploration of chemo-enzymatic routes to complex oligosaccharides. Membrane Leloir-type glycosyltransferases are not all easily available. Their nucleotide activated sugar substrates are expensive, or generated *in situ* at the cost of additional enzymes¹⁴. Alternatively, transglycosidases proceed from diverse non-nucleotide sugar donors and are compatible with a broader variety of acceptors¹⁵. Interestingly, modified donors were occasionally used successfully¹⁶. Nonetheless, enzyme availability is often critical, especially when considering the appropriate regio- and stereospecificity required for a given target^{15, 17, 18}. Moreover, *in vitro* enzymatic glycosylation, relying on enzyme/acceptor compatibility, mostly involves fully deprotected acceptors. Thus, it often comes at the latest stages in the synthesis of complex carbohydrates^{14, 19, 20}.

To circumvent these limitations, protein engineering has proven to be extremely efficient to generate biocatalysts, including carbohydrate-active enzymes with novel substrate specificities^{21, 22}. However, although rather appealing, the quite challenging design of an appropriate enzymatic glycosylation tool that would allow an optimal combination of the chemical and enzymatic steps at any stage of the synthesis of complex oligosaccharides has never been attempted. Along this line, in the present article, our goal is to demonstrate the applicability of glycoenzymes engineering to obtain complex bacterial carbohydrates by innovative *in-vitro* chemo-enzymatic synthetic methodologies involving an enzymatic glycosylation step at an early stage in the synthesis. We turned to the use of glucansucrases, α -retaining transglucosidases belonging to families 13 and 70 of glycoside-hydrolases, known to accept sucrose as low cost glucosyl donor and to be flexible in terms of acceptor specificity, to probe the potential of our strategy. The approach was applied to the synthesis of oligosaccharide fragments of *Shigella flexneri* lipopolysaccharide. *Shigella*, the causal agent of shigellosis, or bacillary dysentery, represents a major burden worldwide. No vaccine is yet available^{23,24}. In countries where shigellose is endemic, a number of *S. flexneri* serotypes have been isolated, emphasizing the need for a multivalent vaccine^{23,24}. Interestingly, protein-conjugates of synthetic oligosaccharides mimicking *S. flexneri* serotype 2a O-antigen induced in mice a potent anti-O-antigen humoral immune response, which was shown to be protective against homologous challenge^{25, 26}.

The diversity, associated to a close resemblance in composition, of the known *S. flexneri* O-antigen repeating units was found of utmost interest to challenge our goal. Indeed, most known *S. flexneri* O-antigen repeating units share a linear tetrasaccharide backbone²⁷. Serotype specificity involves O-acetyl and/or α -D-glucopyranosyl decorations^{27, 28}. Interestingly, 4 different patterns of α -D-glycosylation, involving either L-rhamnopyranose or N-acetyl-D-glucopyranosamine residues, have been characterized for *S. flexneri* polysaccharides. Herein, focus is the enzymatic synthesis of α -D-glucopyranosyl-(1 \rightarrow 4)-N-acetyl- α -D-glucopyranosaminyl (ED) branching pattern shared by serotypes 1a and 1b^{27, 28} (**FIGURE 38**), and on the α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranosyl (EA) side chain common to serotypes 2b, 3a, 5a, 5b, and X^{27, 28} (**FIGURE 39**).

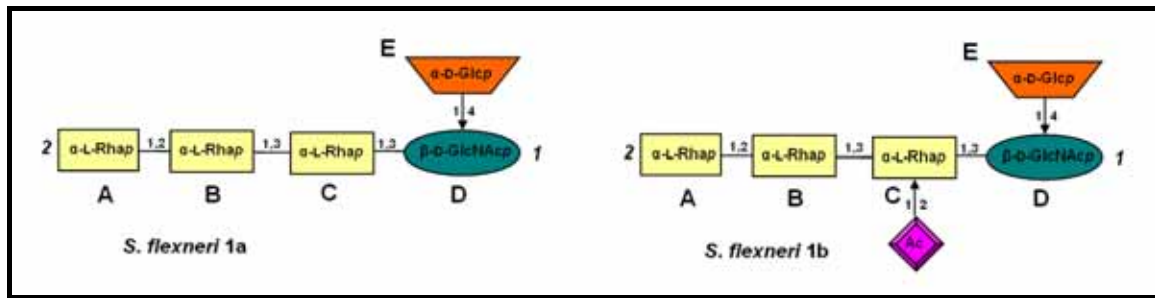


FIGURE 38| Repeating unit of *S. flexneri* serotypes 1a and 1b O-Ags.
 Rhap=L-Rhamnopyranosyl - GlcNAc=N-acetyl-D-glucosaminopyranosyl - Glcp=D-Glucopyranosyl

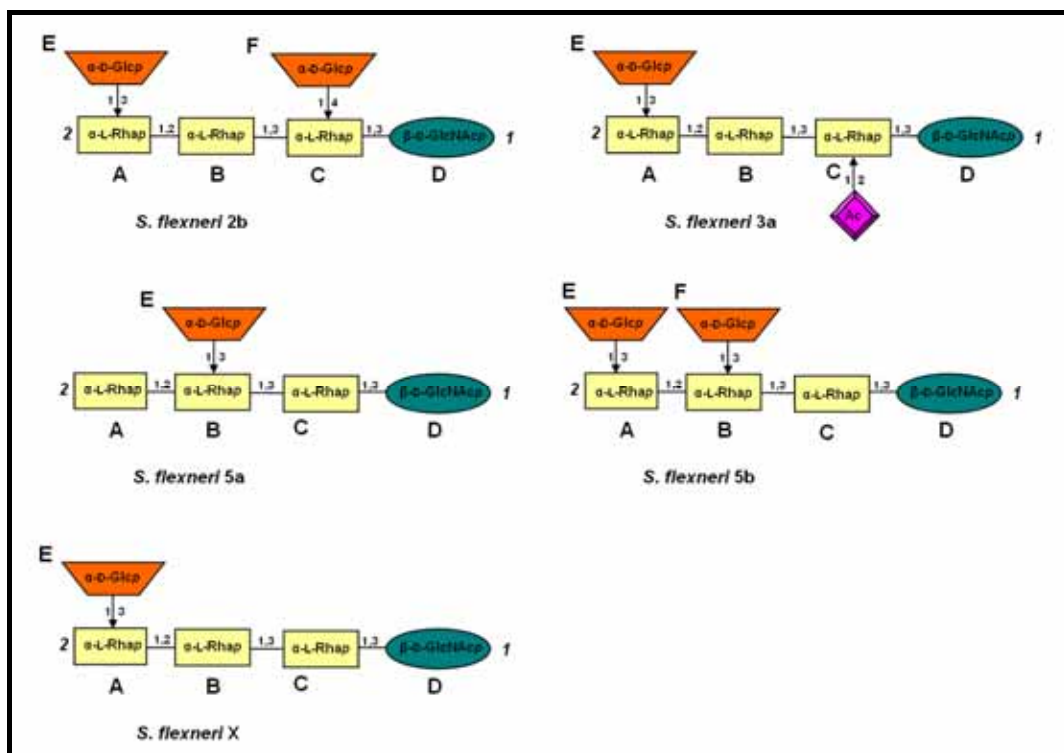


FIGURE 39| Repeating unit of *S. flexneri* serotypes 2b, 3a, 5a, 5b and X O-Antigens.
 Rhap=L-Rhamnopyranosyl - GlcNAc=N-acetyl-D-glucosaminopyranosyl - Glcp=D-Glucopyranosyl

C. Results

C.1. Chemo-enzymatic pathway

Our goal was to find combinations of transglucosidase/low cost donors systems and protected glucosaminyl and rhamnopyranosyl acceptors compatible with the production of large amounts of disaccharides **EA** and **ED** in a form allowing multiple-step chemical chain elongation at both ends of the acceptor residues. To avoid the need for sugar nucleotide donor, we favoured the use of transglucosidases commonly named glucansucrases (**FIGURE 40**). These enzymes are highly specific for sucrose, their donor substrate, but tolerate diverse sugar acceptors²⁹⁻³².

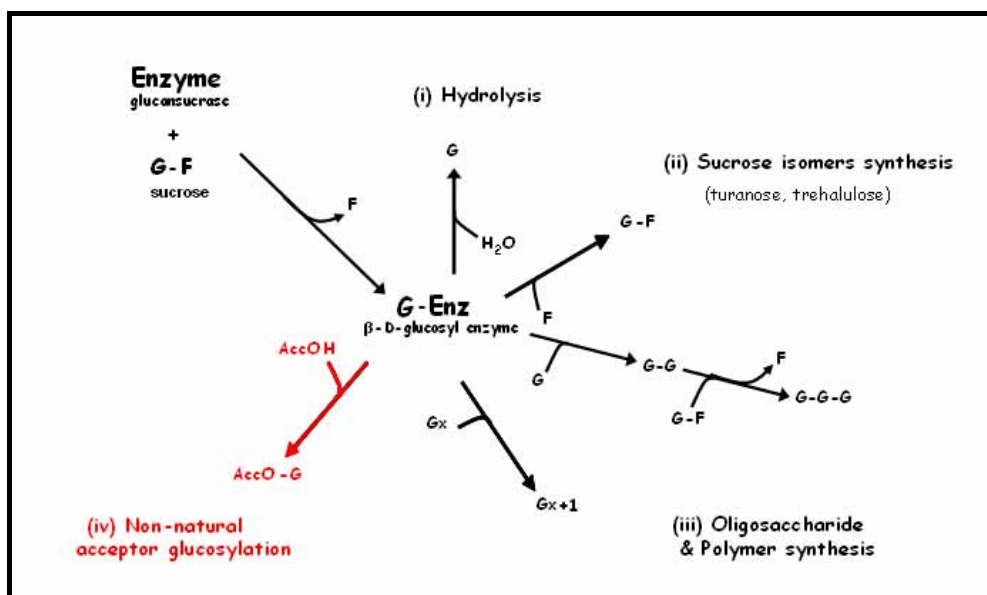


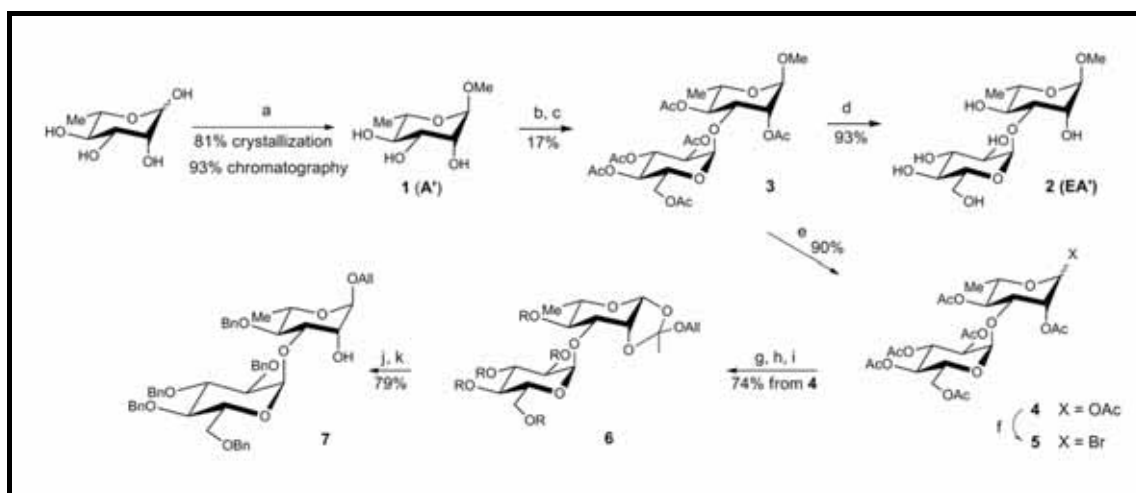
FIGURE 40| Reactions catalyzed by glucansucrases.

Glucansucrases are α -retaining transglucosidases found in families 13 and 70 of glycoside-hydrolases. Glucansucrases follow a double displacement retaining mechanism, in which a β -glucosyl enzyme covalent intermediate is first formed from sucrose substrate. In a second step, the glucosyl moiety is transferred onto an acceptor which can be depending on the conditions of reaction: (i) either water what leads to the production of glucose (ii) fructose to form sucrose isomers (iii) the glucose moiety released from hydrolysis to form soluble oligosaccharides and α -glucan polymers (iv) or an exogeneous hydroxylated acceptor, when the latter is well recognized. Depending on the enzyme regiospecificity, distinct types of glucosidic linkage are found in the polymer formed. The polymerization reaction can be redirected toward the glucosylation of non-natural acceptors, reaction of interest shown in red. α -L-RhaOMe (**A'**) and D-GlcpNAc (**D**) target acceptors are not recognized by the recombinant glucansucrases tested and specific for α -1,2, α -1,3, α -1,4 and α -1,6 glucosidic bond formation. G: glucosyl; F: fructosyl

C.2. Choice of acceptor substrates

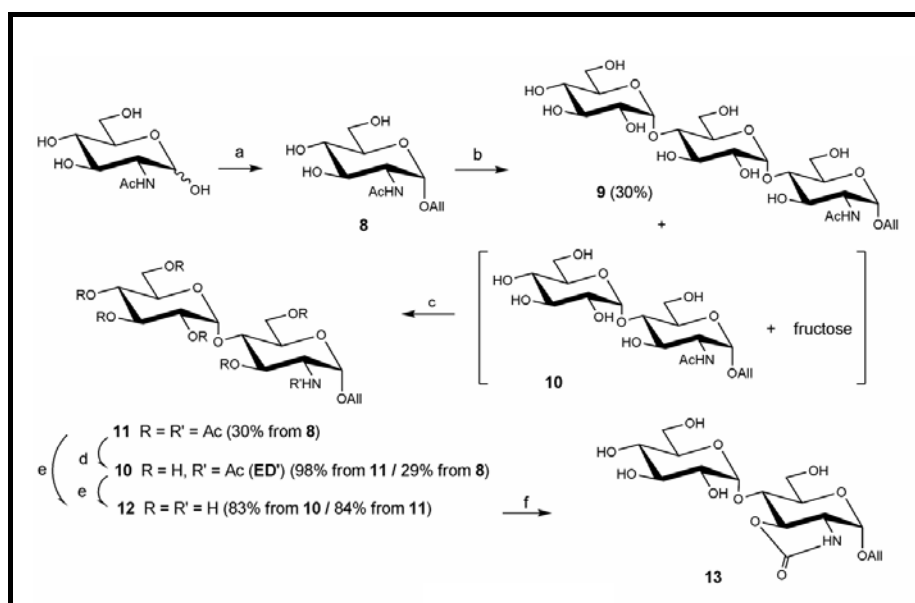
The choice of suitable precursors to residues **A** and **D** to be used as acceptors in the enzymatic step was crucial. It obeyed three rules: (i) light protecting pattern compatible with the ability of selected glucansucrases to modulate their acceptor binding site, (ii) easy synthetic access, and (iii) possible site-selective chemical elongation of the glucosylation product at either the reducing or non reducing end.

Defining an appropriate precursor **A'** to residue **A** matching those requirements was an easy task, thanks to previous achievements on *S. flexneri* 5a oligosaccharides³³. Methyl α -L-rhamnopyranoside³⁴ (**A'**, **1**), a crystalline material obtained in one step from L-rhamnose (**SCHEME 1** and **SUPPLEMENTARY METHODS**), was selected accordingly. Indeed, conversion of the glucosylation product α -D-Glcp-(1 \rightarrow 3)- α -L-RhapO-Me³⁵ (**EA'**, **2**) into allyl (2,3,4,6-tetra-*O*-benzyl- α -D-glucopyranosyl)-(1 \rightarrow 3)-4-*O*-benzyl- α -L-rhamnopyranoside³⁶ (**7**) was thought feasible. Acting as an acceptor and potential donor, disaccharide **7** was as an exquisite building block in the synthesis of several *S. flexneri* 5a³⁷, 3a, and X O-antigen fragments.


SCHEME 1| Chemo-enzymatic synthesis of key building block 7.

a. MeOH, H⁺; b. non-purified I228Y extract, sucrose; c. Ac₂O, Pyridine; d. K₂CO₃, MeOH; e. Ac₂O, H⁺; f. HBr/AcOH; g. MeOH, Lutidine; h. K₂CO₃, MeOH; i. NaH, BnBr; j. AlIOH, TMSOTf, k. MeONa, MeOH.

The choice of an appropriate precursor (**D'**) to residue **D** obeyed the same criteria. In the absence of any model disaccharide to probe the chemo-enzymatic route, it relied extensively on literature data. Indeed, allyl 2-*N*-acetyl-2-deoxy- α -D-glucopyranoside (**D'**, **11**) a crystalline material prepared in one step (84%) from *N*-acetyl-D-glucosamine (**SCHEME 2** and **SUPPLEMENTARY METHODS**) was selected based on the assumption that the 3_{D'}-OH group would be easily differentiated at the disaccharide level providing that **16**, bearing a 2,3-oxazolidinone moiety would derive from **15**, itself resulting from *N*-deacetylation of the product of enzymatic glucosylation **13**. Regioselective differentiation of the 3_{D'}-OH is indeed a pre-requirement to any specific chain elongation at this position as present in *S. flexneri* 1a and 1b O-antigens.


SCHEME 2| Chemo-enzymatic synthesis of disaccharide 10.

a. AlIOH, BF₃.OEt₂; b. non-purified F290K extract, sucrose; c. Ac₂O, Pyridine; d. MeONa, MeOH; e. Ba(OH)₂, H₂O; f. pNO₂C₆H₄OCOCI, MeONa, MeOH

C.3. Screening for native glucansucrases able to synthesize the starting building blocks

Glucansucrases had never been tested against the acceptors of interest. Glucosylation of both **A'** and D-GlcpNAc (**D**) was thus attempted with four recombinant glucansucrases, which were selected for their very distinct specificities, namely their ability to form α -1,2, α -1,3, α -1,4 and α -1,6 glucosidic bond, respectively. Note that this first screening was run with commercially available **D** rather than the allyl glycoside **D'**. Indeed, relying on earlier studies on glucosides³⁸, we assumed that the anomeric allyl aglycone would not significantly affect acceptor recognition. None of the glucansucrases could glucosylate these acceptors in reasonable yields, with the required regiospecificity³⁹. Noteworthy one of the glucansucrases namely the recombinant amylosucrase from *Neisseria polysaccharea* (**wtAS**), was shown to be unable to glucosylate α -L-RhapOMe, but in the presence of D-GlcpNAc as acceptor, this enzyme synthesized the desired α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc disaccharide, although in very low yield³⁹ ($\leq 2\%$).

C.4. Engineering of amylosucrases able to glucosylate the target acceptors

Engineering novel transglucosidases with altered acceptor substrate specificity was our next goal. Interestingly, there is yet no description of substrate acceptance enlargement for α -retaining transglycosidases. Among the enzyme tested, **wtAS** is the only glucansucrase whose three-dimensional structure is available in complex with either the substrate or the natural reaction product⁴⁰. It was therefore chosen as a model enzyme to authorize a semi-rational engineering approach aiming at changing the enzyme specificity. Indeed, our in-depth knowledge of the structural features of the enzyme active site and their contribution to function were used to guide site-directed evolution targeted at the binding pocket (**FIGURE 41A**). Our approach consisted in (i) mapping the binding site residues important for functional plasticity, (ii) identifying the most promising positions to be modified to favour acceptor recognition without interfering with sucrose substrate binding at subsites -1 and +1, and (iii) limiting side-reactions involving natural acceptors present in the reaction mixture (ie. hydrolysis, polymerisation and sucrose isomerisation) by reducing their binding to mutants. Mutation selection relied on docking studies of disaccharides **EA'**, and **ED**, in the active site of **wtAS**^{41,42}.

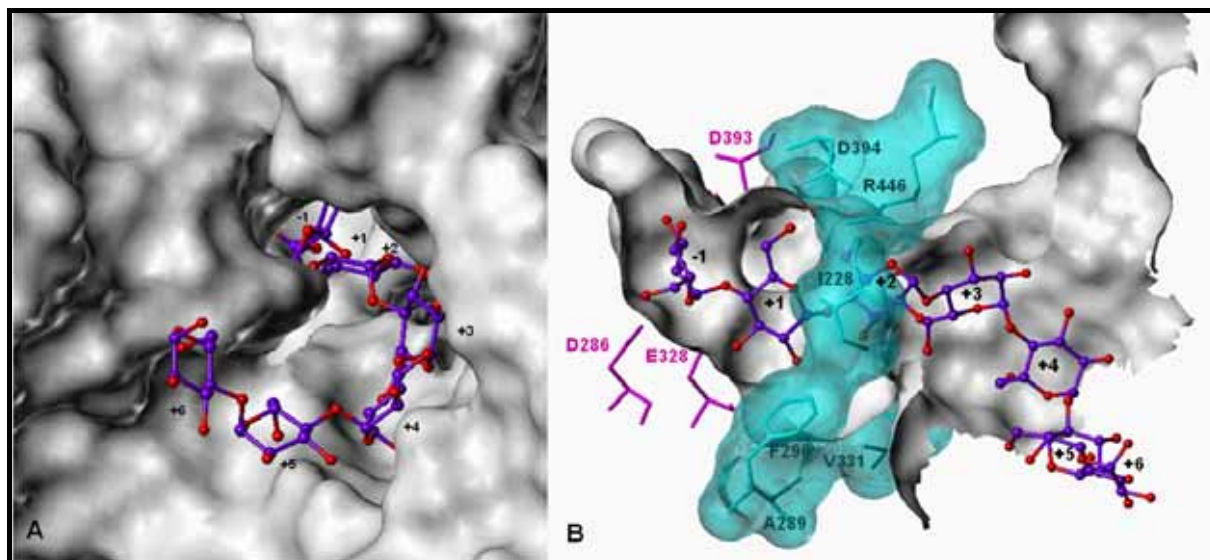


FIGURE 41| View of amylosucrase active site in complex with maltoheptaose (PDB: 1MW0).

(A) Amylosucrase is viewed as a Connolly surface, as calculated by the MOLCAD module implemented in Sybyl 7.3. The active site is located at the bottom of a narrow and 15 Å deep pocket. Mapping of maltoheptaose subsites -1 to +6 was done according to the nomenclature earlier described for glycoside hydrolases⁴³.

(B) Cross view of the active site. Residues forming the catalytic triad (E286, E328 and D393), located at the bottom of the pocket, are coloured in magenta. The 7 positions identified for site-directed mutagenesis (I228, A289, F290, I330, V331, D394 and R446) are shown in cyan colour. I330 is not shown on the figure, as it is located on the opposite side. These residues form a crown that encircles subsite +1.

The subsite -1 is responsible for the specificity towards sucrose and is occupied by the glucosyl unit which will be transferred during the reaction whereas the subsite +1 ensures a correct positioning of the acceptor and is also responsible for specificity of synthesis of the (α -1 \rightarrow 4) glucan linkage.

C.4.a. Structural analysis of AS: α -D-Glcp-(1 \rightarrow 3)- L-RhapOMe [EA'] complex

Of all docking solutions, the conformation of EA' in the enzyme pocket with the most favourable binding energy was considered for detailed structural analysis (FIGURE 42 and SUPPLEMENTARY FIGURES 1B and 1C). A detailed comparison of both disaccharides bound in the enzyme pocket allowed to pinpoint the key amino acid side-chains in close contact with EA' that could be targeted by site-directed mutagenesis to facilitate the binding of the acceptor in a catalytically productive position and provide a new network of interactions specific to A'. After an extensive review of the residues surrounding EA', the corresponding amino acids Ile228, Ile330, Asp394 and Arg446 were proposed for amino acid exchange. We reasoned that introduction of less hindered hydrophobic residues at positions 228 and 446 could bring beneficial van der Waals interactions with the 6-methyl group of A' and assist a better positioning of the acceptor in the catalytic pocket. In addition, the introduction of longer polar side chains at position 394 and 330 were thought to be important to recover some stabilizing hydrogen bonds with the acceptor. (SUPPLEMENTARY FIGURES 1B and 1C).

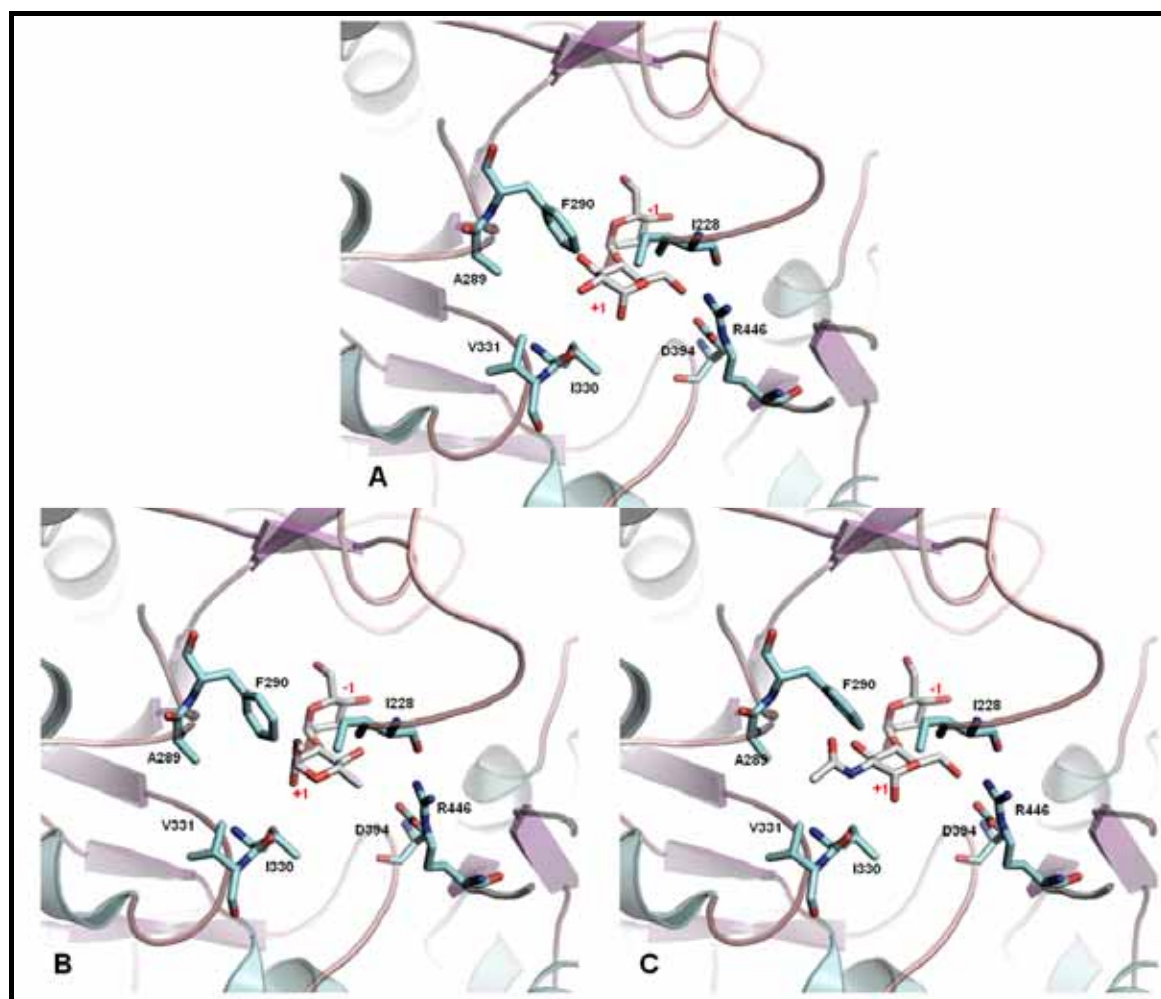


FIGURE 42| Comparison of docking modes.

(A) Maltose moiety from the crystallographic maltoheptaose (PDB: 1MW0) occupying binding subsites (-1) and (+1) of amylosucrase from *Neisseria polysaccharea*, (B) α -D-Glcp-(1 \rightarrow 3)- α -L-RhapOMe (EA'), (C) α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc (ED) in the active site of AS. The 7 amino acid residues (I228, A289, F290, I330, V331, D394 and R446) selected for mutagenesis are shown on the figures. Hydrogen atoms have been omitted for clarity purpose.

The α -D-glucopyranosyl moiety (E) of EA' at subsite -1 is more tightly bound to the enzyme than A' at subsite +1. Comparison of the docking modes adopted by EA' with crystallographic maltose (α -D-Glcp-(1 \rightarrow 4)-D-Glcp), in particular at subsite +1, revealed a 'flip' of the pyranose ring, due to the inversion in the ring chair conformation adopted by A' (1C₄ vs 4C₁, respectively) as well as a 'rotation' of the pyranose ring due to the altered osidic linkage (α -1 \rightarrow 3 vs α -1 \rightarrow 4, respectively). As a result of these main structural differences, the network of interactions between with the wild type enzyme was drastically impaired, particularly at subsite +1.

C.4.b. Structural analysis of AS: α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc [ED] complex

The docking mode adopted by ED (target product) in AS binding site emerged very close to that observed by crystallography for maltose glucosyl units at subsites -1 and +1 (FIGURE 42). As amylosucrase synthesizes naturally an α -(1 \rightarrow 4) glucosidic linkage, the unique difference between both disaccharides resides in the *N*-acetyl group at C_{2D} that replaces the hydroxyl group in D-Glcp. Docking results indicated that to accommodate D at subsite +1, several hydrophobic residues (Ala289, Phe290, Ile330 and Val331) forming a pocket surrounding the *N*-acetyl group had to move away to provide enough fitting space. Thus, we foresee that substitution of these residues by smaller polar amino acids could not only ease the accommodation of the *N*-acetyl

group in the pocket but also provide additional and beneficial hydrogen bond interactions to reshape the pocket for its specific recognition (**SUPPLEMENTARY FIGURES 1A** and **1C**).

In summary, 7 positions of subsite +1, alleged to be not critical for sucrose binding but beneficial for the glucosylation of target acceptors, have been selected for mutagenesis: 228, 289, 290, 330, 331, 394 and 446 (**FIGURE 41B**). Aware that the outcome of the amino acid substitution remains uncertain as it implies a new refolding of subsite +1, we mutated systematically the selected amino acids by the 19 other possible residues to create a small size library focused on subsite +1 of amylosucrase variants.

C.5. Detection of mutants able to glucosylate the target acceptors

Site-directed mutagenesis was preferred to saturation mutagenesis, as each monomutant generated is easily identified, isolated and characterized. A library of 133 monomutants (7x19) was thus obtained. Mutants were tested against acceptors **A'** and **D** in microtiter format experiments combined to HPLC screening. Both sucrose consumption and disaccharide formation were determined in order to calculate the glucosylation rate (**FIGURE 43**).

C.5.a. Target building block α -D-Glcp-(1→3)-L-RhapOMe [EA']

No glucosylated forms of **A'** were detected in our analysis showing thus that wtAS did not recognize **A'** (< 5 μ mol of product formed / min / g_{enzyme}). In contrast, 15 mutants displaying a totally novel specificity towards **A'** were isolated (**FIGURE 43A**). Comparison of HPLC and NMR data of the glucosylation products with those of chemically synthesized **EA'**^{35, 44} revealed that position 228 is crucial for **EA'** enzymatic synthesis. Of the 19 corresponding mutants, 11 formed the desired disaccharide. Mutations at position 290 were also beneficent since 4 mutants could form disaccharide **EA'**, albeit in lower yield ($\leq 5\%$). Mutations at positions 289, 330, 331, 394, and 446, failed to improve **A'** glucosylation.

C.5.b. Target building block α -D-Glcp-(1→4)-D-GlcpNAc [ED]

In screening conditions, D-GlcpNAc was poorly recognized by wild type AS (conversion rate $\sim 2\%$). However, remodelling of subsite +1 led to remarkable results. Indeed, 22 mutants catalyzed **ED** formation with a glucosylation rate comprised between 10 and 50 % for 17 mutants, and over 50 % for 5 mutants (**FIGURE 43B**). Clearly, position 290 appeared strategic to improve recognition of D-GlcpNAc. Notably, F290D and F290K yielded disaccharide **ED** with glucosylation rates over 90%, which represents a 45 fold increase compared to that of wtAS. Mutations at positions 228, 289, 331, and 446, also improved **ED** synthesis (not detailed).

Overall, two mutants were further characterized: I228Y yielding **EA'**, and F290K providing **ED**.

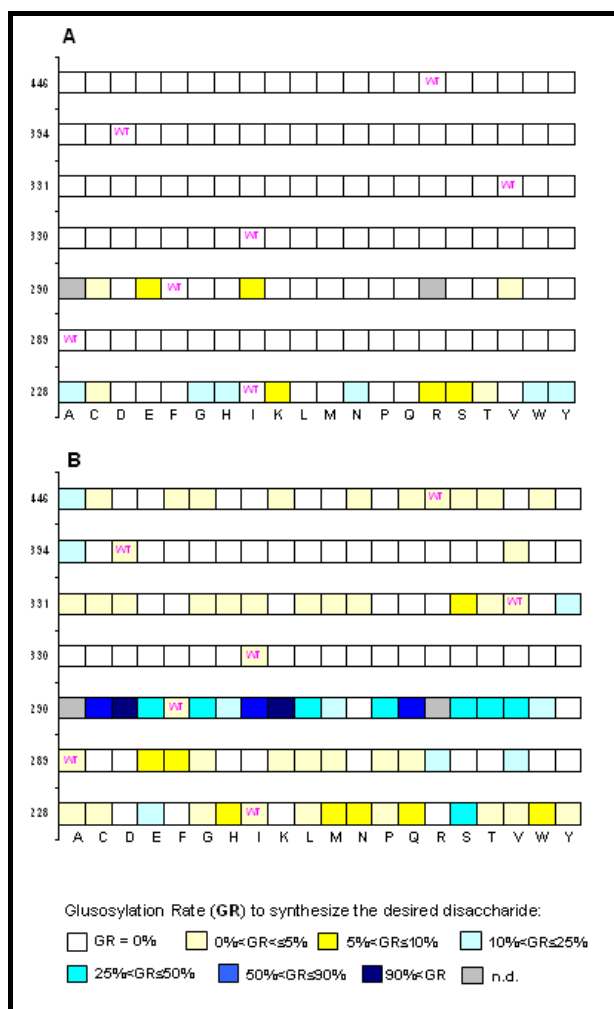


FIGURE 43| Screening of the library for their ability to synthesize the desired disaccharide: (A) α -D-Glc-(1→3)-L-Rhap-OMe (EA') and (B) α -D-Glc-(1→4)-D-GlcpNAc (ED).

Mutants were tested on both acceptors α -L-Rhap-OMe and D-GlcpNAc. Rows indicate the 7 mutated positions and columns represent the 20 possible amino acid mutations including the corresponding wt amino acid (indicated in magenta)

C.6. Characterization of I228Y and F290K mutants

Mutants I228Y and F290K were produced and purified to homogeneity to determine i) the profile of their acceptor reaction products and ii) their kinetic parameters.

In the absence of **A'** acceptor, I228Y did not synthesize high molecular weight amylose. Sucrose hydrolysis was the main reaction catalyzed and only traces of maltooligosaccharides of DP ranging from 3 to 5 were formed (**FIGURE 44**). Compared to the wild type AS, the transglucosylation activity onto glucose and longer maltooligosaccharides of I228Y mutant is clearly impaired. On the other hand, when **A'** was added to the reaction mixture, we observed that 47% of the glucosyl residues coming from sucrose were transferred onto the **A'** acceptor. The

disaccharide of interest (**EA'**) was formed with a remarkable yield of 93%. Meanwhile, addition of **A'** to the reaction mixture containing the wild type enzyme had no effect on the acceptor reaction profile and no trace of **EA'** could be detected within the limit of detection of the screening conditions used. Regarding the characterization of the products formed by F290K mutant, we observed that addition of **D'** acceptor totally suppressed maltooligosaccharide formation. More than 95% of the glucosyl residues were transferred onto **D'**, yielding 69 % of the disaccharide of interest **ED'**. Note that in the absence of **D'** acceptor, this mutant was still able to synthesize longer maltooligosaccharides (up to DP 20). Our assumption regarding the influence of the allyl aglycone in **D'** on the glucosylation rate and regiospecificity of glucosaminyl acceptors was thus valid.

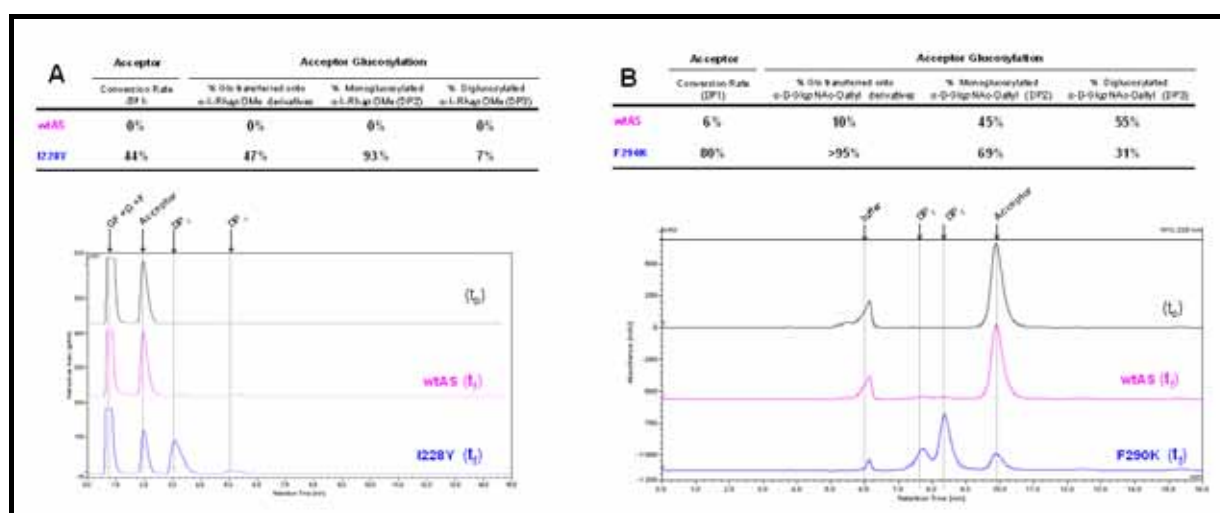


FIGURE 44| Transglucosylation of (a) α -L-RhapOMe (A'**) and (b) α -D-GlcpNAcOallyl (**D**) using the most improved variants I228Y and F290K.**

A) α -L-RhapOMe glucosylation and corresponding HPLC chromatogram (with RI detection) comparing I228Y (blue) and wtAS (magenta).

B) α -D-GlcpNAcOallyl glucosylation and corresponding HPLC chromatogram (with $UV_{\lambda=220nm}$ detection) comparing F290K (blue) and wtAS (magenta).

Initial reaction conditions: Sucrose=Acceptor=146mM. At final time, sucrose was fully consumed. Conversion Rate = $(Q(\text{Acceptor})_t - Q(\text{Acceptor})_0) / Q(\text{Acceptor})_0$ where $Q(X)$ = Quantity of X in moles. % Glc transferred onto acceptor derivatives = $Q(\text{Glucosyl units transferred onto acceptor derivatives}) / Q(\text{Glucosyl units transferrable from initial sucrose})$. % Monoglucosylated acceptor = $Q(\text{Monoglucosylated acceptor}) / Q(\text{acceptor derivatives})$. % Diglucosylated acceptor = $Q(\text{Diglucosylated acceptor}) / Q(\text{acceptor derivatives})$. G=Glucose; F=Fructose; GF=Sucrose; DP1= Acceptor; DP2= Acceptor monoglucosylated; DP3= Acceptor diglucosylated.

Kinetic measurements showed that both I228Y and F290K could not be saturated with sucrose, indicating thus a poor affinity for sucrose. k_{cat}/K_M values were then determined by linear regression both in the presence and in the absence of acceptor (**SUPPLEMENTARY FIGURES 2-5** and **TABLE 10**). In the presence of sucrose alone, k_{cat}/K_M values decreased by 100 and 40 fold for I228Y and F290K mutants, respectively. In the presence of the acceptors (**A'** for I228Y and **D'** for

F290K), we observe, for both mutants, a formidable increase of catalytic efficiency towards both sucrose and the acceptor. Such an improvement can be attributed to the high specificity of each mutant for the **EA'** and **ED'** formation, respectively. The adaptation of the acceptor binding site to each acceptor enhanced the rate of the de-glucosylation step, which is no more the limiting step of the reaction. For I228Y, the gain is spectacular, but could not be calculated as the wild type enzyme did not catalyze at all the synthesis of **EA'**. In the case of F290K, a remarkable 130 fold increase of the efficiency of the mutant compared to **wtAS** was observed and the expected regioselectivity was achieved.

TABLE 10| Determination of kinetic parameters of the improved variants I228Y and F290K compared to wtAS.

¹ Catalytic efficiency k_{cat}/K_M (GF) was determined by measurement of sucrose (GF) consumption.

² Catalytic efficiency k_{cat}/K_M (GF) was determined by the formation of the desired disaccharide at a constant concentration (250mM) of acceptor.

³ Catalytic efficiency k_{cat}/K_M (Acc) was determined by the formation of the desired disaccharide at a constant concentration (250mM) of sucrose.

⁴ Specific activity corresponds to the amount of micromole of desired product formed per min per g of enzyme at a concentration of 250 mM of both donor and acceptor.

| Enzymes | Kinetic parameters without acceptor | Kinetic parameters with acceptor | | Specific activity of disaccharide formation ⁴ |
|--------------|--|--|---|--|
| | k_{cat}/K_M (GF) ¹ (mM ⁻¹ min ⁻¹) | k_{cat}/K_M (GF) ² (mM ⁻¹ min ⁻¹) | k_{cat}/K_M (Acc) ³ (mM ⁻¹ min ⁻¹) | |
| wtAS | 1,54 | 0 | 0 | <5 |
| I228Y | 0,017 | 0,080 | 0,085 | 210 |
| wtAS | 1,54 | 0,074 | 0,122 | 287 |
| F290K | 0,038 | 3,630 | 15,880 | 6690 |

C.7. Structural insight on mutation effects

All 7 targeted positions are located on flexible loops that connect β -strands with α -helices of the $(\beta/\alpha)_8$ catalytic barrel that constitutes the *N. polysaccharea* amylosucrase fold (**SUPPLEMENTARY FIGURE 6**), and encircling subsite +1 (**FIGURE 41B**). Interestingly, positions 290 and 228 are in close vicinity although they belong to distinct regions of the structure (C and B domains, respectively). Screening analysis of monomutants indicated that both positions 290 and 228 can tolerate a wide range of substitutions of distinct nature, and still result in an increased glucosylation rate of **A'** and **D**, suggesting thereby multiple causes to the observed enhancement (**FIGURE 43**). Docking of **ED** and **ED'** indicated a favourable binding of the disaccharide product in the active site of F290K mutant compared to **wtAS**, with a stabilization of the predicted free energy of binding by

up to 3 kcal/mol. However, docking of EA' in I228Y appeared more tricky without the rearrangement of several residues in the vicinity of I228, including F290. This suggests that an exchange at either 228 or 290 position might affect the other position. Noteworthy, position 290 is part of a conserved region of the $(\beta/\alpha)_8$ catalytic barrel, known as region II in α -amylases, and thought to be involved in GH13 amylase specificity⁴⁵. On the other hand, position 228 belongs to domain B which is formed by the domain A extension loop 3, known to be involved in reaction and substrate specificity in GH13. Compared to alpha-amylases, domain B of amylosucrase is larger and together with domain B', they cover the active site of the enzyme to confer the enzyme active site a pocket architecture unlike the large cleft displayed by hydrolases of family 13. As one of the residues forming the interface between domains B and B' and covering up the active site, position 228 appears critical to confer the active site pocket topology and thus the glucosyl transfer specificity of wtAS.

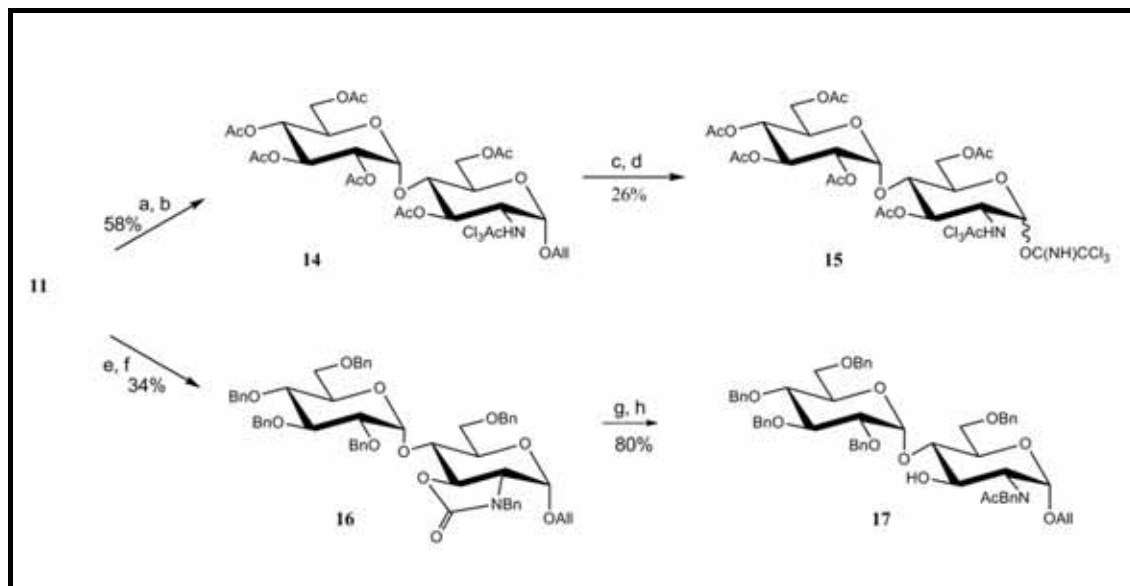
C.8. EA' as starting material for the synthesis of *S. flexneri* building blocks

To synthesize *S. flexneri* oligosaccharides bearing the EA branching pattern, EA' had to be converted to an acceptor compatible with chain elongation at position 3_{A'}, and/or a donor allowing chain elongation at the reducing end. A three-step process provided EA' (**2**) from A' (**1**) in 16% overall yield: (i) obtaining of **2** contaminated by fructose by column chromatography of the crude enzymatic glucosylation mixture, (ii) *O*-acetylation to give pure **3**, and (iii) transesterification of **3** (SCHEME 1). Alternately, compound **3** (17% yield from A'), was converted into building block³⁶ **10** (SCHEME 1, data not shown). Briefly, **3** was acetylated into **4** (90%), then turned into bromide **5**, which was subsequently cyclized, transesterified, and benzylated into orthoester **6** (one purification from **4**, 74%). Regioselective opening of **6**, and subsequent deacetylation provided the target **7** (79%).

C.9. ED' as starting material for the synthesis of *S. flexneri* building blocks

Isolation of ED' (**13**) was best performed following rough chromatography of the crude enzymatic glucosylation mixture issued from D' (**8**), peracetylation into pure **11**, then transesterification of intermediate **11** (30% from **8**) (SCHEME 2). To access *S. flexneri* oligosaccharides bearing the ED branching pattern, the product of D' enzymatic glucosylation, disaccharide **10**, had to be turned into a donor allowing chain elongation at the reducing end **15** and/or an acceptor allowing chain extension at position 3_{D'} (**17**) (SCHEME 3). Therefore, selective *N*-trichloroacetylation and per-*O*-acetylation of **11** gave **14** (58%), which was converted to trichloroacetimidate **15** (26%) bearing a participating group at position 2_D as required. Alternatively, **11** was saponified into **12** (83%)

(SCHEME 2), which was turned into the benzylated 2_D,3_D-oxazolidinone **16** via **13** (34%). Oxazolidinone cleavage and subsequent *N*-acetylation gave acceptor **17** (80%) bearing a free hydroxyl group at position 3 and a participating group at position 2.



SCHEME 3 | Conversion of disaccharide 10 in key building blocks 15 or 17.

a. $(\text{Cl}_3\text{Ac})_2\text{O}$, NaOMe; b. Ac_2O , Pyridine; c. cat. $[\text{Ir}(\text{COD})[\text{PCH}_2(\text{C}_6\text{H}_5)_2]_2]^+\text{PF}_6^-$, THF; d. CCl_3CN , DBU; e. $\text{Ba}(\text{OH})_2$, H_2O ; f. (i) $\text{CCl}_3\text{CH}_2\text{OCOCI}$, NaOMe, (ii) BnBr, NaH; g. 1M aq NaOH, 1,4-dioxane; Ac_2O , Pyridine

D. Discussion

Surprisingly, the quite challenging design of an appropriate enzymatic glycosylation tool that would allow an optimal combination of the chemical and enzymatic steps involved in the synthesis of complex oligosaccharides has never been attempted, although it would be of major interest to develop new synthetic pathways. To demonstrate the efficiency of such a global chemo-enzymatic pathway, we concentrate our effort on the synthesis of *S. flexneri* O-antigen oligosaccharides, which are of prime interest for the synthesis of third generation vaccines against shigellose. We reasoned that enzymatic glycosylation could overcome the poor α/β stereospecificity of the chemical glycosylation process. The use of the native Leloir glycosyltransferases was envisioned, but rapidly found inappropriate as they are not easily available and involve nucleoside diphosphate monosaccharide donors¹³. In addition, no known transglucosidases are able to mimic the serotype specific *Shigella* regioselective glycosylation of complex oligosaccharide acceptors⁴⁶. Consequently, we considered that the enzymatic glycosylation of appropriate monosaccharide acceptors could open the way to disaccharide building blocks compatible with chain elongation at both the reducing and non reducing ends able to act as donors as well as acceptors. To produce these building blocks, we turned our choice toward original α -transglucosidases, which naturally synthesize glucans from sucrose, a largely

available substrate. Then, the remaining major challenge consisted in setting up the appropriate transglucosidase/donor/acceptor combinations. Non-natural lightly protected acceptors, which may not be recognized by available native enzymes, were anticipated to be needed. This issue was characteristic of an *a priori* critical limitation. Indeed, among the wild type recombinant glucansucrases tested, none had yet been considered for the glycosylation of **A'** or **D'**, the starting building blocks envisaged for the synthesis of serotype 3a and 1b pentasaccharides. Amylosucrase from *Neisseria polysaccharea*, an enzyme of glycoside-hydrolases family 13, was found to be the only enzyme that recognized **D** to synthesize **ED**, although with a dramatically low yield, and failed to recognize **A'** as an acceptor³⁹.

At this stage and given the poor recognition of the target acceptors by wild type recombinant glucansucrases, a strategy of protein engineering was considered to re-design glucansucrase acceptor substrate specificity. Rational and combinatorial engineering has proven to be extremely efficient to generate biocatalysts, including carbohydrate-active enzymes, with improved properties and novel substrate specificities^{21, 22}. Glycosyltransferase substrate specificity has been modified by site-directed mutagenesis assisted by computational modelling or directed evolution for the synthesis of biologically relevant structures²². Promiscuous glycosidases showing altered and new specificities towards donor and acceptor sugars have been generated⁴⁷⁻⁵⁰. Engineering of new glycosynthases from β -glycosidases also emerged as a powerful way to generate modified transferases which use fluoride donors^{51,52, 53}. However, only one active glycosynthase derived from an α -retaining enzyme has been described up to now⁵⁴, and there was yet no description of substrate acceptance enlargement for α -retaining transglycosidases such as glucansucrases. Consequently, the re-designing of such enzymes was also quite challenging.

A fully combinatorial approach based on directed molecular evolution techniques could have been envisaged but it is now admitted that without any rational data, the size of the library to be screened for the evolution of enzyme substrate specificity is exorbitant and high throughput screens and selections would need to be set up. The absence of fluorescence property linked to the acceptor reaction of interest prevented in our case any use of elegant and powerful screening methods based on FACS or microfluidic devices to isolate the correct variant^{55,56}. Thus, we opted for a structure-based engineering strategy relying on a combination of rational and combinatorial engineering approaches. *N. polysaccharea* amylosucrase was naturally chosen as candidate enzyme to apply our strategy, taking advantage of its available crystal structure^{44,41,42}. The difficulty of engineering was double : modified amylosucrase had not only to keep its specificity toward the donor (sucrose) which binds at subsites -1 and +1 but also accept at subsite +1 the acceptor of interest once the glucosyl-enzyme intermediate is formed. Our computational modelling study enabled us to select 7 positions from the first shell of residues centred on subsite

+1 (**FIGURE 41B**). Systematic changes of these amino acids by site-directed mutagenesis rapidly led to a small size library of 133 variants. Notably, 50 % of these variants kept their ability to use sucrose. Without rescuing to high throughput screening capabilities, our reasoned strategy led to the identification of 43 mutants of interest, thus demonstrating the efficiency of the approach. In particular, two spatially neighbouring positions (228 and 290) were identified as playing a major role in the glucosylation of the acceptors of interest with the correct regiospecificity and a very high selectivity. A totally new specificity for the glucosylation of **A'** was observed with mutant I228Y and an exceptionally high improvement ($\times 130$) of **D'** glucosylation was also obtained with mutant F290K. To our knowledge, this is the first engineering of amylosucrase, that led to such improvement of non-natural acceptor glucosylation and this is also the first creation of a new specificity toward non-natural acceptor substrates for α -transglucosidases.

To spare time and foresee the scale-up of the enzymatic process, a brut extract of these mutants was employed to synthesize the starting building blocks. They served to synthesize gram amounts of the selected disaccharides **EA'** and **ED'**, which were best purified as the corresponding per-acetylated **3** (17% from **A'**) and **11** (30% from **D'**), respectively (**SUPPLEMENTARY RESULTS 1**). The efficient transformation, 7 steps requiring 3 purifications (53% overall yield), of intermediate **3** into the key building block **7** validates the proof of concept at best. Indeed, as a whole, this newly disclosed chemo-enzymatic protocol to **7** comprises 1 enzymatic step involving easily available low cost donor and acceptor, and 8 additional chemical steps. It compares favourably with the full chemical synthesis of **7** (11 steps) since the number of chemical steps is lowered and most importantly the limiting chemical glucosylation step³⁶ (57%) is avoided. Besides, having demonstrated that AS mutants bearing one single mutation could turn **A'** into **EA'**, we are confident that additional enzyme engineering will result in improved enzymatic conversion. The 17% conversion yield of **1** into **3** obtained so far is thus not seen as a limiting step. The demonstration that monomutants of the same wtAS could efficiently convert **D'** into **ED'** (30%) is an additional argument, if needed, ascertaining the potential of the concept. In support to the overall strategy, the chemical transformation of **11** into donor **15** (4 steps) and acceptor **17** (5 steps) was shown feasible. This innovative chemo-enzymatic process may thus be seen as an alternative to that proposed thus far for the synthesis of *S. flexneri* 1a O-antigen fragments, which relies on the sophisticated prearranged-glycosides mediated intramolecular glycosylation strategy which was developed to ensure the correct stereochemical outcome of the glucosylation step⁵⁷.

Well-illustrating the potential of enzyme semi-rational engineering, this is to our knowledge the first report on α -transglucosidase engineering resulting in both a tremendous enhancement of a non-natural acceptor glucosylation and a totally new specificity toward another non-natural acceptor. Applied to other α -retaining transglycosidases, the approach should take advantage of

the natural promiscuity generally encountered in these enzyme families. Besides, mutants selected in this study could also transfer to allyl α -L-rhamnopyranoside, 2-*N*-trichloroacetyl-D-glucosamine, methyl 2-acetamido-2-deoxy- α -D-glucopyranoside, methyl 2-acetamido-2-deoxy- β -D-glucopyranoside, and allyl β -D-glucopyranoside (**SUPPLEMENTARY RESULTS 2**), thus offering additional alternatives. The concept validation disclosed herein offers new opportunities in the synthesis of complex oligosaccharides providing that appropriate combinations of donors, acceptors and transglucosidases are defined. For example, enzymes disclosed herein might be applicable to the synthesis of selected fragments of *Streptococcus pneumoniae* type 2 capsular polysaccharide, or *Klebsiella* K-antigen type K18 and K41, as well as *S. flexneri* 5a and 5b, which all bear the α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranosyl moiety in their linear backbone⁵⁸. Overall, this innovative concept is applicable to the construction of any complex glycosidic linkage, provided that an appropriate combination of engineered transglycosidase/donor/acceptor system is made available based on compatibility with chemical chain elongation.

E. Methods

E.1. Bacterial strains, plasmids and chemicals

Plasmid pGST-AS, derived from the pGEX-6P-3 (GE Healthcare Biosciences) and containing the *N. polysaccharea* amylosucrase encoding gene⁵⁹ was used for the construction of the AS single mutant library. *E. coli* JM109 was used as host for the plasmid library transformation, gene expression and large-scale production of the selected mutants. Sucrose, *N*-acetyl-D-glucosamine and glycogen were purchased from Sigma-Aldrich (Saint-Louis, MO, USA). Rhamnose was purchased from Carbosynth (Newbury, Berkshire, UK). Acceptor and **8**⁶⁰, were prepared according to known protocols. Reference disaccharides α -D-Glcp-(1 \rightarrow 3)- α -L-Rhap-OMe³⁵, α -D-Glcp-(1 \rightarrow 3)- α -L-Rhap-OMe⁶¹ and α -D-Glcp-(1 \rightarrow 4)- α -L-Rhap-OMe⁶² were previously synthesized at the UCB. Disaccharides α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc and Glcp-(1 \rightarrow 6)-D-GlcpNAc were enzymatically synthesized and characterized at LISBP³⁹. Ampicillin, lysozyme and isopropyl- β -D-thiogalactopyranoside (IPTG) were purchased from Euromedex (Souffelweyersheim, France), and *DpnI* restriction enzyme from New England Biolabs (Beverly, MA, USA). Oligonucleotides were synthesised by Eurogenetec (Liege, Belgium). DNA extraction (QIASpin) and purification (QIAQuick) columns were purchased from Qiagen (Chatsworth, CA).

E.2. Molecular modelling

Disaccharides **EA'** and α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc were constructed with monosaccharides obtained from a database of carbohydrates three-dimensional structures. All molecular modelling

calculations were performed using the SYBYL 7.3 software (TRIPOS inc, Saint Louis, USA). The coordinates of amylosucrase were taken from the 2.0 Å resolution crystal structures of amylosucrase from *N. polysaccharea* in complex with sucrose (PDB: 1JGI) and maltoheptaose, a reaction product (PDB: 1MW0). All hydrogen atoms were added to the enzyme and their position optimized with the Tripos force field. Disaccharide EA' and both anomers of α -D-Glcp-(1→4)-D-GlcpNAc, were subjected to a systematic grid search study of the glycosidic linkage conformation. Starting from minimized disaccharides, a two-dimensional systematic conformational search was performed by rotating the two torsion angles defining the glycosidic linkages, Φ and Ψ by 20° steps: $\Phi = O5'-C1'-O3-C3'$ and $\Psi = C1'-O3-C3'-C2'$ for EA' and $\Phi = O5'-C1'-O4-C4$ and $\Psi = C1'-O4-C4-C3$ for α -D-Glcp-(1→4)-D-GlcpNAc. The MM3 force field⁶³ implemented in SYBYL 7.3 software was used for this purpose together with the energy parameters appropriate for carbohydrates. The geometries were optimized at each point of the grid with the driver option that keeps fixed the atoms defining the torsion angles. The solvent specific relaxed conformational maps (data not shown) obtained for all disaccharides were then used to locate the different energy minima that were subsequently fully relaxed. The lowest energy conformations identified on the disaccharides potential energy maps were used as starting structures to be docked in the binding site of amylosucrase. This was performed by superimposing the glucosyl unit of the disaccharides at subsite -1 onto the glucosyl unit of the crystallographic maltoheptaose. Each of these AS:disaccharide complexes was optimized by means of the appropriate energy parameters. The annealing method implemented in SYBYL 7.3 software was used to optimize the complexes. Two shells of amino acids were considered: a 12Å shell centred on the binding site was taken into account for the energy calculations. A 6Å shell region closest to the carbohydrate was defined as the hot region to be optimized. The position of all atoms included in this region was optimized using Powell's method

E.3. Construction of mutant library

Single mutagenesis, focused on subsite +1 amino acids selected from ligand docking, was carried out with the QuickChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions, and using pGST-AS G537D as vector template. It was checked that this mutation had no impact on the native enzyme catalytic properties. The complementary primers listed below were used to obtain the single mutant library (TABLE 11). XXX codon indicates the bases which were used to obtain the replacement by the desired amino acids and are listed in TABLE 12.

PCR amplification was carried out with Pfu DNA polymerase (2.5 U) for 16 cycles (95°C, 30s; 55°C, 30s; 72°C, 12min). The DNA was digested with DpnI to eliminate methylated parental

template and purified using Qiaquick spin column, following manufacturer's recommendations. *E. coli* JM109 was transformed with the plasmid and plated on LB agar supplemented with 100µg/mL ampicillin. For each construction, two clones were isolated and their corresponding plasmids stored at -20°C. 17 mutants (I228A₁, I228V₁, I228Y₁, A289D₁, F290D₁, F290K₁, F290Q₁, I330A₁, I330D₁, I330E₁, I330F₁, I330T₁, I330W₁, V331A₁, V331S₁, D394V₁ and R446K₁) were sequenced on the entire gene by Millegen (Labège, France) or Cogenics (Meylan, France) and showed no other mutations.

TABLE 11 | Primers used to generate the 19 monomutants for the positions 228, 289, 290, 330, 331, 394 and 446.

| Primer Name | Nucleotide Sequence |
|-------------|--|
| I228for | 5'-ACC CTG CGC GAA XXX TTC CCC GAC CAG CA-3' |
| I228rev | 5'-TG CTG GTC GGG GAA XXX TTC GCG CAG GGT-3' |
| A289for | 5'-T ATG GAT GCG GTT XXX TTT ATT TGG AAA CAA AT-3' |
| A289rev | 5'-AT TTG TTT CCA AAT AAA XXX AAC CGC ATC CAT A-3' |
| F290for | 5'-T ATG GAT GCG GTT GCC XXX ATT TGG AAA CAA AT-3' |
| F290rev | 5'-AT TTG TTT CCA AAT XXX GGC AAC CGC ATC CAT A-3' |
| I330for | 5'-TC AAA TCC GAA GCC XXX GTC CAC CCC GAC CAA GT-3' |
| I330rev | 5'-AC TTG GTC GGG GTG GAC GAT XXX TTC GGA TTT GA-3' |
| V331for | 5'-TC AAA TCC GAA GCC ATC XXX CAC CCC GAC CAA GT-3' |
| V331rev | 5'-AC TTG GTC GGG GTG XXX GAT GGC TTC GGA TTT GA-3' |
| D394for | 5'-TC CGC AGC CAC GAC XXX ATC GGC TGG ACG TTT-3' |
| D394rev | 5'-AAA CGT CCA GCC GAT XXX GTC GTG GCT GCG GA-3' |
| R446for | 5'-ACA GGC GAC TGC XXX GTC AGT GGT ACA-3' |
| R446rev | 5'-TGT ACC ACT GAC XXX GCA GTC GCC TGT-3' |

TABLE 12| Sequence of XXX codon used to replace each selected amino acids by the 19 other ones

| | Ala | Cys | Asp | Glu | Phe | Gly | His | Ile | Lys | Leu | Met | Asn | Pro | Gln | Arg | Ser | Thr | Val | Trp | Tyr | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 228 | for | GCC | TGC | GAC | GAG | TTC | GGC | CAC | wt | AAG | CTC | ATG | AAC | CCC | CAG | CGC | AGC | ACC | GTC | TGG | TAC |
| | rev | GGC | GCA | GTC | CTC | GAA | GCC | GTG | wt | CTT | GAG | CAT | GTT | GGG | CTG | GCG | GCT | GGT | GAC | CCA | GTA |
| 289 | for | wt | TGC | GAC | GAA | TTC | GGC | CAC | ATC | AAA | CTC | ATG | AAC | CCC | CAA | CGC | AGC | ACC | GTC | TGG | TAC |
| | rev | wt | GCA | GTC | TTC | GAA | GCC | GTG | GAT | TTT | GAG | CAT | GTT | GGG | TTG | GCG | GCT | GGT | GAC | CCA | GTA |
| 290 | for | GCT | TGT | GAT | GAA | wt | GGT | CAT | ATT | AAA | CTT | ATG | AAT | CCT | CAA | CGT | TCT | ACT | GTT | TGG | TAT |
| | rev | AGC | ACA | ATC | TTC | wt | ACC | ATG | AAT | TTT | AAG | CAT | ATT | AGG | TTG | ACG | AGA | AGT | AAC | CCA | ATA |
| 330 | for | GCC | TGC | GAC | GAA | TTC | GGC | CAC | wt | AAA | CTC | ATG | AAC | CCC | CAA | CGT | TCT | ACC | GTC | TGG | TAC |
| | rev | GGC | GCA | GTC | TTC | GAA | GCC | GTG | wt | TTT | GAG | CAT | GTT | GGG | TTG | ACG | AGA | GGT | GAC | CCA | GTA |
| 331 | for | GCC | TGC | GAC | GAA | TTC | GGC | CAT | ATC | AAA | CTC | ATG | AAC | CCC | CAA | CGC | TCC | ACC | wt | TGG | TAC |
| | rev | GGC | GCA | GTC | TTC | GAA | GCC | ATG | GAT | TTT | GAG | CAT | GTT | GGG | TTG | GCG | GGA | GGT | wt | CCA | GTA |
| 394 | for | GCC | TGC | wt | GAG | TTC | GGC | CAC | ATC | AAA | CTC | ATG | AAC | CCC | CAA | CGC | AGC | ACC | GTC | TGG | TAC |
| | rev | GGC | GCA | wt | CTC | GAA | GCC | GTG | GAT | TTT | GAG | CAT | GTT | GGG | TTG | GCG | GCT | GGT | GAC | CCA | GTA |
| 446 | for | GCT | TGT | GAT | GGT | TTT | GGT | CAT | ATT | AAA | CTT | ATG | AAT | CCT | CAA | wt | AGT | ACT | GTT | TGG | TAT |
| | rev | AGC | ACA | ATC | ACC | AAA | ACC | ATG | AAT | TTT | AAG | CAT | ATT | AGG | TTG | wt | ACT | AGT | AAC | CCA | ATA |

E.4. Expression of mutant library

The protocol was established to enable the rapid identification of (i) clones able to glucosylate α -L-RhaOMe (**A'**), and (ii) clones for which D-GlcNAcp (**D**) glucosylation was improved. To obtain larger amounts of enzymes and facilitate detection of glucosylated compounds by HPLC, the mutants were produced in 96-DeepWell Format plates. Storage microplates containing monomutants were thawed and replicated to inoculate a starter culture in 96-well microplates containing, in each well, 150 μ L LB medium supplemented with ampicillin (100 μ g/mL). After growth for 24h at 30°C under agitation (200 rpm), plates were duplicated into 96-Deep Well plates containing, in each well, 1.1 mL LB medium supplemented with ampicillin (100 μ g/mL) and IPTG (1mM) to induce GST-AS expression. Cultures were then grown for 24h at 30°C under agitation (200 rpm). Plates were centrifuged (20 min, 3000g, 4°C) and the supernatant was removed. The cell pellet was resuspended in 200 μ L of lysozyme (0.5 mg/mL), followed by freezing at -80°C for 8 to 12h. After thawing at room temperature, 100 μ L of sucrose and 100 μ L of acceptor (each at a final concentration of 73 mM) were added to each well. Enzymatic reaction was incubated at 30°C during 24h under agitation. The DeepWell plates were then centrifuged (20 min, 3000g, 4°C) and 300 μ L of the supernatant was transferred to a filter micro-plate (Glass fiber membrane, PS, 0.25 mm pore) to be clarified. Supernatant filtration was carried out by centrifugation of the filter micro-plate (5 min, 2000g, 4°C) into a novel microplate for screening.

E.5. Screening of mutant library

Efficiency of the glucosylation reaction was evaluated by HPLC analysis of the acceptor reaction product synthesized when using either **A'** or **D** as acceptors using a Dionex P 680 series pump, a Shodex RI 101 series refractometer, a Dionex UVD 340 UV/Vis detector and an autosampler HTC PAL. HPLC analyses were performed using two columns: (1) a Biorad HPLC Carbohydrate Analysis column (HPX-87K column (300x7.8 mm)) maintained at 65°C, using ultra-pure water as eluent with a flow rate of 0.6 mL/min (2) a reversed phase analytical column (Synergi C18RPFusion, 4 μ m, 30x4.6 mm) kept at room temperature, with 1ml/min of ultra-pure water as eluent. HPX-87K column was used to determine sucrose consumption by RI detection and concomitant α -D-Glcp-(1 \rightarrow 4)-D-GlcpNAc formation by UV_{=220nm} detection. C18RPFusion column served to detect EA' formation.

E.6. Characterization of the Selected Variants: I228Y and F290K

E.6.a. Production and Purification of I228Y and F290K.

Production and purification of AS variants were performed as previously described⁵⁹. Since pure GST/AS fusion protein possesses the same function and the same efficiency as pure AS, enzymes were purified to the GST/AS fusion protein stage (96 kDa). The enzymes were desalted by size exclusion chromatography using P6DG columns (GE Healthcare Biosciences) at +4°C and stored in elution buffer (50 mM Tris-HCl, pH 7.0, 150 mM NaCl) at -80°C. The protein content was determined by micro-Bradford method, using bovine serum albumin as standard⁶⁴.

E.6.b. Biochemical Characterization of I228Y and F290K.

All assays were performed at 30°C in 50 mM Tris buffer, pH = 7.0. I228Y was tested on **A'**, F290K on **D'**, and **wtAS** on both **A'** and **D'** acceptors.

Specific activity of the purified enzymes was determined by measuring the initial rate of released fructose under standard conditions (146mM sucrose). Fructose concentration was determined using the dinitrosalicylic acid (DNS) method⁶⁵. One unit of AS variant corresponds to the amount of enzyme that catalyzes the production of 1µmole fructose per minute in the assay conditions.

Comparison of products synthesized by wild-type and AS variants.

Reactions were performed in the presence of 146 mM sucrose alone or supplemented with 146 mM acceptor. The purified wild-type or mutated GST/AS were employed at 0.5 U/mL. The reactions were stopped by heating at 95°C for 5 min. The final mixture was centrifuged at 18 000g for 10 min and filtered on 0.22 µm membrane. In addition to HPLC analyses previously described, different carbohydrate analyses were also performed to compare the product profiles synthesized by wild-type and AS variants (I228Y and F290K): soluble and insoluble oligosaccharides produced during the reaction were identified by HPAEC using a Dionex Carbo-Pack PA100 column at 30°C. Before analysis, the insoluble fraction was solubilized in KOH at a final total sugar concentration of 10 g/kg. Mobile phase (150mM NaOH) was set at 1mL/min flow rate with a sodium acetate gradient (going from 6 to 500 mM within 120 min). Detection was performed using a Dionex ED40 module with a gold working electrode and an Ag/AgCl pH reference. Note that **A'**, **D'** (acceptors) and their derivatives are not oxidable products and thus are not detectable by HPAEC. Sucrose, glucose, fructose, **A'**, **D'** (acceptors) and their derivatives (glucosylation products) were quantified by HPLC, as previously described.

Determination of kinetic parameters.

Enzyme assays were carried out in a total volume of 2 mL containing pure enzyme (0.115 mg, 2.600 mg and 0.106 mg when using **wtAS**, I228Y and F290K, respectively). Catalytic efficiency (k_{cat}/K_M) of **wtAS** and AS variants was determined with (i) sucrose alone and with (ii) both sucrose and acceptor **A'** (or **D'**) as variable substrates. In the presence of sucrose alone (GF varying between 0 and 600 mM), k_{cat}/K_M (GF) values were determined by measurement of the released fructose. The fructose formation reflects the consumption of sucrose. In the presence of sucrose and acceptor, both catalytic efficiencies, k_{cat}/K_M (GF) and k_{cat}/K_M (**A'**) (or k_{cat}/K_M (**D'**)) values were determined by the formation of the desired disaccharide (called Vi(Dp2)). For the determination of k_{cat}/K_M (GF), **A'** (or **D'**) was held constant at 250 mM and GF was varied between 0 and 600 mM. For the determination of k_{cat}/K_M (**A'**) (or k_{cat}/K_M (**D'**)), GF was held constant at 250 mM and **A'** (or **D'**) was varied between 0 and 250 mM. Experiments were performed until sucrose and **A'** (or **D'**) solubility limits were reached

Initial velocities were fitted to the Michaelis-Menten equation using Sigma-Plot. When saturation was not achieved with the mutant, efficiency was calculated by linear regression analysis of the velocity versus substrate concentration plot. Aliquots (200 μ L) were removed between 0 and 60 min (at which time product formation was still linear with respect to time), heated at 95°C for 5 min and centrifuged at 18 000g for 5 min. The final mixtures were filtered on a 0.22 μ m membrane and analyzed using HPLC material previously described.

E.7. Preparative synthesis of the acceptor reaction product.

In order to characterize the products of D-GlcNAc β glucosylation, a synthesis was carried out at the preparative scale. 10 ml mixture containing 146 mM sucrose, 146 mM D-GlcNAc β (**D**) and 1.5 U/ml of purified F290K AS mutant were incubated at 30°C for 24h. Then, the reaction mixture was centrifuged (4800 rpm, 20 min, 4°C) to remove proteins and filtered on a 0.22 μ m membrane. Mono- and di-glucosylated acceptor reaction products were separated by preparative chromatography on C18 reverse-phase chromatography column (Bischoff Chromatography). Ultra pure water was used as eluent at a constant flow rate of 50 mL/min. Glucosyl detection was carried out with a refractometer (Bischoff) and each peak was collected separately, concentrated and reinjected into the analytical HPLC system described above, to ascertain purity.

Oligosaccharides issued from **D'** glucosylation were produced in a 100 mL batch (146 mM sucrose + 146 mM **D'**), using 4 U/mL of non-purified F290K extract (sonication supernatant), whereas **EA'** was synthesized in a 500 mL batch (146 mM sucrose + 100 mM **A'**) using 0.25 U/ml of non-purified I228Y extract.

F. Supplementary Results

F.1. Supplementary Results 1 : Scale-up of the enzymatic processes

EA' was synthesized in a 500 mL batch using 0.25 U/mL of non-purified I228Y extract. The low activity of this mutant did not allow a full consumption of sucrose and consequently, affected the glucosylation yield (17%). Concentration of purified I228Y activity increased significantly this yield (47%) (SEE C.6.) but was not adapted for the scale-up of the enzymatic process.

Disaccharide and trisaccharide issued from D' glucosylation were produced in a 100 mL batch, using 4 U/mL of non-purified F290K extract (sonication supernatant). These oligosaccharides were obtained in the same proportions (30% of glucosylation yield for each). The use of this less active mutant reduced the formation of the co-product (i.e. trisaccharide) as shown in FIGURE 45 (SEE also C.6.).

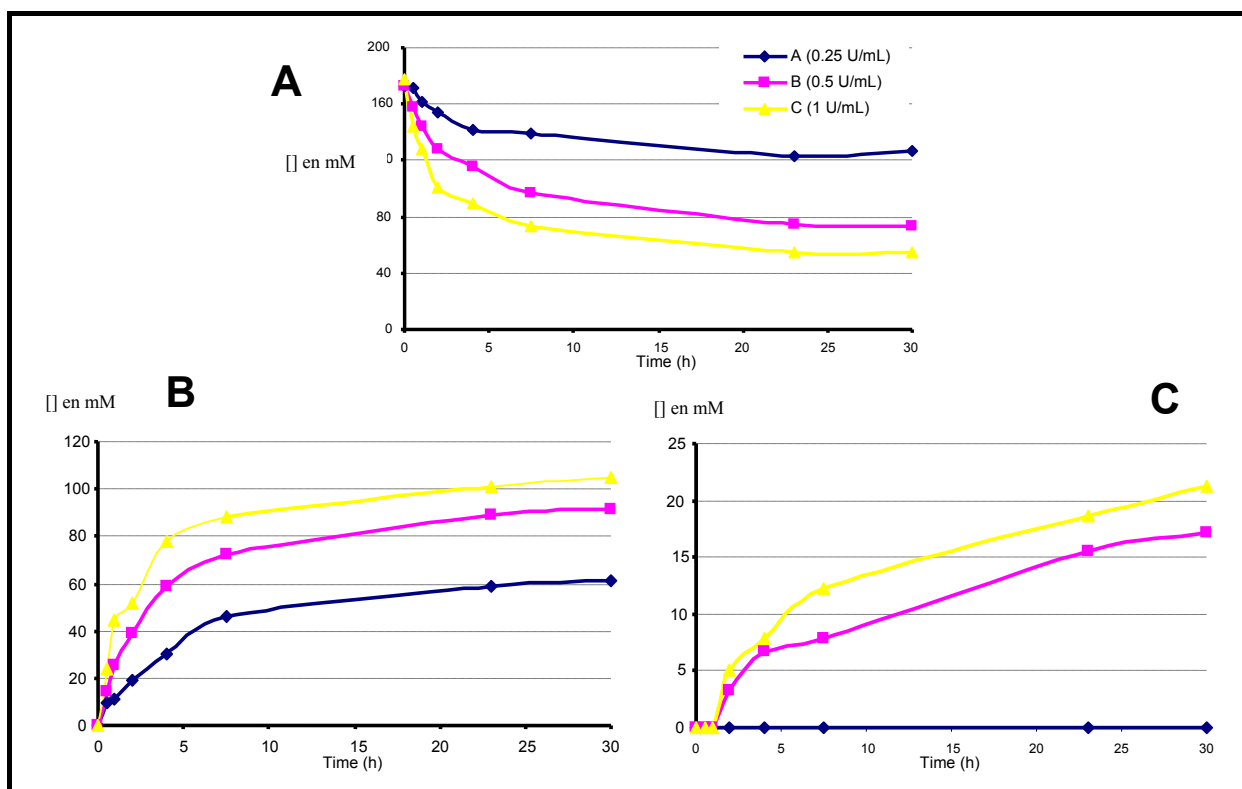


FIGURE 45 | Glucosylation Rates of α -D-GlcpNAc-Oallyl at different times using 0.25, 0.5 or 1 U/mL of non-purified F290K extract.

- (A) Consumption of α -D-GlcpNAc-Oallyl
- (B) Formation of mono-glucosylated α -D-GlcpNAc-Oallyl.
- (C) Formation di-glucosylated α -D-GlcpNAc-Oallyl.

F.2. Supplementary Results 2

F.2.a. Enzymatic glucosylation of L-Rhap and α -L-RhapOAll derivative

The 7 mutants I228A, I228G, I228H, I228N, I228S, I228W and I228Y were tested on other rhamnose derivatives as potential acceptors, L-Rhap and α -L-RhapOAll, respectively. The enzymatic reaction was carried out using sucrose as donor and the rhamnose derivative as acceptor in a 1:1 ratio (146 mM for L-Rhap, 50 mM for α -L-RhapOAll).

Regarding the glucosylation of L-rhamnose, an α -(1 \rightarrow 3) linkage was achieved for all mutants, with the exception of I228Y. All mutants, except I228W, were also able to glucosylate α -L-RhapOAll. This is an additional demonstration of the potency of the strategy considering that the wild-type amylosucrase lacks this newly described specific activity (FIGURE 46)

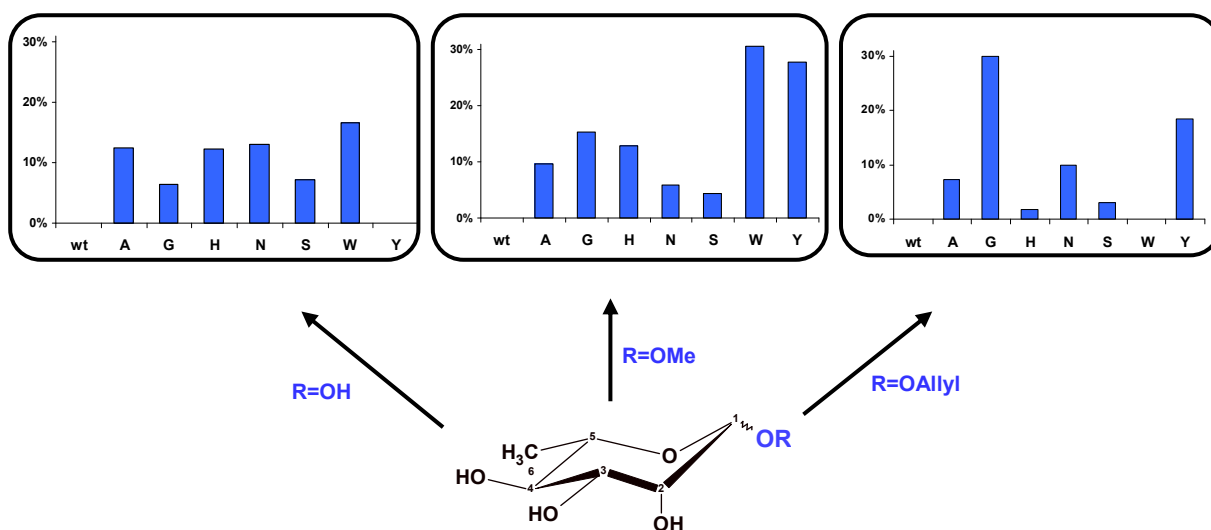


FIGURE 46 | Glucosylation Rates obtained for L-Rhap, α -L-Rhap-OMe and α -L-Rhap-Oall with formation of an α -(1-3) linkage using wild-type amylosucrase (wtAS) and its variants I228A, G, H, N, S, W, and Y.

Efficiency of the glucosylation reaction was evaluated by HPLC analysis of the acceptor reaction products synthesized using a reversed phase analytical column (Synergi C18RPFusion, 4 μ m, 30x4.6 mm) (when testing L-Rhap) and a Biorad HPLC Carbohydrate Analysis column (HPX-87K column (300x7.8 mm)) when testing α -L-RhapOAll.

F.2.b. Enzymatic glucosylation of D-GlcpNAc derivatives

The mutant F290K as well as the wtAS were tested on other D-GlcpNAc derivatives (α -D-GlcpNAcOMe, β -D-GlcpNAc-OMe, β -D-GlcpNAcOallyl and D-GlcpNHTCA) as potential acceptors. The enzymatic reaction was carried out using sucrose as donor and the D-GlcpNAc derivative as acceptor in a 1:1 ratio (146mM)

The distribution of the acceptor reaction products is shown in TABLE 13. Compared to results obtained with wild-type amylosucrase, the variant F290K appears highly specific for the

glucosylation of all tested D-GlcpNAc derivatives. (>95% for α -D-GlcpNAcOMe, β -DGlcpNAc-Oallyl and D-GlcpNHTCA and 65-70% for β -D-GlcpNAcOMe).

TABLE 13 | Glucosylation Rates of α -D-GlcpNAcOMe, β -D-GlcpNAcOMe, β -D-GlcpNAcOallyl and D-GlcNHTCA using wtAS and variant F290K.

| | Acceptor | | Acceptor Glucosylation | | | |
|--------------|---------------------------|-----------------------|---|-----------------------------------|---------------------------------|----------------------------------|
| | D-GlcpNAc derivative | Conversion Rate (DP1) | % Glc transferred onto acceptor derivatives | % Monoglucosylated acceptor (DP2) | % Diglucosylated acceptor (DP3) | % Triglucosylated acceptor (DP4) |
| wtAS | α -D-GlcpNAc-OMe | <5% | <5% | 50-55% | 45-55% | 0% |
| F290K | α -D-GlcpNAc-OMe | 50-55% | >95% | 50-55% | 45-55% | 0% |
| wtAS | β -D-GlcpNAc-OMe | <5% | <5% | 55-60% | 40-45% | 0% |
| F290K | β -D-GlcpNAc-OMe | 30-35% | 65-70% | 45-50% | 50-55% | 0% |
| wtAS | β -D-GlcpNAc-Oallyl | <5% | <5% | 55-60% | 40-45% | 0% |
| F290K | β -D-GlcpNAc-Oallyl | 50-55% | >95% | 50-55% | 45-50% | 0% |
| wtAS | D-GlcpNHTCA | 0% | 0% | 55-60% | 40-45% | 0% |
| F290K | D-GlcpNHTCA | 70-75% | >95% | 55-60% | 30-35% | 5-10% |

Efficiency of the glucosylation reaction was evaluated by HPLC analysis of the acceptor reaction product synthesized using a Biorad HPLC Carbohydrate Analysis column (HPX-87K column (300x7.8 mm)). Percentages of glucosylation rates were estimated on the basis of the peak areas as no reference was available.

G. Supplementary Methods

General methods.

TLC were performed on precoated slides of Silica Gel 60 F₂₅₄ (Merck). Detection was effected with UV light, and/or by charring in 5% sulfuric acid in ethanol. Preparative chromatography was performed by elution from columns of Silica Gel 60 (particle size 0.040-0.063 mm). Products were routinely analyzed by ¹H and ¹³C NMR spectroscopy and by mass spectrometry (MS). NMR spectra were recorded at 25 °C (400 MHz for ¹H, 100 MHz for ¹³C). Proton-signal assignments were made by first-order analysis of the spectra, as well as analysis of 2D ¹H-¹H correlation maps (COSY). The ¹³C NMR assignments were supported by 2D ¹³C-¹H correlations maps (HSQC and HMBC).

Methyl (2,3,4,6-tetra-O-acetyl- α -D-glucopyranosyl)-(1→3)-2,4-di-O-acetyl- α -L-rhamnopyranoside (3)³⁵. The crude mixture issued from the enzymatic glucosylation of **1** (9.1 g, 50 mmol) by use of non-purified I228Y extract was purified by column chromatography (CH₃CN-H₂O, 9:1) to give contaminated unreacted acceptor (7.7 g) as the first eluting product, followed by

the target disaccharide contaminated by the more polar fructose (17.5 g). Remaining sucrose was identified as the last eluting product. Acetic anhydride (100 mL) was added dropwise to a solution of the whole mixture (17.5 g) in anhydrous pyridine (100 mL) stirred at 0°C. The solution was stirred overnight at room temperature. TLC (Toluene-EtOAc, 6:4) showed the complete disappearance of the starting materials and the presence of 3 less polar products. The mixture was concentrated under reduced pressure, and volatiles were eliminated by repeated coevaporation with toluene. The residue was purified by column chromatography (Toluene-EtOAc, 7:3) to give **3** (4.9 g, 17% from **1**) as a colourless amorphous solid. $R_f = 0.45$ (6:4 Toluene-EtOAc); $^1\text{H NMR}$ (CDCl_3) δ 5.30 (dd, 1H, $J_{3,4} = 9.4$ Hz, H-3_E), 5.18 (dd, 1H, $J_{2,3} = 3.2$ Hz, H-2_A), 5.11 (d, 1H, H-1_E), 5.08 (dd, 1H, $J_{3,4} = 9.8$ Hz, H-4_A), 5.05 (dd, 1H, $J_{4,5} = 9.4$ Hz, H-4_E), 4.91 (dd, 1H, $J_{1,2} = 3.4$ Hz, $J_{2,3} = 10.3$ Hz, H-2_E), 4.61 (d, 1H, $J_{1,2} = 1.7$ Hz, H-1_A), 4.15 (dd_{overlaped}, 1H, $J_{5,6a} = 4.3$ Hz, $J_{6a,6b} = 12.4$ Hz, H-6a_E), 4.13 (dd_{overlaped}, 1H, $J_{5,6bE} = 2.5$ Hz, H-6b_E), 4.07-4.02 (m, 2H, H-5_E, H-3_A), 3.72 (dq, 1H, $J_{4,5} = 9.8$ Hz, H-5_A), 3.33 (s, 1H, OCH₃), 2.13, 2.12, 2.06, 2.00, 1.98, 1.94 (6s, 18H, 6Ac), 1.21 (d, 3H, $J_{5,6} = 6.3$ Hz, H-6_A); $^{13}\text{C NMR}$ (CDCl_3) δ 170.5, 170.0, 169.9, 169.8, 169.6, 169.3 (6C, 6C=O), 98.5 (C-1_A, $^1J_{\text{CH}} = 166.5$ Hz), 93.6 (C-1_E, $^1J_{\text{CH}} = 174.6$ Hz), 72.1 (C-3_A), 71.8 (C-4_A), 69.7 (2C, C-2_E, C-3_E), 68.5 (C-4_E), 67.9 (C-2_A), 67.7 (C-5_E), 66.6 (C-5_A), 61.6 (C-6_E), 55.0 (OCH₃), 20.9, 20.8, 20.7, 20.6, 20.5, 20.4 (6C, 6Ac), 17.5 (C-6_A). HRMS (ESI⁺) of C₂₅H₃₆O₁₆Na (M+Na, 615.1901) m/z 615.1861 ([M+Na]⁺); HRMS (ESI⁺) of C₂₅H₃₆O₁₆NH₄ (M+NH₄, 610.2347) m/z 610.2309 ([M+NH₄]⁺).

Methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside (2**)³⁵.** K₂CO₃ (21 mg, 0.15 mmol), was added to peracetylated **3** (300 mg, 0.51 mmol) in methanol (5 mL) and the suspension was stirred at room temperature. After 24 h, the reaction mixture was filtered and volatiles were evaporated. The residue was purified by reverse phase chromatography to give the target disaccharide as a lyophilized powder (160 mg, 93%). HRMS (ESI⁺) of C₁₃H₂₄O₁₀Na (M+Na, 363.1267) m/z 363.1304 ([M+Na]⁺). NMR data were identical to that of a reference compound issued from chemical synthesis.

Allyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- α -D-glucopyranoside (9**).** The crude mixture issued from the enzymatic glucosylation of **8** (3.80 g, 14.54 mmol) was purified by silica gel column chromatography (9:1 CH₃CN-H₂O) to give a mixture of **10** and fructose (4.16 g), and the pure trisaccharide **9** (2.53 g, 4.32 mmol, 30%) resulting from the enzyme-mediated transfer of glucose onto **10**. Compound **9** had $R_f = 0.3$ (9:1 CH₃CN-H₂O). $^1\text{H NMR}$ (D₂O, 400 MHz) δ (ppm): 5.91 (m, 1H, CH=), 5.35-5.32 (m, 2H, H-1_E, H-1_{E'}), 5.29 (m, 1H, =CH₂), 5.20 (m, 1H, =CH₂), 4.86 (d, 1H, $J_{1,2} = 3.5$ Hz, H-1_D), 4.15 (m, 1H,

OCH₂), 4.01-3.93 (m, 2H, H-3_D, OCH₂), 3.92-3.85 (m, 2H, H-2_D, H-3_E), 3.83-3.58 (m, 12H, H-4_D, H-5_D, H-6a_D, H-6b_D, H-4_E, H-5_E, H-6a_E, H-6b_E, H-3_F, H-5_F, H-6a_F, H-6b_F), 3.56 (dd overlapped, 1H, $J_{1,2} = 4.0$ Hz, $J_{2,3} = 10.0$ Hz, H-2_E), 3.52 (dd overlapped, 1H, $J_{1,2} = 3.8$ Hz, $J_{2,3} = 9.8$ Hz, H-2_{E'}), 3.35 (pt, 1H, $J_{3,4} = J_{4,5} = 9.3$ Hz, H-4_E), 1.97 (s, 3H, COCH₃); ¹³C NMR (D₂O, 100 MHz) δ (ppm): 174.5 (C=O), 133.7 (CH=), 118.1 (=CH₂), 99.7 (C-1_E, $J_{C,H} = 171.6$ Hz), 99.9 (C-1_{E'}, $J_{C,H} = 172.2$ Hz), 96.0 (C-1_D), 77.7 (C-4_D), 76.9 (C-4_E), 73.4 (C-3_E), 72.9 (C-3_{E'}), 72.8 (C-5_{E'}), 71.8, 71.6 (2C, C-2_E, C-2_{E'}), 71.5 (C-3_D), 71.3 (C-5_E), 70.5 (C-5_D), 69.4 (C-4_{E'}), 68.6 (OCH₂), 60.6 (3C, C-6_D, C-6_E, C-6_{E'}), 53.5 (C2_D), 21.9 (COCH₃). HRMS (ESI⁺) of C₂₃H₃₉NO₁₆Na ([M+Na]⁺, 608.2167) m/z 608.2172.

Allyl 2,3,4,6-tetra-*O*-acetyl- α -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-*O*-acetyl-2-deoxy- α -D-glucopyranoside (11). Acetic anhydride (80 mL, 725 mmol) was added dropwise to a solution of the above mentioned disaccharide **10** in mixture with fructose (4.16 g) in anhydrous pyridine (80 mL). The resulting mixture was stirred at room temperature under argon. After 2 days, TLC showed the complete disappearance of the starting materials. The mixture was concentrated under reduced pressure, and volatiles were eliminated by repeated coevaporation with toluene. The residue was purified by column chromatography (Cyclohexane-EtOAc, 6:4) to give **11** (2.96 g, 30% from **8**) as a colourless amorphous solid. Compound **11** had R_f = 0.6 (95:5 CH₂Cl₂-MeOH). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 5.90 (m, 1H, CH=), 5.71 (d, 1H, $J = 9.7$ Hz, NH), 5.45 (d, 1H, $J_{1,2} = 4.0$ Hz, H1_E), 5.35 (dd, 1H, $J_{3,4} = 9.6$ Hz, H3_E), 5.34-5.23 (m, 3H, =CH₂, H3_D), 5.04 (pt, 1H, $J_{4,5} = 9.8$ Hz, H-4_E), 4.85 (dd, 1H, $J_{2,3} = 10.5$ Hz, H-2_E), 4.79 (d, 1H, $J_{1,2} = 3.6$ Hz, H-1_D), 4.42 (dd, 1H, $J_{5,6a} = 2.0$ Hz, $J_{6a,6b} = 12.0$ Hz, H-6a_D), 4.23-4.15 (m, 4H, OCH₂, H-2_D, H-6b_D, H-6a_E), 4.04-3.91 (m, 5H, OCH₂, H-4_D, H-5_D, H-5_E, H-6b_E), 2.13, 2.08, 2.01, 2.00, 1.98, 1.91 (7s, 21H, COCH₃); ¹³C NMR (CDCl₃, 100 MHz) δ (ppm): 171.6, 170.6, 170.5, 170.2, 170.0, 169.4 (7C, C=O), 133.0 (CH=), 118.7 (=CH₂), 96.1 (C-1_D), 95.5 (C-1_E), 74.1 (C-3_D), 72.4 (C-4_D), 70.1 (C-2_E), 69.4 (C-3_E), 68.8 (OCH₂), 68.4 (C-5_E), 68.0 (C-5_D), 67.9 (C-4_E), 62.8 (C-6_D), 61.4 (C-6_E), 52.2 (C-2_D), 23.1, 21.0, 20.8, 20.7, 20.6, 20.5 (7C, COCH₃). HRMS (ESI⁺) of C₂₉H₄₁NO₁₇Na ([M+Na]⁺, 698.2272) m/z 698.2277.

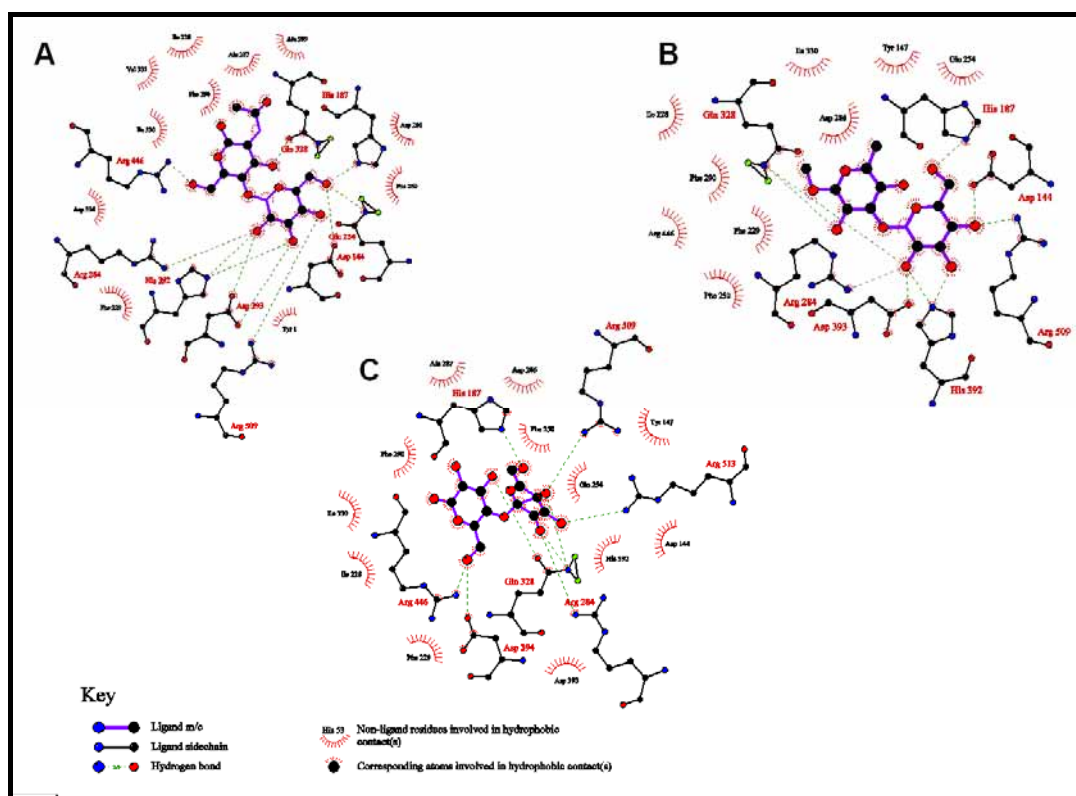
Allyl α -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- α -D-glucopyranoside (10).

Method a: Methanolic sodium methoxide (0.5 M solution, 8 mL, 4 mmol) was added to a stirred solution of **11** (2.68 g, 3.97 mmol) in anhydrous methanol (100 mL). The reaction mixture was stirred for 6.5 h at room temperature by which time all the starting material had been consumed (R_f = 0.3 in 9:1 CH₃CN-H₂O). Excess base was neutralized with Dowex-H⁺ resin, filtered, concentrated, and purified by silica gel chromatography, eluting with 10% H₂O in CH₃CN to

afford compound **10** as a brownish foamy solid (1.65 g, 98%); ^1H NMR (D_2O , 400 MHz) δ (ppm): 5.95 (m, 1H, CH=), 5.38 (d, 1H, $J_{1,2} = 3.9$ Hz, H-1_E), 5.34 (m, 1H, =CH₂), 5.25 (m, 1H, =CH₂), 4.91 (d, 1H, $J_{1,2} = 3.5$ Hz, H-1_D), 4.20 (m, 1H, OCH₂), 4.02 (m overlapped, 1H, OCH₂), 4.01 (dd overlapped, 1H, $J_{3,4} = 8.4$ Hz, H-3_D), 3.94 (dd, 1H, $J_{2,3} = 10.8$ Hz, H-2_D), 3.89-3.67 (m, 7H, H-4_D, H-5_D, H-6_{aD}, H-6_{bD}, H-5_E, H-6_{aE}, H-6_{bE}), 3.66 (pt, 1H, H-3_E), 3.56 (dd, 1H, $J_{2,3} = 9.9$ Hz, H-2_E), 3.40 (pt, 1H, $J_{3,4} = J_{4,5} = 9.5$ Hz, H-4_E), 2.02 (s, 3H, COCH₃); ^{13}C NMR (D_2O , 100 MHz) δ (ppm): 174.5 (C=O), 133.7 (CH=), 118.1 (=CH₂), 99.9 (C-1_E), 96.0 (C-1_D), 77.6 (C4_D), 72.9 (C-3_E), 72.8 (C-5_E), 71.8 (C-2_E), 71.5 (C-3_D), 70.6 (C-5_D), 69.4 (C-4_E), 68.6 (OCH₂), 60.6 (2-C, C-6_D, C-6_E), 53.5 (C-2_D), 21.9 (COCH₃). HRMS (ESI⁺) of C₁₇H₃₀NO₁₁ ([M+H]⁺, 424.1819) m/z 424.1821.

Method b: Trisaccharide **9** (2.19 g, 3.74 mmol) in solution in 110 mL of acetate buffer (pH 4.8) was incubated with 1120 IU of amyloglucosidase from *Aspergillus niger* under magnetic stirring for 90 min at 50 °C. After freeze-drying, the products were purified by silica gel chromatography eluting with 9:1 CH₃CN–H₂O. Concentration and freeze-drying gave a 1:1.1 mixture of disaccharide **10** and D-glucose (1.43 g, 62%) as seen by NMR.

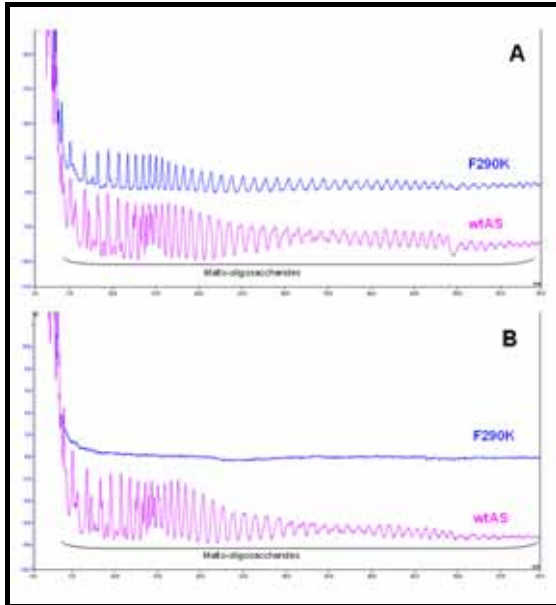
H. Supplementary Figures



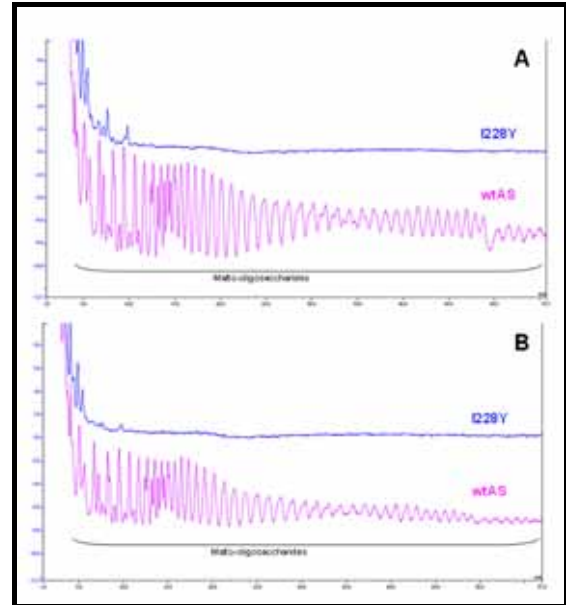
SUPPLEMENTARY FIGURE 1| Schematic diagramm of protein-ligand interactions (LigPlot)

(A) AS- α -D-Glcp-(1→4)-D-GlcpNAc (ED) interactions
 (B) AS- α -D-Glcp-(1→3)- α -L-Rhap (EA') interactions
 (C) AS-maltose interactions

SUPPLEMENTARY FIGURE 2| Comparison of Dionex HPAEC product profiles obtained at the end of the reaction ($t = 24$ h) with wild-type AS (magenta) and the variant F290K (blue) using (A) 146 mM sucrose and (B) using 146 mM sucrose supplemented with 146 mM acceptor (α -D-GlcpNAc-Oallyl).

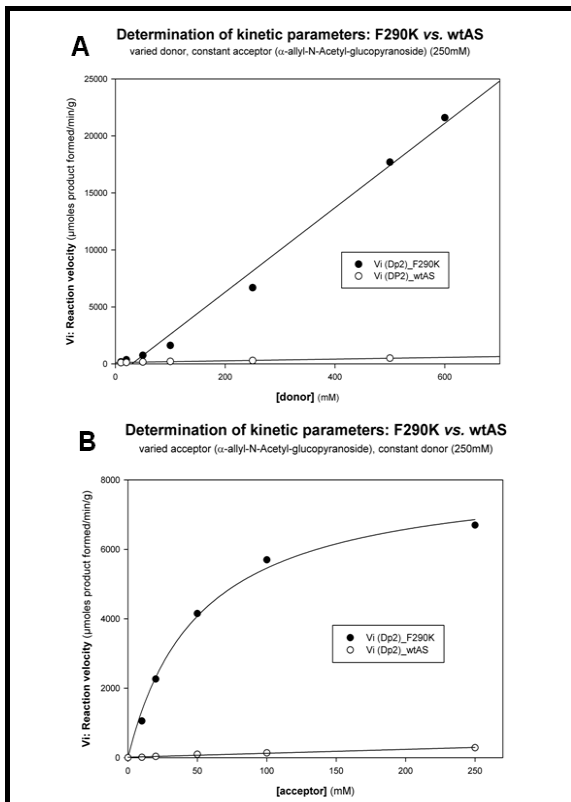


SUPPLEMENTARY FIGURE 3| Comparison of Dionex HPAEC product profiles obtained at the end of the reaction ($t = 24$ h) with wild-type AS (magenta) and the variant I228Y (blue) using (A) 146 mM sucrose and (B) using 146 mM sucrose supplemented with 146 mM acceptor (α -L-RhapOMe).



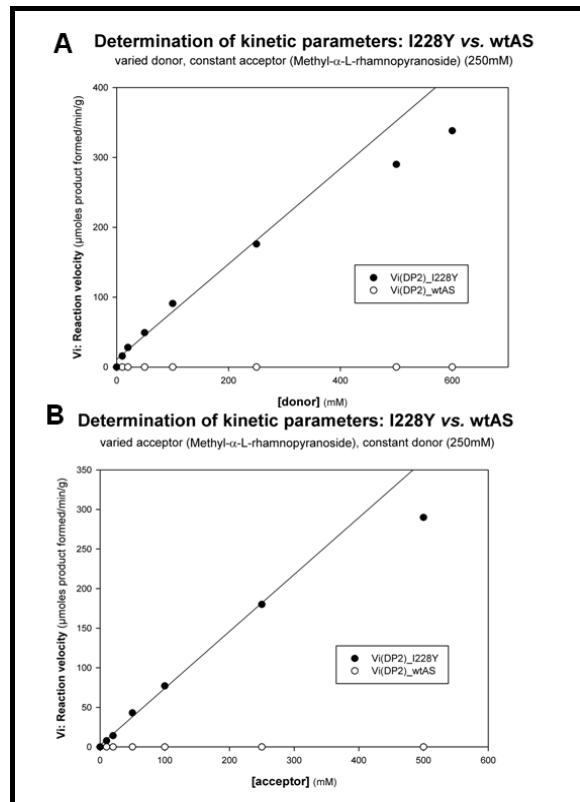
SUPPLEMENTARY FIGURE 4| Comparison of kinetic parameters for the acceptor glucosylation catalysed by the variant F290K and wtAS.

Rate of acceptor glucosylation is calculated as the desired disaccharide formed per min per gramm of enzyme. (a) varied donor, constant acceptor (250 mM); (b) varied donor, constant acceptor (250 mM)



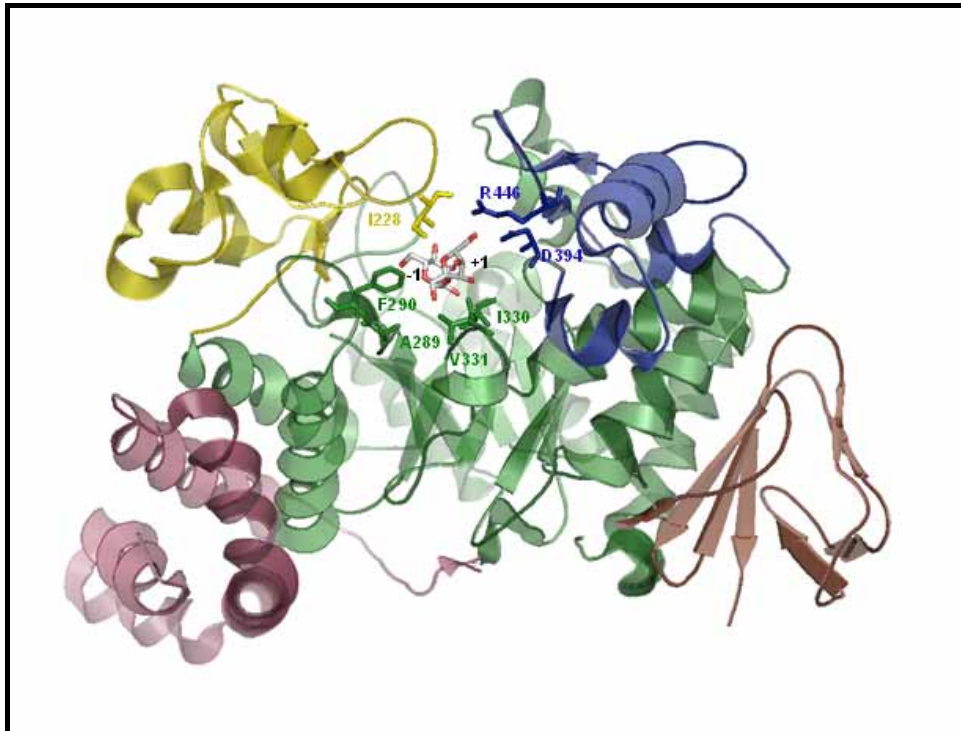
SUPPLEMENTARY FIGURE 5| Comparison of kinetic parameters for the acceptor glucosylation catalysed by the variant I228Y and wtAS.

Rate of acceptor glucosylation is calculated as the desired disaccharide formed per min per gramm of enzyme. (a) varied donor, constant acceptor (250 mM); (b) varied donor, constant acceptor (250 mM)



SUPPLEMENTARY FIGURE 6| Overall structure of *N. polysaccharea* amylosucrase.

The structure consists of 5 domains, three of which are characteristic of GH family 13 enzymes: the characteristic non-permuted (β/α)8-barrel folded catalytic domain A (coloured in green); domain B (coloured in yellow), an extension of domain A between β -strand 3 and α -helix 3; the C-terminal Greek key domain (coloured in brown). Beside these common features, AS structure also contains two unique additional domains, namely the N-terminal α -helical domain (coloured in red), and a B'-domain between β -strand 7 and α -strand 7 of the catalytic barrel (coloured in blue)⁴². Sucrose is shown in the active site for reference. The 7 positions (228, 289, 290, 330, 331, 394 and 446) are coloured according to the domain they belong to.



I. References

1. Persidis, A., The carbohydrate-based drug industry. *Nat Biotechnol*, **1997**, 15, (5), 479-80.
2. Seeberger, P. H.; Werz, D. B., Synthesis and medical applications of oligosaccharides. *Nature*, **2007**, 446, (7139), 1046-51.
3. Kilcoyne, M.; Joshi, L., Carbohydrates in therapeutics. *Cardiovasc Hematol Agents Med Chem*, **2007**, 5, (3), 186-97.
4. Lee, J. C.; Greenberg, W. A.; Wong, C. H., Programmable reactivity-based one-pot oligosaccharide synthesis. *Nat Protoc*, **2006**, 1, (6), 3143-52.
5. Plante, O. J.; Palmacci, E. R.; Seeberger, P. H., Automated solid-phase synthesis of oligosaccharides. *Science*, **2001**, 291, (5508), 1523-7.
6. Wang, C. C.; Lee, J. C.; Luo, S. Y.; Kulkarni, S. S.; Huang, Y. W.; Lee, C. C.; Chang, K. L.; Hung, S. C., Regioselective one-pot protection of carbohydrates. *Nature*, **2007**, 446, (7138), 896-9.
7. Joe, M.; Bai, Y.; Nacario, R. C.; Lowary, T. L., Synthesis of the docosanasaccharide arabinan domain of mycobacterial arabinogalactan and a proposed octadecasaccharide biosynthetic precursor. *J Am Chem Soc*, **2007**, 129, (32), 9885-901.
8. Ruffing, A.; Chen, R. R., Metabolic engineering of microbes for oligosaccharide and polysaccharide synthesis. *Microb Cell Fact*, **2006**, 5, 25.
9. Drouillard, S.; Driguez, H.; Samain, E., Large-scale synthesis of H-antigen oligosaccharides by expressing *Helicobacter pylori* α -1,2-fucosyltransferase in metabolically engineered *Escherichia coli* cells. *Angew Chem Int Ed Engl*, **2006**, 45, (11), 1778-80.
10. Bettler, E.; Samain, E.; Chazalet, V.; Bosso, C.; Heyraud, A.; Joziassé, D. H.; Wakarchuk, W. W.; Imberty, A.; Geremia, A. R., The living factory: in vivo production of N-acetylglucosamine containing carbohydrates in *E. coli*. *Glycoconj J*, **1999**, 16, (3), 205-12.
11. Thiem, J., Applications of enzymes in synthetic carbohydrate chemistry. *FEMS Microbiol Rev*, **1995**, 16, (2-3), 193-211.
12. Koeller, K. M.; Wong, C. H., Enzymes for chemical synthesis. *Nature*, **2001**, 409, (6817), 232-40.
13. Yan, F.; Wakarchuk, W. W.; Gilbert, M.; Richards, J. C.; Whitfield, D. M., Polymer-supported and chemoenzymatic synthesis of the *Neisseria meningitidis* pentasaccharide: a methodological comparison. *Carbohydr Res*, **2000**, 328, (1), 3-16.
14. Hanson, S.; Best, M.; Bryan, M. C.; Wong, C. H., Chemoenzymatic synthesis of oligosaccharides and glycoproteins. *Trends Biochem Sci*, **2004**, 29, (12), 656-63.
15. Crout, D. H.; Vic, G., Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. *Curr Opin Chem Biol*, **1998**, 2, (1), 98-111.
16. Schroven, A.; Meinke, S.; Ziegelmüller, P.; Thiem, J., Transsialidase from *Trypanosoma cruzi* for Regio- and Stereoselective Synthesis of N-Acyl-Modified Sialylated Oligosaccharides and Measurement of Transfer Rates. *Chemistry*, **2007**.
17. Schmidt, D.; Sauerbrei, B.; Thiem, J., Chemoenzymatic synthesis of sialyl oligosaccharides with sialidases employing transglycosylation methodology. *J Org Chem*, **2000**, 65, (25), 8518-26.
18. Yoon, J. H.; Ajisaka, K., The synthesis of galactopyranosyl derivatives with beta-galactosidases of different origins. *Carbohydr Res*, **1996**, 292, 153-63.
19. Niggemann, J.; Kamerling, J. P.; Vliegthart, J. F., β -1,4-Galactosyltransferase-catalyzed synthesis of the branched tetrasaccharide repeating unit of *Streptococcus pneumoniae* type 14. *Bioorg Med Chem*, **1998**, 6, (9), 1605-12.

20. Eichler, E.; Jennings, H. J.; Gilbert, M.; Whitfield, D. M., Synthesis of a disialylated hexasaccharide of type VIII group B *Streptococcus* capsular polysaccharide. *Carbohydr Res*, **1999**, 319, (1-4), 1-16.
21. Hancock, S. M.; Vaughan, M. D.; Withers, S. G., Engineering of glycosidases and glycosyltransferases. *Curr Opin Chem Biol*, **2006**, 10, (5), 509-19.
22. Jank, T.; Reinert, D. J.; Giesemann, T.; Schulz, G. E.; Aktories, K., Change of the donor substrate specificity of *Clostridium difficile* toxin B by site-directed mutagenesis. *J Biol Chem*, **2005**, 280, (45), 37833-8.
23. Kotloff, K. L.; Winickoff, J. P.; Ivanoff, B.; Clemens, J. D.; Swerdlow, D. L.; Sansonetti, P. J.; Adak, G. K.; Levine, M. M., Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ*, **1999**, 77, (8), 651-66.
24. Levine, M. M.; Kotloff, K. L.; Barry, E. M.; Pasetti, M. F.; Sztein, M. B., Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol*, **2007**, 5, (7), 540-53.
25. Phalipon, A.; Costachel, C.; Grandjean, C.; Thuizat, A.; Guerreiro, C.; Tanguy, M.; Nato, F.; Vulliez-Le Normand, B.; Belot, F.; Wright, K.; Marcel-Peyre, V.; Sansonetti, P. J.; Mulard, L. A., Characterization of functional oligosaccharide mimics of the *Shigella flexneri* serotype 2a O-antigen: implications for the development of a chemically defined glycoconjugate vaccine. *J Immunol*, **2006**, 176, (3), 1686-94.
26. Mulard, L. A.; Phalipon, A., From epitope characterization to the design of semi-synthetic glycoconjugate vaccines against *Shigella flexneri* 2a infection, in "Carbohydrate-based vaccines". *ACS symp. ser. 899, in press*, **2008**.
27. Lindberg, A. A.; Karnell, A.; Weintraub, A., The lipopolysaccharide of *Shigella* bacteria as a virulence factor. *Rev Infect Dis*, **1991**, 13 Suppl 4, S279-84.
28. Simmons, D. A., The structural basis of serological specificity in *Shigella flexneri* O-antigens. *Biochem Soc Trans*, **1990**, 18, (6), 1271-2.
29. Yoon, S. H.; Robyt, J. F., Synthesis of acarbose analogues by transglycosylation reactions of *Leuconostoc mesenteroides* B-512FMC and B-742CB dextranases. *Carbohydr Res*, **2002**, 337, (24), 2427-35.
30. Cote, G. L.; Dunlap, C. A.; Appell, M.; Momany, F. A., Alternansucrase acceptor reactions with D-tagatose and L-glucose. *Carbohydr Res*, **2005**, 340, (2), 257-62.
31. Richard, G.; Yu, S.; Monsan, P.; Remaud-Simeon, M.; Morel, S., A novel family of glucosyl 1,5-anhydro-D-fructose derivatives synthesised by transglucosylation with dextranase from *Leuconostoc mesenteroides* NRRL B-512F. *Carbohydr Res*, **2005**, 340, (3), 395-401.
32. Demuth, K.; Jordening, H. J.; Buchholz, K., Oligosaccharide synthesis by dextranase: new unconventional acceptors. *Carbohydr Res*, **2002**, 337, (20), 1811-20.
33. Mulard, L. A.; Ughetto-Monfrin, J., Synthesis of the methyl glycosides of a tri- and tetrasaccharide fragment specific for the *Shigella flexneri* serotype 5a O-antigen. A reinvestigation. *J Carbohydr Chem*, **1999**, 18, 721-53.
34. Binkley, R. W.; Goewey, G. S.; Johnston, J. C., Regioselective ring opening of selected benzylidene acetals: a photochemically initiated reaction for partial deprotection of carbohydrates. *J. Org. Chem.*, **1984**, 49, 992-996.
35. Bakinovskii, L. V.; Gomtsian, A. R.; Bairamova, N. E.; Kochetkov, N. K., [Synthesis of oligosaccharide fragments of O-specific *Shigella flexneri* polysaccharides. II. Synthesis of trisaccharide $\alpha\text{Glc}(1-3)\alpha\text{Rha}(1-2)\alpha\text{RhaOMe}$ and tetrasaccharide $\beta\text{GlcNAc}(1-2)\alpha\text{Glc}(1-3)\alpha\text{Rha}(1-2)\alpha\text{RhaOMe}$]. *Bioorg Khim*, **1985**, 11, (2), 254-63.

36. Mulard, L. A.; Clément, M. J.; Segat-Dioury, F.; Delepierre, M., Synthesis and NMR study of a linear pentasaccharide fragment of the *Shigella flexneri* 5a O-specific polysaccharide. *Tetrahedron* **2002**, 58, 2593-604.
37. Mulard, L. A.; Clément, M. J.; Imberty, A.; Delepierre, M., Convergent synthesis, NMR and conformational analysis of tetra- and pentasaccharide haptens of the *Shigella flexneri* 5a O-specific polysaccharide. *Eur J Org Chem*, **2002**, 2486-98.
38. unpublished results. **2007**.
39. Champion, E.; Morel, S.; Mulard, L.; Monsan, P.; Remaud-Siméon, M.; André, I., Enzymatic synthesis of disaccharide motives entering in the composition of bacterial polysaccharides by use of glucansucrases. *unpublished results*.
40. Skov, L. K.; Mirza, O.; Henriksen, A.; Potocki de Montalk, G.; Remaud-Simeon, M.; Sarcabal, P.; Willemot, R. M.; Monsan, P.; Gajhede, M., Crystallization and preliminary X-ray studies of recombinant amylosucrase from *Neisseria polysaccharea*. *Acta Crystallogr D Biol Crystallogr*, **2000**, 56, (Pt 2), 203-5.
41. Mirza, O.; Skov, L. K.; Remaud-Simeon, M.; Potocki de Montalk, G.; Albenne, C.; Monsan, P.; Gajhede, M., 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*, **2001**, 40, (30), 9032-9.
42. Skov, L. K.; Mirza, O.; Sprogoe, D.; Dar, I.; Remaud-Simeon, M.; Albenne, C.; Monsan, P.; Gajhede, M., Oligosaccharide and sucrose complexes of amylosucrase. Structural implications for the polymerase activity. *J Biol Chem*, **2002**, 277, (49), 47741-7.
43. Davies, G. J.; Wilson, K. S.; Henrissat, B., Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*, **1997**, 321 (Pt 2), 557-9.
44. Lipkind, G. M.; Shashkov, A. S.; Mamyán, S. S.; Kochetkov, N. K., The nuclear overhauser effect and structural factors determining the conformations of disaccharide glycosides. *Carbohydr Res*, **1988**, 181, 1-12.
45. André-Leroux, G. In *A.C.S. symposium series*, 2006; 2006; pp 170-185.
46. Allison, G. E.; Verma, N. K., Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol*, **2000**, 8, (1), 17-23.
47. Zhang, J. H.; Dawes, G.; Stemmer, W. P., Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci U S A*, **1997**, 94, (9), 4504-9.
48. Parikh, M. R.; Matsumura, I., Site-saturation mutagenesis is more efficient than DNA shuffling for the directed evolution of β -fucosidase from β -galactosidase. *J Mol Biol*, **2005**, 352, (3), 621-8.
49. Osanjo, G.; Dion, M.; Drone, J.; Solleux, C.; Tran, V.; Rabiller, C.; Tellier, C., Directed evolution of the α -L-fucosidase from *Thermotoga maritima* into an α -L-transfucosidase. *Biochemistry*, **2007**, 46, (4), 1022-33.
50. Hancock, S. M.; Corbett, K.; Fordham-Skelton, A. P.; Gatehouse, J. A.; Davis, B. G., Developing promiscuous glycosidases for glycoside synthesis: residues W433 and E432 in *Sulfolobus solfataricus* β -glycosidase are important glucoside- and galactoside-specificity determinants. *Chembiochem*, **2005**, 6, (5), 866-75.
51. Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G., Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J. Am. Chem. Soc.*, **1998**, 120, 5583-5584.
52. Kim, Y. W.; Lee, S. S.; Warren, R. A.; Withers, S. G., Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire. *J Biol Chem*, **2004**, 279, (41), 42787-93.
53. Honda, Y.; Kitaoka, M., The first glycosynthase derived from an inverting glycoside hydrolase. *J Biol Chem*, **2006**, 281, (3), 1426-31.

54. Okuyama, M.; Mori, H.; Watanabe, K.; Kimura, A.; Chiba, S., α -glucosidase mutant catalyzes " α -glycosynthase"-type reaction. *Biosci Biotechnol Biochem*, **2002**, 66, (4), 928-33.
55. Aharoni, A.; Thieme, K.; Chiu, C. P.; Buchini, S.; Lairson, L. L.; Chen, H.; Strynadka, N. C.; Wakarchuk, W. W.; Withers, S. G., High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat Methods*, **2006**, 3, (8), 609-14.
56. Griffiths, A. D.; Tawfik, D. S., Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *Embo J*, **2003**, 22, (1), 24-35.
57. Lemanski, G.; Ziegler, T., Synthesis of pentasaccharide fragments related to the O-specific polysaccharide of *Shigella flexneri* serotype 1a. *Eur J Org Chem*, **2006**, (11), 2618-2630.
58. Ovodov, Y. S., Bacterial capsular antigens. Structural patterns of capsular antigens. *Biochemistry (Mosc)*, **2006**, 71, (9), 937-54.
59. De Montalk, G. P.; Remaud-Simeon, M.; Willemot, R. M.; Planchot, V.; Monsan, P., Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J Bacteriol*, **1999**, 181, (2), 375-81.
60. Ghosh, M.; Dulina, R. G.; Kakarla, R.; Sofia, M. J., Efficient Synthesis of a Stereochemically Defined Carbohydrate Scaffold: Carboxymethyl 2-Acetamido-6-Azido- 4-O-Benzyl-2-deoxy- α -D-Glucopyranoside. *J Org Chem*, **2000**, 65, (24), 8387 - 8390.
61. Ray, A. K.; Roy, N., Synthesis of the tetrasaccharide repeating-unit of the polysaccharide from *Klebsiella* type 23. *Carbohydr Res*, **1990**, 196, 95-100.
62. Mulard, L. A.; Costachel, C.; Sansonetti, P. J., Synthesis of the methyl glycosides of a di- and two trisaccharide fragments specific for the *Shigella flexneri* serotype 2a O-antigen. *J Carbohydr Chem* **2000**, 19, 849-77.
63. Allinger, N. L.; Zhou, X.; Bergsma, J., Molecular mechanics parameters. *Theochem*, **1994**, 118, 69-83.
64. Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, **1976**, 72, 248-54.
65. Sumner, J.; Howell, S. A., A method for the determination of invertase activity. *J. Biol. Chem*, **1935**, 108, 51-54.

**Chapitre IV : Exploration du sous-site
(+1) de l'amylosaccharase pour la
glucosylation d'accepteurs non-naturels
à l'aide d'une nouvelle méthode de
criblage**

La stratégie semi-rationnelle adoptée pour le remodelage du sous-site (+1) de l'amylosaccharase a permis d'obtenir avec succès, plusieurs enzymes améliorées et de spécificité nouvelle pour la synthèse spécifique des motifs α -D-Glcp-(1→3)- α -L-Rhap et α -D-Glcp-(1→4)-D-GlcNAcp. Cette approche a démontré que l'introduction d'une seule mutation, sur une des 7 positions précédemment sélectionnées, pouvait modifier et améliorer la spécificité de l'amylosaccharase vis-à-vis d'un substrat accepteur mais aussi altérer fortement l'affinité de cette enzyme vis-à-vis de son substrat donneur.

Nous avons décidé de poursuivre l'exploration du sous-site +1 en générant davantage de diversité sur la base de ces mêmes positions. Cette exploration pourrait permettre d'établir un nouveau jeu d'interactions bénéfiques entre les substrats (donneur et accepteur) et les positions mutées et créer ainsi de nouveaux mutants plus performants.

Cependant, la recombinaison des 20 acides aminés possibles sur les 7 positions cibles conduirait à une banque de taille importante (20^7 variants), qui serait difficile d'explorer dans son intégralité. Par modélisation moléculaire, nous avons cherché, à rationaliser l'approche combinatoire de façon à réduire le nombre de combinaisons générées et à créer des effets synergiques entre les différentes positions considérées. Or, nos précédents travaux ont démontré que 228 et 290 étaient des positions stratégiques pour l'accommodation de nos accepteurs cibles. Ces 2 positions, ainsi que les positions 229, 289 et 330 se trouvent dans un environnement spatial très proche, qui nous est apparu, dans un premier temps, plus judicieux d'explorer.

Dans cette troisième publication, nous mettons à profit cette ingénierie raisonnée en construisant et criblant 4 banques de double-mutants. Nous décrivons également la mise au point d'une nouvelle méthode de criblage, adaptée à la taille plus large de ces banques.

EXPLORATION OF AMYLOSUCRASE SEQUENCE-FUNCTION SPACE FOR NON-NATURAL ACCEPTOR GLUCOSYLATION USING A NOVEL SCREENING METHOD

Claire Moulis^{1-a,b,c}, Elise Champion^{1-a,b,c}, Isabelle André^{a,b,c}, Pierre Monsan^{a,b,c,d} and Magali Remaud-Siméon^{a,b,c}.

^a Université de Toulouse; INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

^b CNRS, UMR5504, F-31400 Toulouse, France

^c INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France

^d Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris

^e Unité de Chimie des Biomolécules, URA CNRS 2128, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris cedex 15, France

¹ These authors contributed equally to this work

A. Summary

Amylosucrase from *Neisseria polysaccharea* is a transglucosidase belonging to glycoside-hydrolase family 13 that catalyses the synthesis of an amylose-like polymer from an inexpensive and abundant substrate, sucrose, as well as the transglucosylation of exogenous hydroxylated acceptors, turning this enzyme into an appealing enzymatic glucosylation tool. A semi-rational engineering approach was undertaken to re-design the enzyme active site (acceptor binding site) and adapt it to non-natural acceptor glucosylation, either methyl α -L-rhamnopyranoside or *N*-acetyl-D-glucosamine, to produce two disaccharides, the methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside and the α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acetyl-D-glucosamine, which enter into the chemo-enzymatic synthesis of *Shigella flexneri* serotypes 1a, 1b, 2b, 3a, 5a, 5b and X lipopolysaccharides.

About 20 000 amylosucrase variants containing either two vicinal or distant mutations at key positions were generated by saturation mutagenesis. These libraries were pre-screened using a novel colorimetric screening method based on the use of a pH indicator for the detection of sucrose-utilizing variants, thereby reducing by 92% the size of the libraries to be subsequently screened for acceptor glucosylation. Active clones (8% of the initial libraries) were screened for acceptor recognition, leading to the isolation of 54 variants of potential interest for the production of the target disaccharides. Among them, 5 double-mutants were identified to efficiently synthesize the methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside and 15 to produce the α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acetyl-D-glucosamine.

B. Introduction

Amylosucrases (EC 2.4.1.4) are glucansucrases from glycoside-hydrolase (GH) family 13, which synthesize an insoluble α -(1,4) glucan from sucrose¹⁻³. In addition, amylosucrases also catalyze a transglucosylation reaction onto exogenous hydroxylated acceptors, mostly oligoaccharides and glucoconjugates⁴⁻⁶. This so-called “acceptor” reaction occurs at the cost of native polymer formation and sucrose hydrolysis; the acceptor efficiency is dependent on its recognition and ability to compete with natural substrates. As amylosucrases use a low cost and renewable glucosyl donor (sucrose), these enzymes are attractive glucosylation tools for the synthesis of structurally-defined glucosylated molecules under mild conditions.

Among the known amylosucrases^{2, 7, 8}, recombinant amylosucrase from *Neisseria polysaccharea* (AS) has been the most studied. Structural analysis of AS oligosaccharide and sucrose complexes coupled to directed mutagenesis and kinetics data allowed to elucidate the α -retaining mechanism of this enzyme and identify the specific structural features involved in sucrose recognition and polymerase activity⁹⁻¹¹. In addition, AS catalytic efficiency and thermal stability were improved by directed molecular evolution^{12, 13, 14}.

Recently, we have undertaken a semi-rational engineering approach to re-design the AS active site and adapt it to non-natural acceptor glucosylation¹⁵. The tailoring of this enzyme required a thorough analysis of the substrate binding site and, more particularly, of the acceptor binding site referred as +1 subsite¹⁶. A molecular modeling study based on molecular docking experiments was carried out to identify residues that could be modified to potentially alter the recognition of non-natural acceptor. Two non-natural acceptors were targeted, namely the methyl α -L-rhamnopyranoside (**A'**, α -L-RhapOMe) and the *N*-acetyl-D-glucosamine (**D**, D-GlcpNAc). The objective was the synthesis of two disaccharides which are not found in nature: the α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside (**EA'**, α -D-Glcp-(1 \rightarrow 3)- α -L-RhapOMe) and the α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acetyl-D-glucosamine (**ED**, α -D-Glcp-(1 \rightarrow 4)- α -D-GlcpNAc). These disaccharides were foreseen as useful starting building blocks that could be used in a chemo-enzymatic pathway¹⁵, designed on purpose, to produce complex oligosaccharide haptens mimicking lipopolysaccharides of *Shigella flexneri* serotypes 1a, 1b, 2b, 3a, 5a, 5b and X^{17, 18}. These haptens are indeed potential candidates for the production of a multivalent conjugate vaccine against the most prevalent serotypes of *Shigella flexneri*^{19, 20}, the human pathogen agent of bacillary dysentery, causing the leading enteric diseases.

From the initial molecular modelling study, 7 residues of AS acceptor subsite +1 were selected¹⁵. A mono-mutant library (133 mutants) focused on these residues was then generated and screened. Notably, mutations on 5 out of the 7 targeted positions improved the acceptor recognition and

their glucosylation, showing that AS is tolerant to those mutations and that a recombination of the mutations could be proposed to enlarge the diversity at the acceptor site and further improve the glucosylation of the target acceptors.

In this paper, our aim was to explore a larger sequence-function space of +1 subsite in order to isolate mutants better adapted to **A'** and **D** binding and glucosylation. First of all, a colorimetric screening test based on the use of a pH indicator was developed and optimized to fasten the isolation from large libraries of active mutants that have kept the ability to use sucrose as substrate. A data-driven exploration of the subsite +1 sequence space was then conducted by generating double-mutant libraries based on vicinal or distant key positions, identified to be beneficial for α -L-RhapOMe (**A'**) and D-GlcpNAc (**D**) glucosylation in a previous study¹⁵.

C. Results

Engineering amylosucrase specificity to enhance the glucosylation of a poor acceptor such as the α -D-GlcpNAc or a naturally not-recognized molecule like the α -L-RhapOMe is a difficult task. Indeed, engineering of the enzyme, while affecting specificity for the acceptor, must not perturb its specificity towards the donor substrate (in our case sucrose). Consequently, sucrose usage is a pre-requisite for any variant and mutant libraries, and all enzymes will be pre-screened for their capability to cleave sucrose in our screening strategy, before being screened for their ability to glucosylate the target acceptors.

C.1 Development of a colorimetric assay for detecting sucrose-utilizing variants

E. coli strains derived from *E. coli* K12 are unable to use sucrose as substrate. This property was used to develop a colorimetric screening test that allows isolation of recombinant *E. coli* clones on solid medium and the determination of the ratio of active clones present in a library of variants. The principle is based on the fact that active amylosucrase produced by recombinant *E. coli* will cleave sucrose and release fructose. This latter can enter the glycolytic pathway to produce acids and induce pH changes that should be easily detected by using an appropriate pH colorimetric indicator. In the absence of an active glucansucrase, no acid production occurs and, thus, no change in the pH indicator color is observed.

To set up this assay and validate the new pH-based screen of active amylosucrase-producing clones, we applied the method to a library of 133 mono mutants previously constructed¹⁵. After growth in microtiter plates, the cells were broken and incubated with sucrose for 24 hours, before adding the BromoThymol Blue (BBT). In parallel, a duplicate plate was analyzed using the dinitrosalicylic assay (DNS)²¹ which enables the detection of the reducing power released from sucrose cleavage. As shown on **FIGURE 47A**, the color of some wells changed from blue (pH=7.5)

to green (pH ~6-7) or yellow (pH <6). This indicates that for some clones, an acidification of the medium occurred, thus revealing their ability to cleave sucrose. The slight change in color observed further shows that some clones produced either low levels of active AS or AS with a lower activity (**FIGURE 47A**). 50% of active clones were detected in this library. By comparison, only 33% of the recombinant clones were detected as active with the DNS assay (**FIGURE 47B**). Notably, all the clones detected with the DNS assay were also detected using the new colorimetric assay.

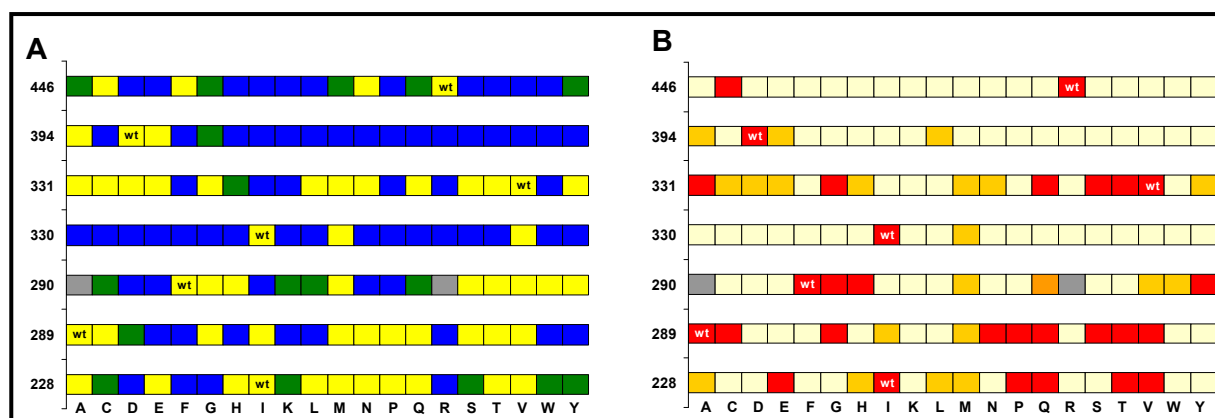


FIGURE 47 | Colorimetric tests for detection of amylosucrase activity

Rows indicate the 7 mutated positions and columns represent the 20 possible amino acid mutations (including the wt)
(A): Mutants activity detected by BBT test (yellow, strong activity, pH < 6; green, significant activity, pH ~ 6-7; blue, no activity, pH ≥ 7.5; grey, not determined)

(B): Mutants activity detected by DNS test (red, strong activity, >2g/l released reduced sugar; orange, significant activity, 0-2g/l released reduced sugar; yellow, no activity, 0 g/l released reduced sugar; grey, not determined)

Screening using BBT colorimetric test at colony level

BBT test was then used to develop a simple and highly sensitive staining method to detect clones producing active amylosucrase on solid medium. BBT pH indicator was directly introduced into solid LB_{amp} medium supplemented with sucrose, IPTG and Tris-Buffer at pH = 7.5 to maintain medium staining (blue color). We first optimized the BBT concentration (from 0.05 g/L to 0.5 g/L) and found that 0.1 g/L of BBT offered the best contrast between blue and yellow colonies after incubation at 30°C during 48h (**FIGURE 48**). We first applied this protocol to freshly transformed cells and observed that transformation yield was much lower than that observed in usual conditions. In addition, colony development was also affected. Therefore, in order to ensure a good revival of the colonies, bacteria were first plated on a square plate containing LB_{amp} agar medium supplemented with 1% glycerol on which a hydrophilic membrane (polyvinylidene fluoride) had been overlaid. The plates were then incubated at 37°C overnight to ensure colony development. The membrane was then transferred onto a second square plate containing LB_{amp} agar medium supplemented with sucrose, IPTG (inductor), Tris-Buffer at pH = 7.5 and BBT. After 24h incubation at 30°C, the colonies were easily differentiated, picked and

cultured without any loss of viability. This procedure was subsequently applied to the library screening.

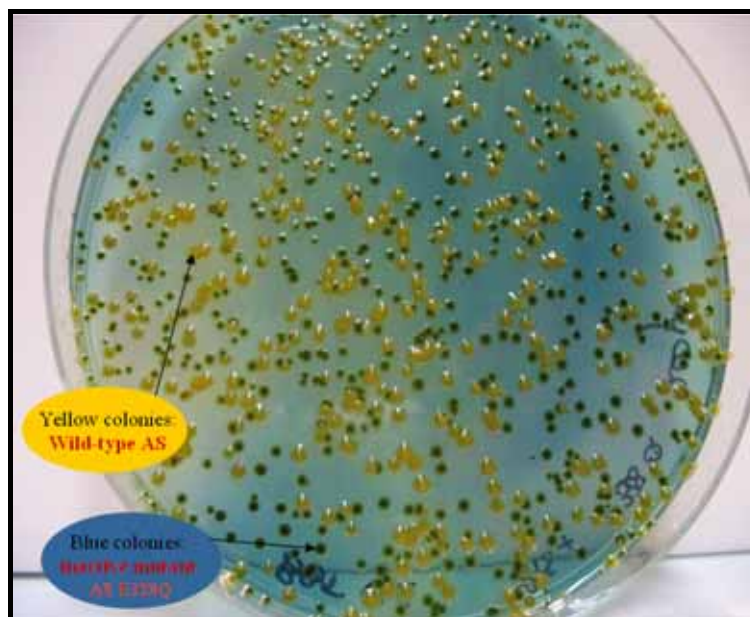


FIGURE 48 | Screening of amylosucrase active clones

on solid LB medium containing Tris-HCl 50 mM pH 7.5 + Amp 100 µg/mL + IPTG 1 mM + Sucrose 50 g/l + BBT 0.1 g/l.

C.2 Library construction and pre-screening of sucrose-utilizing variants

As shown on

FIGURE 49, we previously showed that positions 228 and 290 are key positions for altering AS selectivity towards either α -L-RhapOMe (**A'**) and/or D-GlcpNAc (**D**). However, given the spatial vicinity of these two sites with positions 229, 289 and 330, and the influence they might have on each other, we chose to construct libraries focused on all 5 positions 228, 289, 290, 330, and 331.

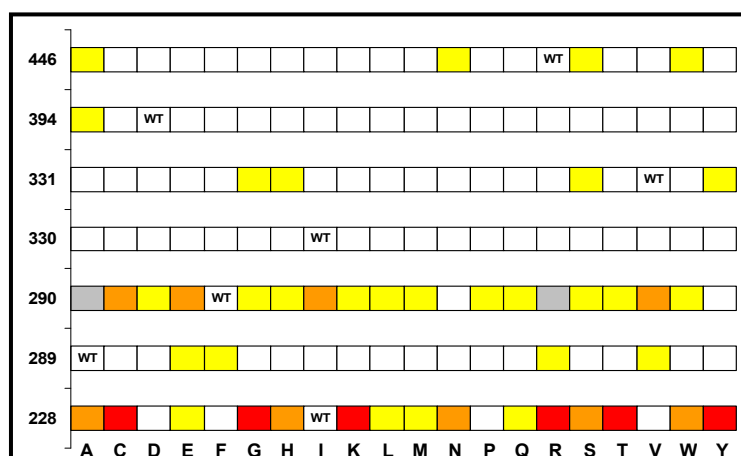


FIGURE 49| Improved single-mutants obtained from previous study¹⁵ compared to wtAS

Rows indicate the 7 mutated positions. Columns represent the 20 possible amino acid mutations (including the wt). (**Yellow**): Improved mono-mutants for D-GlcpNAc (**D**) glycosylation (**Red**): Improved mono-mutants for α -L-RhapOMe (**A'**) glycosylation. (**Orange**): Improved mono-mutants for α -L-RhapOMe (**A'**) and D-GlcpNAc (**D**) glycosylation. (**grey**): Not determined

To enlarge the exploration of subsite +1 sequence space, our strategy was thus to generate four libraries of double-mutants on adjacent or distant positions before attempting a full recombination involving all positions. Indeed, such double-mutants are not easily accessible from error-prone PCR because no polymerase is able to generate successive errors on a sequence and also because of the degenerescence of the genetic code. Libraries **1**, **2** and **3** targeting positions 228-229, 289-290, 330-331, respectively, were constructed by PCR using a set of degenerate primers designed to generate 24 possible codons for each position (**FIGURE 50**). Only His, Gln and Pro residues were not represented but we considered this choice to be a good compromise. In this way, 289 (17*17) possible double mutations were encoded. To ensure a good representation of all the variants, it was estimated that 3000 recombinant clones had to be screened for each library. The strategy employed for the construction of library **4**, corresponding to the combination of positions 228 and 290, was slightly different. PCR was carried out using successively two sets of primers, each of them encoding the 19 possible amino acids. All possible amino acid changes are considered possible in this library.

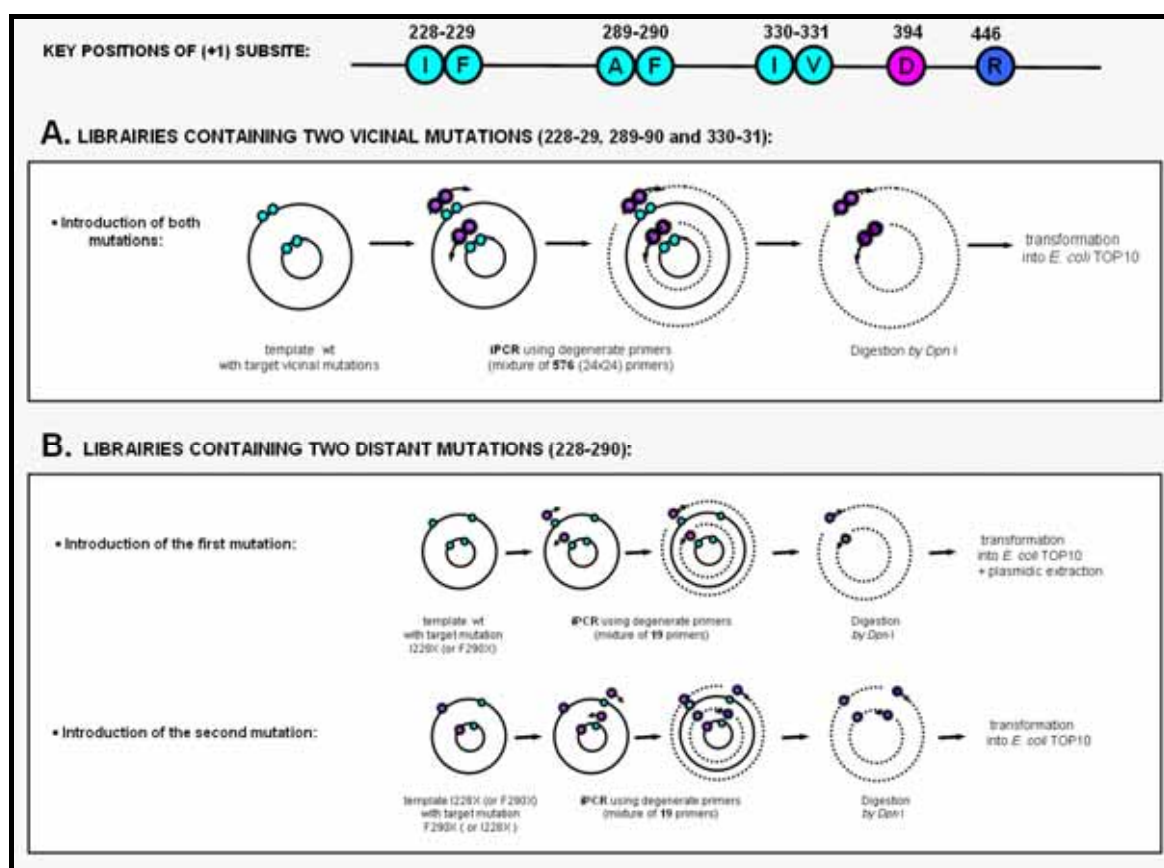


FIGURE 50| Strategy adopted for the construction of the four double-mutant libraries
 Strategy adopted for construction of (A) libraries 1, 2 and 3 (B) library 4

The libraries were then screened to determine both their size and the ratio of active clones using the pH indicator screen (**TABLE 14**). A total of 20 000 recombinant clones were screened. Over 3000 recombinant clones were obtained for each library except for library 2, for which only 2000 clones were generated. Altogether, 1716 clones were found to be able to use sucrose as a glucosyl donor. Depending on the library, the ratio of active clones was comprised between 2 and 29%. Based on the color change of BBT, 384 clones induced a strong pH change whereas 1332 were found to be moderately active.

TABLE 14| Number of clones and active clones obtained for the four libraries

| | Library 1 228-229 | Library 2 289-290 | Library 3 330-331 | Library 4 228-290 |
|---|------------------------------|------------------------------|------------------------------|------------------------------|
| Number of clones | ~ 8000 | ~ 2000 | ~ 3200 | ~ 8000 |
| Number of active clones isolated (%) | 180 (2,3%) | 576 (29%) | 576 (18%) | 384 (4,8%) |
| Number of active clones (yellow) | 96 | 96 | 96 | 96 |
| Number of moderately active clones (green) | 84 | 480 | 480 | 288 |

C.3. Isolation of improved Mutants for EA' and ED synthesis

All clones producing sucrose-utilizing variants were then picked and cultured in 96-well microplates. Following cell lysis, the acceptor reaction in the presence of α -L-RhapOMe (**A'**) or D-GlcpNAc (**D**) was then carried out in microplate format. The acceptor reaction products were analyzed by HPLC analysis at a cadence of about 10 reactions /hour. Analysis time was reduced to 5 min and 7 min for detection of α -D-Glcp-(1→3)- α -L-RhapOMe (**EA'**) and α -D-Glcp-(1→4)- α -D-GlcpNAc (**ED**) respectively. When disaccharide formation was detected, complementary analyses were performed to determine sucrose consumption and disaccharide formation with more accuracy to calculate the glycosylation yield.

Improved Mutants for EA' synthesis

Of the 1716 tested mutants, 12 were able to catalyze **EA'** disaccharide synthesis (**TABLE 15**). Most of them are mono-mutants (7). Some of them were identified twice (I228S, F290V) or four times (I228N). Five double-mutants were also detected. All of them show a mutation at position 228; a position that was previously identified as critical for A' recognition. Single mutation at positions 229 or 290 also favours A' recognition, although to a lesser extent. Of note, the glycosylation rate obtained for the four I228N mutants appears to change from 4 to 16%. This

variation reflects the poor reproducibility of the glycosylation assay in microtiter plates. Microtiter experiments combined with HPLC analysis are designed for high-throughput screening. Such experiments allow the detection of positive mutants but remain inappropriate to determine with precision glycosylation rates. Consequently, to evaluate the impact of the mutations on the catalytic efficiency for **EA'** formation only, the positive mutants, particularly the double-mutants I228C-F229L and I228V-F229Y, will have to be produced in larger scale, purified and further characterized.

TABLE 15| Improved mutants for EA' synthesis

| Mutants | Library number | Single-Mutant (SM)/ Double-Mutant (DM) | Amino acid Mutation(s) | Mutation(s) (codon) | Glycosylation Yield to synthesize EA' |
|-------------|----------------|--|------------------------|---------------------|---------------------------------------|
| wtAS | | | | | 0% |
| DM1_P2_F1 | 1 | SM | I228N | AAC | 16% |
| DM1_P1_D11 | 1 | DM | I228C F229L | TGC CTC | 11% |
| DM1_P1_H3 | 1 | DM | I228V F229Y | GTG TAC | 11% |
| DM1_P2_G5 | 1 | SM | I228N | AAC | 10% |
| DM4_P1_D1 | 4 | SM | I228H | CAC | 7% |
| DM4_P1_H10 | 4 | SM | F290V | GTT | 5% |
| DM1_P2_F8 | 1 | SM | I228S | AGC | 4% |
| DM4_P1_C11 | 4 | SM | I228A | GCC | 4% |
| DM4_P1_D8 | 4 | SM | I228S | AAC | 4% |
| DM4_P1_H9 | 4 | SM | I228N | AAC | 4% |
| DM4_P3_D8 | 4 | DM | I228A F290L | GCC CTT | 4% |
| DM4_P3_E4 | 4 | DM | I228L F290C | CTC TGT | 4% |
| DM4_P3_F10 | 4 | SM | F290V | GTT | 4% |
| DM1_P1_E2 | 1 | SM | I228M | ATG | 3% |
| DM1_P1_E7 | 1 | SM | F229N | AAC | 3% |
| DM4_P1_G4 | 4 | SM | I228N | AAC | 3% |
| DM4_P3_F7 | 4 | DM | I228P F290C | CCC TGT | 3% |

Improved Mutants for ED synthesis

Of the active mutants isolated from libraries **1** to **4**, 25 were able to catalyze **ED** disaccharide synthesis with a glycosylation rate higher than 10% (**TABLE 16**). Most of them are double-mutants (15) or new single-mutants (4: F290R, F229L, F229M and F229V). 18 mutants were also detected in library **2**, but not sequenced (

TABLE 17). These mutants will now be produced on larger scale for comparison and selection of the best one.

TABLE 16| Improved mutants for ED synthesis (Libraries 1 and 4)

| Mutants | Library number | Single-Mutant (SM) Double-Mutant (DM) | Amino acid Mutation(s) | Mutation(s) (codon) | Glucosylation Yield to synthesize ED |
|-------------|----------------|--|---------------------------|------------------------|--|
| wtAS | | | | | 2% |
| DM4_P1_G7 | 4 | DM | I228V F290V | GTC GTT | 100% |
| DM4_P1_H10 | 4 | SM | F290V | GTT | 100% |
| DM4_P1_H11 | 4 | SM | F290R | CGT | 100% |
| DM4_P1_C7 | 4 | DM | I228V F290K | GTC AAA | 89% |
| DM4_P1_C1 | 4 | DM | I228A F290H | GCC CAT | 88% |
| DM1_P1_C1 | 1 | SM | F229V | GTG | 88% |
| DM1_P1_C12 | 1 | SM | F229M | ATG | 79% |
| DM1_P1_D11 | 1 | DM | I228C F229L | TGC CTC | 78% |
| DM4_P3_B8 | 4 | DM | I228T F290K | ACC AAA | 72% |
| DM4_P1_C10 | 4 | DM | I228V F290H | GTC CAT | 61% |
| DM4_P1_A6 | 4 | DM | I228T F290H | ACC CAT | 56% |
| DM1_P2_F7 | 1 | DM | I228V F229L | GTG TTG | 56% |
| DM4_P2_A2 | 4 | SM | F290L | CTT | 55% |
| DM4_P2_H10 | 4 | SM | F290G | GGT | 53% |
| DM4_P2_C2 | 4 | DM | I228K F290W | AAG TGG | 50% |
| DM1_P2_C8 | 1 | SM | F229L | CTC | 49% |
| DM4_P2_B10 | 4 | DM | I228V F290V | GTC GTT | 47% |
| DM4_P2_G4 | 4 | DM | I228K F290L | AAG CTT | 47% |
| DM4_P3_C8 | 4 | SM | F290H | CAT | 46% |
| DM4_P2_H4 | 4 | SM | F290K | AAA | 42% |
| DM4_P2_D12 | 4 | SM | F290K | AAA | 41% |
| DM1_P1_H2 | 1 | DM | I228L F229L | TTG CTC | 35% |
| DM1_P2_F2 | 1 | DM | I228V F229M | GTG ATG | 29% |
| DM1_P2_C3 | 1 | DM | I228M F229M | ATG ATG | 28% |
| DM4_P3_G9 | 4 | DM | I228V F290R | GTC CGT | 25% |
| DM4_P3_F7 | 4 | DM | I228P F290C | CCC TGT | 22% |
| DM4_P2_G9 | 4 | SM | F290L | CTT | 20% |
| DM4_P4_B12 | 4 | SM | F290I | ATT | 16% |

TABLE 17| Improved and not sequenced mutants for ED synthesis (Library 2)

| Mutants | Library number | Glucosylation Rate to synthesize ED |
|------------|----------------|--|
| DM2_P1_A2 | 2 | 84% |
| DM2_P1_H10 | 2 | 59% |
| DM2_P1_E12 | 2 | 47% |
| DM2_P1_E7 | 2 | 44% |
| DM2_P1_B6 | 2 | 34% |
| DM2_P2_H12 | 2 | 34% |
| DM2_P1_D12 | 2 | 32% |
| DM2_P1_C12 | 2 | 30% |
| DM2_P3_H7 | 2 | 28% |
| DM2_P3_C12 | 2 | 26% |
| DM2_P4_A3 | 2 | 26% |
| DM2_P2_A2 | 2 | 20% |
| DM2_P1_C11 | 2 | 20% |
| DM2_P2_B12 | 2 | 19% |
| DM2_P3_F5 | 2 | 17% |
| DM2_P3_G1 | 2 | 16% |
| DM2_P1_D8 | 2 | 15% |
| DM2_P2_H11 | 2 | 13% |

D. Discussion

In a previous work, we used targeted randomization at the +1 acceptor binding site of amylosucrase to successfully modify the enzyme specificity towards two acceptors, α -L-RhapOMe (**A'**) and D-GlcpNAc (**D**). Guided by a molecular modelling study, a first library of single-mutants was generated and several improved variants were isolated. In the current study, our goal was to attempt the generation of further improved variants displaying a better efficiency towards the formation of the disaccharide of interest. In particular, we focused our efforts on the construction of double-mutant libraries, involving the single mutation positions identified to be beneficial for the recognition of α -L-RhapOMe (**A'**) or D-GlcpNAc (**D**). Although the choice of the mutation positions was rationally driven, the possibility for amino acid changes was broad and large libraries had to be created and screened to select variants able to use sucrose as glucosyl donor and to glucosylate the molecule of interest (**A'** or **D**). As sucrose usage remains a pre-requisite, the libraries were first screened to identify clones able to use sucrose substrate. For this purpose a screen was developed. This test was preferred to the selection method that was previously¹² reported. This latter consisted in growing recombinant clones on solid mineral medium containing sucrose as the sole carbon source. Due to the fact that *E. coli* strains do not use sucrose as substrate, only the clones producing active AS were able to grow. However, using this method, it would not have been possible to estimate the ratio of active recombinants. In addition, 5 days of culture at 30°C would have been necessary to obtain sufficient cell growth. The new test developed in this work is far more rapid. Only two days were required to isolate positive clones on colony levels. In addition, it allowed a better analysis of the library in terms of size and proportion of active/inactive clones. When implemented in microtiter plates, the method was shown to be much more sensitive than the DNS assay previously described. In addition, this test could be applied to the library screening of all types of sucrose utilizing enzymes. Application of the BBT screen enabled to reduce the size of the library by 92% and thus significantly limit the size of the library to be screened for acceptor recognition. The results obtained further show that double mutations may often lead to inactive mutants. Sequencing of inactive clones combined to analysis of their structural models would be at this stage very informative to improve the rational construction of mutants.

The HPLC screening enabled us to isolate 11 clones that were able to recognize α -L-RhapOMe. These variants show a totally new specificity. Among them, 4 had previously been identified in our previous work¹⁵. Notably, the strategy developed, here, did not permit to isolate I228G, I228W and I228Y single-mutants which were previously generated by site-directed mutagenesis. This may reflect either a bias in the library construction or be due to the very low activity of these

single-mutants when assayed in the presence of sucrose alone without acceptor (TABLE 18). To avoid this problem, the acceptor could be added to sucrose during the BBT test, what could further increase the sensitivity of this test. In addition, only 4 double-mutants were isolated, some of them being equivalent to the mono-mutants previously isolated. They will have now to be further characterized to determine whether or not they are more efficient and show an interest compared to the isolated improved single-mutants.

TABLE 18| Sucrose activity of the most improved mono-mutants for EA' synthesis

| | Sucrose activity (U/mL) | % of activity in comparison with wtAS | Glucosylation Rate to synthesize EA' ¹⁵ (%) |
|--------------|-------------------------|---------------------------------------|--|
| wtAS | 2.7 | 100 | 0 |
| I228A | 2.56 | 94 | 11 |
| I228N | 0.57 | 21 | 14 |
| I228H | 0.38 | 14 | 18 |
| I228S | 0.35 | 13 | 9 |
| I228Y | 0.18 | 7 | 23 |
| I228W | 0.08 | 3 | 17 |
| I228G | 0.01 | 0.4 | 11 |

Regarding D-Glc_pNac glucosylation, we have been able to isolate 25 improved mutants and 15 of them are double-mutants. This clearly indicates that libraries were well constructed. Notably, most of the double-mutants combine 228 and 290 positions and recurrent mutations (F290C, G, H, K, L, H and V) previously identified as the most beneficial. The characterization of these mutants is now under progress to determine their degree of improvement in terms of catalytic efficiency and ability to produce the disaccharide of interest without producing additional products of glucosylation.

In conclusion, a new efficient screening method has been developed to analyze libraries of sucrose-utilizing enzymes and isolate active mutants. Applied to the screening of more than 20 000 mutants, it enabled the isolation of 54 variants of potential interest for the production of the disaccharides of interest for the chemo-enzymatic synthesis of *Shigella flexneri* antigenic oligosaccharides. These variants will now be further characterized. This test will also be applied to screen libraries of larger size obtained by full recombination of the identified positions. In addition, the set of data acquired in this work is large. Crystallographic and molecular modelling studies will be carried out to better understand the effect of mutation(s) on the three-dimensional structure of amylosucrase and its catalytic properties as well as to better characterize the structural features involved in the recognition of the non-natural acceptors and provide a molecular interpretation of the mutants properties. Our aim is to make use of this information to develop novel computational methodologies enabling a virtual search of the sequence space.

E. Methods

E.1. Plasmids, bacterial strains and chemicals

Plasmid pGST-AS, derived from the pGEX-6P-3 (GE Healthcare Biosciences) and containing the *N. polysaccharea* amylosucrase encoding gene² was used for the construction of the AS double-mutant libraries. Fusion DNA-polymerase was purchased from Finnzymes (Espoo, Finland), and *DpnI* restriction enzyme from New England Biolabs (Beverly, MA, USA). Oligonucleotides were synthesised by Eurogenetec (Liege, Belgium). DNA extraction (QIASpin) and purification (QIAQuick) columns were purchased from Qiagen (Chatsworth, CA). *E. coli* TOP 10 electrocompetent cells (Invitrogen, Carlsbad, USA) were used as host for the plasmid library transformation and gene expression. DNA sequencing was performed by Cogenics (Meylan, France). All positive clones for **A'** and **D** glycosylation were sequenced on the mutated region (~600 bp) using the primer pGEX_int: CCAACGAACACGAATGGGC.

Ampicillin (Amp), lysozyme, and isopropyl-β-D-thiogalactopyranoside (IPTG) were purchased from Euromedex (Souffelweyersheim, France); Bromothymol Blue sodium salt, sucrose and *N*-acetyl-D-glucosamine (D-GlcpNAc) from Sigma-Aldrich (Saint-Louis, MO, USA). Reference disaccharides α-D-Glcp-(1→3)-α-L-Rhap-OMe²² and α-D-Glcp-(1→4)-α-L-Rhap-OMe²³ were available in the UCB (L. Mulard). Reference disaccharides α-D-Glcp-(1→4)-D-GlcpNAc and Glcp-(1→6)-D-GlcpNAc were enzymatically synthesized and characterized at LISBP²⁴.

E.2. Construction of libraries containing two vicinal mutations (libraries 1-3)

Site-saturation mutagenesis, focused on the vicinal positions (I228-F229, A289-F290, I330-V331) of AS +1 acceptor subsite was carried out using pGST-AS G537D as vector template. It was checked that this mutation had no impact on the native enzyme catalytic properties

We designed three partial overlapping primer pairs surrounding double codons. Each of these codons was replaced in the primers with degenerate DNS or DNW sequence, where D=A,G or T; N= A,C, G or T; S= C or G and W=A or T (**TABLE 19**). Such degenerate primers were designed to generate 24 codons encoding 17 amino acids (all except His (H),Gln (Q) and Pro (P))

TABLE 19 | Degenerate primers used for the construction of librairies (I228-F229, A289-F290, I330-V331)

| Primer Name | Nucleotide Sequence |
|-------------|--|
| 228-229 for | 5'-CTG CGC GAA DNS DNS CCC GAC CAG CAC CCG GGC G-3' |
| 228-229 rev | 5'-CTG GTC GGG SND SND TTC GCG CAG GGT GCG GTC G-3' |
| 289-290 for | 5'-GCG GTT DNS DNW ATT TGG AAA CAA ATG GGG ACA AGC TGC G-3' |
| 289-290 rev | 5'-CCA AAT WND SND AAC CGC ATC CAT ACG CAG GAT GTC AAC GCC-3' |
| 330-31 for | 5'-TCC GAA GCC DNS DNS CAC CCC GAC CAA GTC GTC C-3' |
| 330-31 rev | 5'- GGG GTG SND SND GGC TTC GGA TTT GAA GAA CAC GGC-3' |

PCR amplification was carried out on the whole plasmid with Phusion DNA-polymerase (1 U) for 30 cycles (98°C, 10s; 75°C, 20s; 57°C,15s; 72°C, 5min). The DNA was digested with DpnI to eliminate methylated parental template and purified using Qiaquick spin column, following manufacturer's recommendations. *E. coli* TOP10 was transformed by electroporation with 4µL of each plasmid library using standard procedures.

E.3. Construction of library containing two distant mutations (library 4)

pGST-AS G537D was also used as vector template for library (I228-F290) construction. First, for each targeted position, we pooled, the 2*19 complementary primers previously constructed by site-directed mutagenesis¹⁵ in equimolar ratio, and formed thus two pairs of degenerate primers, for positions 228 and 290, respectively. Then, we proceeded following two PCR amplification steps using the same PCR method, DNA treatment and cell transformation as described above. The first step consisted in generating individually two mono-mutants libraries by saturation mutagenesis. After plasmid library extraction, a second PCR step was carried out to introduce the second mutation.

E.4. Development of the BBT test in liquid medium

Mono-mutant library was cultured in 96-well microplates as previously described²⁵. After thawing at room temperature, 80µL of the lysed cells were transferred into a new microtiter plate. Enzymatic reaction was carried out by adding 80 µL of sucrose to a final concentration of 146 mM followed by incubation at 30°C for 24h.

BromoThymol Blue (BBT) test: Medium acidification due to acid production consecutive to fructose release by action of active amylosucrase onto sucrose was determined by adding 100µL of the reaction mixture to 10µL BBT solution (0.25% (g/v) dissolved in 1% ethanol) in a polystyrene plate. Absorbance was measured at 620 nm with the sunrise microplate reader.

The fructose production was also followed by DNS assay²¹ for comparison. A volume of 100µL of reaction mixture was added to 100µL of dinitrosalicylic acid (DNS) reagent in a propylene plate, incubating at 95°C for 10 minutes. 100µL of this mixture and 100µL H₂O were transferred in a polystyrene microtiter plate. Absorbance at 540 nm was measured.

For Iodine staining assay²⁶, amylose formation was detected by adding 10µL of iodine solution (100 mM KI, 6mM I₂, 0.02 M HCl) to 50µL of the reaction mixture. Absorbance was measured at 550 nm, the iodine forming a blue complex with the helical form of amylose.

E.5. Development of the BBT test at colony level

On day 1, libraries were transformed into electrocompetent *E. coli* TOP10 and plated on membranes (Durapore® membrane filters, 022µm GV- Millipore, Ireland) which were previously soaked in physiological water and placed onto 22 cm square plates (Corning, USA) containing 200 mL solid LB agar, 1% glycerol and 100 µg/mL ampicillin.

On day 2, after overnight growth at 37°C (~16h), each membrane was transferred on a 22 cm square plate containing inducing medium (200 mL solid LB agar + 1 mM IPTG, 100 µg/mL ampicillin) supplemented with 50 g/L sucrose and stained in blue by adding 50 mM Tris-HCl pH=7,5 and 0,1 g/l pH indicator BromoThymol Blue (BBT, dissolved in 1% ethanol). The plates were incubated overnight (24h) at 30°C.

On day 3, active clones (green and yellow) were isolated in microplates containing 200 µL LB medium, 12% glycerol and 100 µg/mL ampicillin. After 24h of growth at 30°C, they were stored at -20°C and -80°C. Inactive clones (blue) were scraped, cultivated and stored in the same conditions.

E.6. HPLC screening of variants able to glycosylate α -L-RhaOMe and D-GlcNAc_p

The protocol, as previously described¹⁵, was used to enable the rapid identification of (i) clones able to glycosylate α -L-RhaOMe (**A'**), and (ii) clones for which D-GlcNAc_p (**D**) glycosylation was improved. To obtain larger amounts of enzymes and facilitate detection of glycosylated compounds by HPLC, mutants were produced in 96-DeepWell Format plates (Abgene, UK). Storage microplates containing mono-mutants were thawed and replicated to inoculate a starter culture in 96-well microplates containing, in each well, 150µL LB medium supplemented with ampicillin (100µg/mL). After growth for 24h at 30°C under agitation (200 rpm), plates were duplicated into 96-Deep Well plates containing, in each well, 1.1 mL LB medium supplemented with ampicillin (200µg/mL) and IPTG (1mM) to induce GST-AS expression. Cultures were then grown for 24h at 30°C under agitation (200 rpm). Plates were centrifuged (20 min, 3700g, 4°C) and the supernatant was removed. The cell pellet was resuspended in 200µL of lysozyme (0.5 mg/mL), followed by freezing at -80°C for 8 to 12h. After thawing at room temperature, 100µL of sucrose and 100µL of acceptor (each at a final concentration of 73 mM) were added to each well. Enzymatic reaction was incubated at 30°C during 24h under agitation. The DeepWell plates were then centrifuged (20 min, 3700g, 4°C) and 300µL of the supernatant was transferred to a filter micro-plate (Glass fiber membrane, PS, 0.25 mm pore- Corning, USA) to be clarified. Supernatant filtration was carried out by centrifugation of the filter micro-plate (5 min, 2000g, 4°C) into a novel microplate for HPLC screening.

Acceptor reaction products were analyzed by HPLC analysis using two reversed phase analytical columns: (1) a monolithic C18 column (ONYX, 125x4 mm, 3 μ m) maintained at 30°C and eluted with ultra-pure water at 2 mL/min to detect **EA'** formation (analysis time: 5min vs. 20 min) (2) a C18-AQ column (Bischoff C18, 125x4 mm, 3 μ m) kept at room temperature and eluted with 0.6 mL/min of ultra-pure water to detect **ED** formation (analysis time: 7 min vs. 20 min)

To evaluate the efficiency of the glucosylation reaction, we carried out complementary HPLC analyses with a Biorad HPLC Carbohydrate Analysis column. This HPX-87K column (maintained at 65°C, and eluted at a flow rate of 0.6 mL/min with ultra-pure water). By this way, it was possible to measure sucrose consumption by RI detection and determine mutant ability to synthesize the desired disaccharide **EA'** (or **ED**) (ie. % Glc transferred onto acceptor derivatives = $Q(\text{Glucosyl units transferred onto A' (or D)}) / Q(\text{Glucosyl units transferred from initial sucrose})$)

F. References

1. Butcher, V.; Welsh, T.; Willmitzer, L.; Kossmann, J., Cloning and characterization of the gene for amylosucrase from *Neisseria polysaccharea*: production of a linear α -1,4-glucan. *J Bacteriol*, **1997**, 179, (10), 3324-30.
2. De Montalk, G. P.; Remaud-Simeon, M.; Willemot, R. M.; Planchot, V.; Monsan, P., Sequence analysis of the gene encoding amylosucrase from *Neisseria polysaccharea* and characterization of the recombinant enzyme. *J Bacteriol*, **1999**, 181, (2), 375-81.
3. Potocki de Montalk, G.; Remaud-Simeon, M.; Willemot, R. M.; Sarcabal, P.; Planchot, V.; Monsan, P., Amylosucrase from *Neisseria polysaccharea*: novel catalytic properties. *FEBS Lett*, **2000**, 471, (2-3), 219-23.
4. Potocki de Montalk, G.; Remaud-Simeon, M.; Willemot, R. M.; Monsan, P., Characterisation of the activator effect of glycogen on amylosucrase from *Neisseria polysaccharea*. *FEMS Microbiol Lett*, **2000**, 186, (1), 103-8.
5. Rolland-Sabaté, A.; Colonna, P.; Potocki-Véronèse, G.; Monsan, P.; Planchot, V., Elongation and insolubilisation of α -glucans by the action of *Neisseria polysaccharea* amylosucrase. *Journal of Cereal Science*, **2004**, 40, (1), 17-30.
6. Bounaix, M. S., *unpublished results*.
7. Pizzut-Serin, S.; Potocki-Veronese, G.; van der Veen, B. A.; Albenne, C.; Monsan, P.; Remaud-Simeon, M., Characterisation of a novel amylosucrase from *Deinococcus radiodurans*. *FEBS Lett*, **2005**, 579, (6), 1405-10.
8. Emond, S., Characterisation of a novel amylosucrase from *Deinococcus geothermalis*. *unpublished results*.
9. Sarcabal, P.; Remaud-Simeon, M.; Willemot, R.; Potocki de Montalk, G.; Svensson, B.; Monsan, P., Identification of key amino acid residues in *Neisseria polysaccharea* amylosucrase. *FEBS Lett*, **2000**, 474, (1), 33-7.
10. Albenne, C.; de Montalk, G. P.; Monsan, P.; Skov, L. K.; Mirza, O.; Gajhede, M.; Remaud-Simeon, M., Site-directed mutagenesis of key amino acids in the active site of amylosucrase from *Neisseria polysaccharea*. *Biologia, Bratislava*, **2002**, 57, 119-128.
11. 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., Molecular basis of the amylose-like polymer formation catalyzed by *Neisseria polysaccharea* amylosucrase. *J Biol Chem*, **2004**, 279, (1), 726-34.
12. van der Veen, B. A.; Potocki-Veronese, G.; Albenne, C.; Joucla, G.; Monsan, P.; Remaud-Simeon, M., Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett*, **2004**, 560, (1-3), 91-7.
13. van der Veen, B. A.; Skov, L. K.; Potocki-Veronese, G.; Gajhede, M.; Monsan, P.; Remaud-Simeon, M., Increased amylosucrase activity and specificity, and identification of regions important for activity, specificity and stability through molecular evolution. *Febs J*, **2006**, 273, (4), 673-81.
14. Emond, S.; André, I.; Jaziri, K.; Potocki-Véonèse, G.; Mondon, P.; Bouayadi, K.; Kharrat, H.; Monsan, P.; Remaud-Siméon, M., Combinatorial engineering to enhance amylosucrase thermostability. *unpublished results*.
15. Champion, E.; André, I.; Moulis, C.; Morel, S.; Boutet, J.; Descroix, K.; P., M.; Mulard, L. A.; M., R.-S., Tailored transglucosidases for innovative chemo-enzymatic synthesis of antigenic oligosaccharides. *unpublished results*.
16. Davies, G. J.; Wilson, K. S.; Henrissat, B., Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*, **1997**, 321 (Pt 2), 557-9.

17. Lindberg, A. A.; Karnell, A.; Weintraub, A., The lipopolysaccharide of *Shigella* bacteria as a virulence factor. *Rev Infect Dis*, **1991**, 13 Suppl 4, S279-84.
18. Simmons, D. A., The structural basis of serological specificity in *Shigella flexneri* O-antigens. *Biochem Soc Trans*, **1990**, 18, (6), 1271-2.
19. Kotloff, K. L.; Winickoff, J. P.; Ivanoff, B.; Clemens, J. D.; Swerdlow, D. L.; Sansonetti, P. J.; Adak, G. K.; Levine, M. M., Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ*, **1999**, 77, (8), 651-66.
20. Levine, M. M.; Kotloff, K. L.; Barry, E. M.; Pasetti, M. F.; Sztein, M. B., Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol*, **2007**, 5, (7), 540-53.
21. Sumner, J.; Howell, S. A., A method for the determination of invertase activity. *J. Biol. Chem*, **1935**, 108, 51-54.
22. Bakinovskii, L. V.; Gomtsian, A. R.; Bairamova, N. E.; Kochetkov, N. K., [Synthesis of oligosaccharide fragments of O-specific *Shigella flexneri* polysaccharides. II. Synthesis of trisaccharide $\alpha\text{Glc}(1-3)\alpha\text{Rha}(1-2)\alpha\text{RhaOMe}$ and tetrasaccharide $\beta\text{GlcNAc}(1-2)\alpha\text{Glc}(1-3)\alpha\text{Rha}(1-2)\alpha\text{RhaOMe}$]. *Bioorg Khim*, **1985**, 11, (2), 254-63.
23. Mulard, L. A.; Costachel, C.; Sansonetti, P. J., Synthesis of the methyl glycosides of a di- and two trisaccharide fragments specific for the *Shigella flexneri* serotype 2a O-antigen. *J. Carbohydr. Chem.*, **2000**, 19, (7), 849-877.
24. Champion, E.; Morel, S.; Mulard, L. A.; Monsan, P.; Remaud-Siméon, M.; André, I., Enzymatic synthesis of disaccharide motives entering in the composition of bacterial polysaccharides by use of glucansucrases. *unpublished results*.
25. Emond, S.; Potocki-Veronese, G.; Mondon, P.; Bouayadi, K.; Kharrat, H.; Monsan, P.; Remaud-Simeon, M., Optimized and Automated Protocols for High-Throughput Screening of Amylosucrase Libraries. *J Biomol Screen*, **2007**.
26. Banks, W.; Greenwood, C. T.; Khan, K. M., The interaction of linear, amylose oligomers with iodine *Carbohydr. Res.*, **1971**, 17, (1), 25-33.

CONCLUSION

Avant de conclure, nous souhaitons brièvement rappeler les motivations et les objectifs de notre étude.

Ces travaux de thèse ont été proposés dans le cadre d'un projet ambitieux, soutenu par l'ANR visant à développer de nouvelles voies de synthèse chimio-enzymatique de vaccins contre certains sérotypes de *Shigella flexneri*, responsables du décès de milliers de personnes chaque année. Récemment, la synthèse par voie chimique d'haptènes oligosaccharidiques mimant les motifs antigéniques du lipopolysaccharide de *Shigella flexneri* sérotype 2a a permis d'élaborer un glycoconjuqué à potentiel vaccinal prometteur. Cette approche vaccinale pourrait en principe être étendue à d'autres sérotypes prévalents tels que les sérotypes 1b et 3a. Cependant, pour ces sérotypes, le procédé de synthèse implique l'introduction d'une étape de glycosylation de résidus *N*-acétyl-D-glucosamine (1b) et L-rhamnose (3a), qui est problématique et difficile à réaliser par voie chimique. Nous avons donc envisagé de simplifier le schéma de synthèse chimique en intégrant, en début de synthèse, une étape de glycosylation enzymatique. Evidemment, cela suppose de disposer d'enzymes capables de glycosyler efficacement et spécifiquement les deux résidus cibles porteurs de groupes protecteurs compatibles avec la poursuite de la synthèse. Or, l'existence d'enzymes parfaitement adaptées à cette fonction est peu probable car les produits de glycosylation recherchés n'ont pas été observés, à ce jour, à l'état naturel.

Dans ce contexte, notre étude a donc eu pour ambition d'exploiter les outils d'ingénierie enzymatique afin de concevoir des transglucosidases spécifiquement adaptées à la glycosylation d'accepteurs qui naturellement ne sont pas (ou très peu) reconnus.

L'approche que nous avons adoptée s'est décomposée en deux étapes :

- Evaluer le potentiel de quatre glucane-saccharases, spécifiques de la synthèse de liaisons osidiques distinctes à glycosyler les accepteurs cibles et caractériser les produits les plus abondants résultant de la glycosylation de l'accepteur.
- Mettre en œuvre une stratégie raisonnée d'ingénierie, associant ingénieries rationnelle et combinatoire, pour un remodelage de l'amylosaccharase de *Neisseria polysaccharea* adapté aux accepteurs cibles.

Nous proposons à présent de dresser le bilan des différents résultats obtenus et de dégager pour chacun des points abordés les perspectives de cette étude.

1- Glucosylation des accepteurs cibles par les glucane-saccharases de type sauvage :

En testant quatre glucane-saccharases de spécificité différente sur les deux types d'accepteurs cibles, nous avons pu caractériser et/ou isoler **six nouveaux produits de glucosylation** : l' α -D-glucopyranosyl-(1 \rightarrow 6)-*N*-acétyl-D-glucosamine, l' α -D-glucopyranosyl-(1 \rightarrow 4)-*N*-acétyl-D-glucosamine, l' α -D-glucopyranosyl-(1 \rightarrow 1)- β -L-rhamnopyranoside, l' α -D-glucopyranosyl-(1 \rightarrow 4)- α -D-glucopyranosyl-(1 \rightarrow 1)- β -L-rhamnopyranoside, le methyl α -D-glucopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranoside et le methyl α -D-glucopyranosyl-(1 \rightarrow 3)- α -L-rhamnopyranoside.

Quatre des cinq motifs disaccharidiques synthétisés interviennent dans la composition de polysaccharides bactériens, justifiant l'intérêt accordé à ces enzymes pour la synthèse de motifs osidiques antigéniques. Notamment, soulignons qu'un taux de glucosylation de 61% a été obtenu pour la synthèse de l' α -D-glucopyranosyl-(1 \rightarrow 6)-*N*-acétyl-D-glucosamine, qui compose, en partie, l'antigène polysidique de *Shigella flexneri* sérotype 4a et 4b.

Ces travaux préliminaires ont permis d'élargir la liste des accepteurs reconnus par les glucane-saccharases. Ils ont également montré les limites pressenties d'utilisation des enzymes sauvages qui n'allient pas efficacité de glucosylation et régiospécificité désirée. Une stratégie de remodelage du site actif de l'amylosaccharase de *Neisseria polysaccharea* a été mise en place pour favoriser les réactions de glucosylation ciblées.

2- Ingénierie semi-rationnelle de l'amylosaccharase de *Neisseria polysaccharea*

Notre approche a fait intervenir des méthodes d'analyse des interactions AS-Substrat afin d'identifier les acides aminés impliqués dans la reconnaissance spécifique des accepteurs (sous-site +1). Ces travaux de modélisation moléculaire ont été effectués sur les 2 motifs glucidiques ciblés : α -D-Glcp-(1 \rightarrow 3)- α -L-Rhap (sérotype 3a) et α -D-Glcp-(1 \rightarrow 4)-D-GlcNAcp (sérotype 1b). Sept positions clés (228, 289, 290, 330, 331, 394 et 446) du sous-site de fixation de l'accepteur ont été choisies comme cibles. Elles ont été individuellement et intégralement explorées en remplaçant les 7 résidus sélectionnés par les 19 autres acides aminés possibles. Une banque de 133 mono-mutants a ainsi été construite par mutagenèse dirigée. Le criblage de cette banque sur les accepteurs cibles a conduit à l'identification de 43 mono-mutants d'intérêt. Les positions 228 et 290 se sont révélées être des positions particulièrement stratégiques pour l'accommodation de l' α -methyl L-rhamnose et de la *N*-acétyl-D-glucosamine respectivement, avec les spécificités ciblées : 9 variants (mutés en position 228) ont permis de créer une nouvelle spécificité vis-à-vis

du premier accepteur et 5 variants (mutés en position 290) glucosylent le deuxième accepteur avec des taux supérieurs à 50% (vs. 2% pour l'amylosaccharase sauvage).

Ces travaux font la démonstration du potentiel de la stratégie d'ingénierie semi-rationnelle adoptée et décrivent des changements de spécificité spectaculaires qui n'ont jamais été rapportés auparavant pour ce type de transglucosidases. Depuis la résolution récente des structures de GTF-A et GTF-180, une démarche similaire peut être désormais envisagée pour les glucane-saccharases de la famille GH70.

D'autre part, la banque composée de 66 nouvelles transglucosidases fonctionnelles et modifiées au niveau du site de reconnaissance de l'accepteur pourra maintenant être exploitée pour la glucosylation d'autres molécules d'intérêt.

3- Caractérisation des deux mutants améliorés I228Y et F290K

Nous avons choisi d'analyser plus finement les propriétés catalytiques des mono-mutants les plus améliorés, I228Y et F290K, pour la glucosylation des 2 accepteurs cibles : l' α -methyl L-rhamnose et l' α -allyl N-acétyl-D-glucosamine, respectivement. Ces mutants ont été produits à plus large échelle (de l'ordre du litre), purifiés et caractérisés en présence de saccharose et des accepteurs cibles.

- Le mutant I228Y **ne présente pas d'activité polymérase** en présence de saccharose. En outre, il est capable de **glucosyler l' α -methyl L-rhamnose et ses dérivés (47%)**. **L'efficacité de ce mutant en présence du α -methyl L-rhamnose est exceptionnelle mais n'a pu être déterminée puisque l'enzyme sauvage ne reconnaît pas cet accepteur.**
- Le mutant F290K conserve une activité polymérase en présence de saccharose **mais la perd au détriment de la glucosylation de l' α -allyl N-acétyl-D-glucosamine et de ses dérivés**. **L'efficacité catalytique du mutant F290K en présence de cet accepteur a été améliorée de façon spectaculaire (facteur 130) en comparaison avec la forme sauvage de l'amylosaccharase.**

Des études complémentaires d'analyse structurale (cristallographie des mutants en complexe avec les disaccharides d'intérêt, simulations de dynamique moléculaire ou techniques de RMN de type STD (Saturation Transfert Difference)) pourront être menées afin de **mettre en évidence les modifications conformationnelles induites par l'introduction de ces mutations au niveau du sous-site +1**. Elles permettraient de mieux comprendre l'affinité accrue de ces mono-mutants envers les accepteurs cibles. Ces expériences pourraient également guider la recherche vers

d'autres **cibles de mutagenèse au niveau du sous-site +1** mais aussi **du sous-site +2 de l'ASNP pour éviter la formation de trisaccharide.**

4- Production des disaccharides cibles par les 2 mutants améliorés : I228Y et F290K

Ces deux mutants ont été utilisés, sous forme d'extraits cellulaires, pour produire à plus grande échelle les deux disaccharides cibles et les fournir à l'Unité de Chimie des Biomolécules de l'Institut Pasteur. **L'obtention de plusieurs grammes de produits a ainsi permis de valider l'approche enzymatique.**

Ces disaccharides ont, ensuite, été convertis, par voie chimique, en intermédiaires de synthèse pour poursuivre l'élongation du motif. La voie chimio-enzymatique a permis, non seulement de contourner l'étape chimique de glucosylation mais aussi de substituer les substrats chimiques onéreux (comme le tétra benzyl-D-glucose) par le saccharose, substrat abondant et peu coûteux. Dans le cas du sérotype 3a, elle offre aussi l'avantage de simplifier les étapes de la chimie et leur nombre (de 9 à 7).

Cette phase a, en définitive, permis de valider le concept d'une démarche doublement innovante pour la synthèse d'oligosides complexes, qui a mis à profit les outils d'ingénierie enzymatique, pour concevoir de nouveaux biocatalyseurs et les intégrer, pour la première fois, en début de synthèse. En outre, ce principe peut être appliqué à d'autres systèmes enzymatiques pour la production de cibles d'intérêt thérapeutiques.

Suite à la validation de l'approche, **nos efforts se sont concentrés sur l'optimisation de l'étape enzymatique.** Il s'agissait alors de créer des mutants améliorés en termes de rendement en disaccharide cible et d'efficacité catalytique vis-à-vis de leur donneur, le saccharose. Celle-ci a été fortement altérée par l'introduction d'une simple mutation. Pour cela, nous avons entrepris une exploration plus approfondie du sous-site +1 de l'ASNP qui requérait au préalable une méthode de criblage plus adaptée à des banques de grande taille.

5- Développement d'une nouvelle méthode de criblage

Afin de sélectionner uniquement les mutants actifs de l'ASNP, **nous avons développé une nouvelle méthode de criblage, réalisée sur milieu solide et basée sur l'utilisation d'un indicateur coloré de pH** (le Bleu de BromoThymol, BBT). Cette méthode permet de différencier, en 24h les clones actifs (colorés en jaune ou vert, pH < 6) des clones inactifs (colorés en bleu, pH > 7,5) puisque les amylosaccharases actives libèrent du fructose qui sera métabolisé par les enzymes d'*E. coli* pour libérer de l'acide. Cette méthode a été appliquée avec succès à des

banques de double-mutants permettant de réduire la taille des banques à seulement 8 % de clones actifs. Ce principe pourra, à l'avenir, être extrapolé au criblage d'autres glucane-saccharases, voire d'autres transglycosidases.

Dans le cadre de notre étude, cette méthode présente, cependant, ses limites puisqu'elle ne sélectionne pas les clones faiblement actifs sur saccharose mais qui seront activés en présence d'accepteurs (tels que I228Y ou F290D). L'introduction des accepteurs cibles dans le test colorimétrique au pH est une possibilité à envisager.

6- Exploration du sous site (+1) de l'amylosaccharase par recombinaison raisonnée

Au vu des résultats positifs obtenus sur les 7 positions de la banque de mono-mutants, nous avons décidé de générer une diversité plus importante, focalisée sur ces mêmes positions, et de rechercher des effets synergiques entre les différentes mutations pour améliorer la réaction de glucosylation des accepteurs cibles. Quatre banques de variants (20 000 clones) présentant deux mutations voisines ou distantes ont ainsi été construites par mutagenèse à saturation. **Vingt double-mutants ont été identifiés comme potentiellement améliorés** (5 pour la glucosylation de l' α -methyl L-rhamnose et 15 pour celle de l' α -allyl N-acétyl-D-glucosamine). Ces mutants devront être produits à plus grande échelle pour évaluer plus précisément leur gain.

Pour établir une cartographie fonctionnelle du sous-site (+1) de l'amylosaccharase, la recombinaison de l'ensemble de ces positions apparaît comme la stratégie idéale. L'exploration de la diversité d'une telle banque dont la taille théorique est d'environ 10^8 variants est très élevée. Cela nécessitera le développement de méthodes de criblage à haut-débit adaptées aux accepteurs cibles. En parallèle, il serait évidemment très avantageux de pouvoir disposer **d'outils de criblage virtuel performants nous permettant de réduire l'exploration expérimentale.** Pour le développement de tels types d'outils, il est absolument indispensable de disposer d'un jeu de données assez large pour valider la justesse des prédictions de changement de spécificité. Nous disposons aujourd'hui de ces données et envisageons de les mettre à profit pour développer des outils innovants de criblage virtuel.

ANNEXE

Séquence nucléotidique et protéique de l'amylosaccharase de *Neisseria polysaccharea*

Les éléments de structure secondaire sont colorés en jaune pour les feuillets β et en rouge pour les hélices α . Les domaines N, B, B' et C sont grisés. Les résidus de la triade catalytique sont indiqués en orange. Les 7 positions qui ont été ciblées dans cette étude pour remodeler le sous-site +1 de l'ASNP encadrées en bleu. La séquence est numérotée d'après la séquence déposée dans la Protein Data Bank (numéro d'accès : 1G5A).

| Domaine N | | | | | | | | | | | | | | | | | | |
|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| S | P | N | S | Q | Y | L | K | T | R | I | L | D | I | Y | T | P | E | 18 |
| TCC | CCG | AAT | TCA | CAG | TAC | CTC | AAA | ACA | CGC | ATC | TTG | GAC | ATC | TAC | ACG | CCC | GAA | 54 |
| Q | R | A | G | I | E | K | S | E | D | W | R | Q | F | S | R | R | M | 36 |
| CAG | CGC | GCC | GGC | ATC | GAA | AAA | TCC | GAA | GAC | TGG | CGG | CAG | TTT | TCG | CGC | CGC | ATG | 108 |
| D | T | H | F | P | K | L | M | N | E | L | D | S | V | Y | G | N | N | 54 |
| GAT | ACG | CAT | TTC | CCC | AAA | CTG | ATG | AAC | GAA | CTC | GAC | AGC | GTG | TAC | GGC | AAC | AAC | 162 |
| E | A | L | L | P | M | L | E | M | L | L | A | Q | A | W | Q | S | Y | 72 |
| GAA | GCC | CTG | CTG | CCT | ATG | CTG | GAA | ATG | CTG | CTG | GCG | CAG | GCA | TGG | CAA | AGC | TAT | 216 |
| S | Q | R | N | S | S | L | K | D | I | D | I | A | R | E | N | N | P | 90 |
| TCC | CAA | CGC | AAC | TCA | TCC | TTA | AAA | GAT | ATC | GAT | ATC | GCG | CGC | GAA | AAC | AAC | CCC | 270 |
| β 1 | | | | | | | | | | | | | | | | | | |
| D | W | I | L | S | N | K | Q | V | G | G | V | C | Y | V | D | L | F | 108 |
| GAT | TGG | ATT | TTG | TCC | AAC | AAA | CAA | GTC | GGC | GGC | GTG | TGC | TAC | GTT | GAT | TTG | TTT | 324 |
| α 1 | | | | | | | | | | | | | | | | | | |
| A | G | D | L | K | G | L | K | D | K | I | P | Y | F | Q | E | L | G | 126 |
| GCC | GGC | GAT | TTG | AAG | GGC | TTG | AAA | GAT | AAA | ATT | CCT | TAT | TTT | CAA | GAG | CTT | GGT | 378 |
| β 2 | | | | | | | | | | | | | | | | | | |
| L | T | Y | L | H | L | M | P | L | F | K | C | P | E | G | K | S | D | 144 |
| TTG | ACT | TAT | CTG | CAC | CTG | ATG | CCG | CTG | TTT | AAA | TGC | CCT | GAA | GGC | AAA | AGC | GAC | 432 |
| G | G | Y | A | V | S | S | Y | R | D | V | N | P | A | L | G | T | I | 162 |
| GGC | GGC | TAT | GCG | GTC | AGC | AGC | TAC | CGC | GAT | GTC | AAT | CCG | GCA | CTG | GGC | ACA | ATA | 486 |
| α 2 | | | | | | | | | | | | | | | | | | |
| G | D | L | R | E | V | I | A | A | L | H | E | A | G | I | S | A | V | 180 |
| GGC | GAC | TTG | CGC | GAA | GTC | ATT | GCT | GCG | CTG | CAC | GAA | GCC | GGC | ATT | TCC | GCC | GTC | 540 |
| β 3 | | | | | | | | | | | | | | | | | | |
| V | D | F | I | F | N | H | T | S | N | E | H | E | W | A | Q | R | C | 198 |
| GTC | GAT | TTT | ATC | TTC | AAC | CAC | ACC | TCC | AAC | GAA | CAC | GAA | TGG | GCG | CAA | CGC | TGC | 594 |
| A | A | G | D | P | L | F | D | N | F | Y | Y | I | F | P | D | R | R | 216 |
| GCC | GCC | GGC | GAC | CCG | CTT | TTC | GAC | AAT | TTC | TAC | TAT | ATT | TTC | CCC | GAC | CGC | CGG | 648 |
| M | P | D | Q | Y | D | R | T | L | R | E | I | F | P | D | Q | H | P | 234 |
| ATG | CCC | GAC | CAA | TAC | GAC | CGC | ACC | CTG | CGC | GAA | ATC | TTC | CCC | GAC | CAG | CAC | CCG | 702 |
| G | G | F | S | Q | L | E | D | G | R | W | V | W | T | T | F | N | S | 252 |
| GGC | GGC | TTC | TCG | CAA | CTG | GAA | GAC | GGA | CGC | TGG | GTG | TGG | ACG | ACC | TTC | AAT | TCC | 756 |
| α 3 | | | | | | | | | | | | | | | | | | |
| F | Q | W | D | L | N | Y | S | N | P | W | V | F | R | A | M | A | G | 270 |
| TTC | CAA | TGG | GAC | TTG | AAT | TAC | AGC | AAC | CCG | TGG | GTA | TTC | CGC | GCA | ATG | GCG | GGC | 810 |
| β 4 | | | | | | | | | | | | | | | | | | |
| E | M | L | F | L | A | N | L | G | V | D | I | L | R | M | D | A | V | 288 |
| GAA | ATG | CTG | TTC | CTT | GCC | AAC | TTG | GGC | GTT | GAC | ATC | CTG | CGT | ATG | GAT | GCG | GTT | 864 |
| A | F | I | W | K | Q | M | G | T | S | C | E | N | L | P | Q | A | H | 306 |
| GCC | TTT | ATT | TGG | AAA | CAA | ATG | GGG | ACA | AGC | TGC | GAA | AAC | CTG | CCG | CAG | GCG | CAC | 918 |
| α 4 | | | | | | | | | | | | | | | | | | |
| A | L | I | R | A | F | N | A | V | M | R | I | A | A | P | A | V | F | 324 |
| GCC | CTC | ATC | CGC | GCG | TTC | AAT | GCC | GTT | ATG | CGT | ATT | GCC | GCG | CCC | GCC | GTG | TTC | 972 |
| β 5 | | | | | | | | | | | | | | | | | | |
| F | K | S | E | A | I | V | H | P | D | Q | V | V | Q | Y | I | G | Q | 342 |
| TTC | AAA | TCC | GAA | GCC | ATC | GTC | CAC | CCC | GAC | CAA | GTC | GTC | CAA | TAC | ATC | GGG | CAG | 1026 |

TABLE DES ILLUSTRATIONS

LISTE DES FIGURES

| | |
|--|-----|
| FIGURE 1 Représentation schématique des étapes essentielles pour la mise au point d'un vaccin par la vaccinologie classique et de la vaccinologie inverse. | 31 |
| FIGURE 2 Type de vaccins enregistrés en 2000 et prédiction des types de vaccins disponibles en 2020 | 32 |
| FIGURE 3 Structure de la paroi des bactéries à Gram positif et à Gram négatif. | 33 |
| FIGURE 4 Structure des LPS..... | 34 |
| FIGURE 5 Production d'anticorps stimulés par des antigènes polysaccharidiques conjugués ou non à des protéines porteuses..... | 38 |
| FIGURE 6 Représentation schématique de la stratégie de construction des vaccins glyco-conjugués..... | 39 |
| FIGURE 7 Synthèses chimiques automatisées | 45 |
| FIGURE 8 Structure du premier vaccin commercialisé semi synthétique anti-Hib..... | 47 |
| FIGURE 9 Pandémies de dysenteries de <i>Shiga</i> | 50 |
| FIGURE 10 Distribution des espèces de <i>Shigella</i> provenant d'études menées en Amérique Latine, Asie et Afrique..... | 50 |
| FIGURE 11 Invasion de l'épithélium par les shigelles | 51 |
| FIGURE 12 Antigènes Ag-O de <i>S. flexneri</i> , <i>S. dysenteriae 1</i> et <i>S. sonnei</i> | 54 |
| FIGURE 13 Glycoconjugués semi-synthétiques à potentiel vaccinal contre <i>S. dysenteriae 1</i> | 57 |
| FIGURE 14 Unité répétitive de <i>S. flexneri 2a</i> | 58 |
| FIGURE 15 Représentation schématique de l'organisation du tonneau (β/α) ₈ des GS des familles GH13 et GH70 et des domaines N, B, B' et C avec l'exemple de l'AS de <i>N. polysaccharea</i> (GH13)..... | 68 |
| FIGURE 16 Schéma du mécanisme réactionnel suivi par l'ASNP..... | 69 |
| FIGURE 17 Réactions catalysées par les glucane-saccharases..... | 71 |
| FIGURE 18 Glucosylation d'accepteurs glucidiques catalysée par l'ASNP..... | 73 |
| FIGURE 19 Amylosaccharases hypothétiques identifiées dans les génomes séquencés <i>via</i> le logiciel BLAST (analyse effectuée en novembre 2007)..... | 76 |
| FIGURE 20 Séquence logo des amylosaccharases et saccharose-hydrolases (famille GH13.4)..... | 77 |
| FIGURE 21 Alignement des séquences consensus des GH13.4 et identification des résidus conservés..... | 79 |
| FIGURE 22 Réactions catalysées par l'ASNP à partir de saccharose seul (100 mM) et répartition des produits..... | 80 |
| FIGURE 23 Structure tridimensionnelle de l'amylosaccharase de <i>N. polysaccharea</i> | 82 |
| FIGURE 24 Représentation des sous-sites de liaison de l'ASNP | 83 |
| FIGURE 25 Représentation schématique des interactions (A) entre le mutant inactif E328Q et le saccharose et (B) entre l'ASNP et le D-Glucose..... | 85 |
| FIGURE 26 Interactions du site OB1..... | 85 |
| FIGURE 27 Mécanismes d'action d'une glycoside-hydrolase agissant avec rétention de configuration (cas d'une β -D-glycoside hydrolase) (A), d'une glycosynthase (B), d'une thioglycoligase (C) et d'une thioglycosynthase (D)..... | 92 |
| FIGURE 28 Sélection <i>in vivo</i> par complémentation chimique..... | 95 |
| FIGURE 29 Alignement de séquences conservées des domaines catalytiques de glucane-saccharases de spécificités variées..... | 97 |
| FIGURE 30 Unités répétitives de <i>S. flexneri 3a</i> et <i>1b</i> | 100 |
| FIGURE 31 Accepteurs retenus ([A'] et [D'])..... | 103 |
| FIGURE 32 Réactions catalyzed by glucansucrases..... | 130 |
| FIGURE 33 Analysis of D-GlcPNAc glucosylation by 4 different GH70 glucansucrases..... | 132 |
| FIGURE 34 Analysis of D-GlcPNAc glucosylation by GH13 glucansucrase, | 132 |

| | |
|--|-----|
| FIGURE 35 Structure of disaccharide P1 obtained by the $\Delta 4N$ SEV663YDA-..... | 133 |
| FIGURE 36 Structure of disaccharide P2 obtained by AS-mediated glucosylation of D-Glc ρ NAc..... | 134 |
| FIGURE 37 Structures of disaccharide P3 and trisaccharide P4..... | 137 |
| FIGURE 38 Repeating unit of <i>S. flexneri</i> serotypes 1a and 1b O-Ags..... | 154 |
| FIGURE 39 Repeating unit of <i>S. flexneri</i> serotypes 2b, 3a, 5a, 5b and X O-Antigens..... | 154 |
| FIGURE 40 Reaction catalyzed by glucansucrases..... | 155 |
| FIGURE 41 View of amylosucrase active site in complex with maltoheptaose (PDB: 1MW0)..... | 158 |
| FIGURE 42 Comparison of docking modes. | 159 |
| FIGURE 43 Screening of the library for their ability to synthesize the desired disaccharide: | 161 |
| FIGURE 44 Transglucosylation of (a) α -L-RhapOMe (A') and (b) α -D-Glc ρ NAcOallyl (D) using the most improved variants I228Y and F290K..... | 162 |
| FIGURE 45 Glucosylation Rates of α -D-Glc ρ NAc-Oallyl at different times | 174 |
| FIGURE 46 Glucosylation Rates obtained for L-Rhap, α -L-Rhap-OMe and α -L-Rhap-Oall..... | 175 |
| FIGURE 47 Colorimetric tests for detection of amylosucrase activity | 193 |
| FIGURE 48 Screening of amylosucrase active clones | 194 |
| FIGURE 49 Improved single-mutants obtained from previous study ¹⁵ compared to wtAS..... | 194 |
| FIGURE 50 Strategy adopted for the construction of the four double-mutant libraries | 195 |

LISTE DES SUPPLEMENTARY FIGURES

| | |
|--|-----|
| SUPPLEMENTARY FIGURE 1 Schematic diagramm of protein-ligand interactions (LigPlot) | 179 |
| SUPPLEMENTARY FIGURE 2 Comparison of Dionex HPAEC product profiles obtained at the end of the reaction ($t = 24$ h) with wild-type AS (magenta) and the variant F290K (blue) using (A) 146 mM sucrose and (B) using 146 mM sucrose supplemented with 146 mM acceptor (α -D-Glc ρ NAc-Oallyl)..... | 180 |
| SUPPLEMENTARY FIGURE 3 Comparison of Dionex HPAEC product profiles obtained at the end of the reaction ($t = 24$ h) with wild-type AS (magenta) and the variant I228Y (blue) using (A) 146 mM sucrose and (B) using 146 mM sucrose supplemented with 146 mM acceptor (α -L-RhapOMe)..... | 180 |
| SUPPLEMENTARY FIGURE 4 Comparison of kinetic parameters for the acceptor glucosylation catalysed by the variant F290K and wtAS..... | 180 |
| SUPPLEMENTARY FIGURE 5 Comparison of kinetic parameters for the acceptor glucosylation catalysed by the variant I228Y and wtAS..... | 180 |
| SUPPLEMENTARY FIGURE 6 Overall structure of <i>N. polysaccharea</i> amylosucrase..... | 181 |

LISTE DES TABLEAUX

| | |
|--|-----|
| TABLEAU 1 Exemples d'unités répétitives de PC et O-PS des bactéries à Gram positif et à Gram négatif..... | 35 |
| TABLEAU 2 Caractéristiques des vaccins polyosidiques conjugués commercialisés contre les méningites à <i>Haemophilus influenzae</i> de sérotype b (Hib). | 42 |
| TABLEAU 3 Espèces et sérotypes de <i>Shigella</i> (D'après Levine et coll. ⁷⁵)..... | 49 |
| TABLEAU 4.1 Glucane-saccharases de la famille GH13 clonées et caractérisées..... | 65 |
| TABLEAU 4.2 Glucane-saccharases de la famille GH70 clonées et caractérisées..... | 54 |
| TABLEAU 5 Valeurs des constantes cinétiques apparentes des vitesses initiales de consommation du saccharose (ViS), d'hydrolyse du saccharose (ViG) et de polymérisation (ViGn)..... | 81 |
| TABLEAU 6 Glycosynthases, thioglycoligases et thioglycosynthases répertoriées (décembre 2007). | 93 |
| TABLEAU 7 Comparaison des principales caractéristiques des 4GS..... | 102 |

LISTE DES TABLES

| | |
|---|-----|
| TABLE 1 Origin and specificity of the recombinant glucansucrases selected for the glucosylation of the target acceptors: D-GlcpNAc, L-Rhap and α -L-RhapOMe. | 131 |
| TABLE 2 Transglucosylation of D-GlcpNAc using different recombinant glucansucrases | 131 |
| TABLE 3 ¹ H and ¹³ C chemical shifts (ppm, D ₂ O) and ³ J coupling constants (Hz) of P1 | 133 |
| TABLE 4 ¹ H and ¹³ C chemical shifts (ppm, D ₂ O) and ³ J coupling constants (Hz) of P2..... | 134 |
| TABLE 5 Transglucosylation of L-rhamnose using different recombinant glucansucrases at Acceptor:Sucrose ratio of 1:1 (146 mM). | 135 |
| TABLE 6 Effect of the Acceptor/Sucrose molar ratio (A/S) at sucrose concentrations of 146 and 292 mM. Percentages of Glcp and Rhap transferred were determined at final time (24h). At final time, sucrose is fully consumed | 136 |
| TABLE 7 ¹ H and ¹³ C chemical shifts (ppm, D ₂ O) and ³ J coupling constants | 137 |
| TABLE 8 Transglucosylation of α -L-RhapOMe using different recombinant glucansucrases | 138 |
| TABLE 9 Examples of repeating units and/or cores of bacterial surface polysaccharides containing the disaccharide motives synthesized by glucansucrases as demonstrated in this work..... | 139 |
| TABLE 10 Determination of kinetic parameters of the improved variants I228Y and F290K compared to wtAS. | 163 |
| TABLE 11 Primers used to generate the 19 monomutants for the positions 228, 289, 290, 330, 331, 394 and 446..... | 170 |
| TABLE 12 Sequence of XXX codon used to replace each selected amino acids by the 19 other ones | 170 |
| TABLE 13 Glucosylation Rates of α -D-GlcpNAcOMe, β -D-GlcpNAcOMe, β -D-GlcpNAcOallyl and D-GlcNHTCA..... | 176 |
| TABLE 14 Number of clones and active clones obtained for the four librairies | 196 |
| TABLE 15 Improved mutants for EA' synthesis | 197 |
| TABLE 16 Improved mutants for ED synthesis (Libraries 1 and 4) | 198 |
| TABLE 17 Improved and not sequenced mutants for ED synthesis (Library 2)..... | 198 |
| TABLE 18 Sucrose activity of the most improved mono-mutants for EA' synthesis | 200 |
| TABLE 19 Degenerate primers used for the construction of librairies (I228-F229, A289-F290, I330-V331) | 201 |