(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 September 2009 (17.09.2009)

(51) International Patent Classification:
C12N 9/10 (2006.01)

(21) International Application Number:
PCT/IB2009/005325

(22) International Filing Date:
12 March 2009 (12.03.2009)

(25) Filing Language:
English

(26) Publication Language:
English

(30) Priority Data:
08290237.0 12 March 2008 (12.03.2008) EP

(71) Applicants (for all designated States except US):
INSTI-
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE [FR/FR]; 3 rue Michel Ange, F-75016 Paris (FR).
INSTITUT NATIONAL DES SCIENCES APPLIQUES DE TOULOUSE [FR/FR]; 135 avenue de Raguell, F-31077 Toulouse Cedex 4 (FR).

(72) Inventors; and
MULARD, Laure
rence [FR/FR]; 11 rue de la Convention, F-94270 Le Kremlin Bicetre (FR).
ANDRE, Isabelle [FR/FR]; 3 rue de l’Aqueduc, F-31400 Toulouse (FR).
CHAMPION, Elise [FR/FR]; 2 rue de Reims, F-31400 Toulouse (FR).
MOULIS, Claire [FR/FR]; Lieu-dit Emprat, F-31380 Garidech (FR).
MOREL, Sandrine [FR/FR]; Chateau d’Auzeville, F-31320 Auzeville (FR).
MONSAN, Pierre [FR/FR]; 22 Chemin Gravette, F-31700 Mondonville (FR).
REMAUD-SIMEON, Magali [FR/FR]; 1 rue Ben-
jamin Charrier, F-31520 Ramonville (FR).
DESCROIX, Karine [FR/FR]; 117 rue Paul Vaillant Couturier, F-92240 Malakoff (FR).

(74) Agents: ORÈS, Beatrice et al; Cabinet Ores, 36 rue de Saint Petersburg, F-75008 Paris (FR).


(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UA, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report (Rule 48.2(g))
— with sequence listing part of description (Rule 5.2(a))

(54) Title: MUTANTS OF GLYCOSIDE HYDROLASES AND USES THEREOF FOR SYNTHESIZING COMPLEX OLIGOSACCHARIDES AND DISACCHARIDE INTERMEDIATES

(57) Abstract: Method for preparing the disaccharide α-D-glucopyranosyl-(1→4)-2-iV-acetyl-2-deoxy-α-D-glucopyranoside, comprising the step of using a mutant of a wild type glycoside hydrolase.
MUTANTS OF GLYCOSIDE HYDROLASES AND USES THEREOF FOR SYNTHESIZING COMPLEX OLIGOSACCHARIDES AND DISACCHARIDE INTERMEDIATES

The present invention relates to mutants of glycoside hydrolases and uses thereof in chemo-enzymatic synthesis of complex oligosaccharides, in particular fragments of *S. flexneri* Ia and Ib O-antigen.

Carbohydrates displayed at the surface of cells and pathogens are involved in a wide range of biological processes, among which several intercellular recognition events, as well as host-pathogen interactions possibly resulting in microbial or viral infections. The understanding of the molecular events involved in carbohydrate-mediated interactions has long been impaired by the difficult access to relevant oligosaccharides and glycoconjugates in pure form and sufficient amounts. The exquisite diversity of possible structures, varying in monosaccharide composition, linkage and branching pattern (ref. 1), is indeed a major roadblock to easy availability. In recent years however, important developments in the preparation of carbohydrate derivatives, based on (i) multistep chemical synthesis, (ii) enzymatic strategies using recombinant glycosyltransferases (i.e., enzymes catalyzing the transfer of a monosaccharide residue from an activated sugar phosphate to an acceptor molecule; EC 2.4) with *in situ* regeneration of sugar nucleotides, (iii) combinations thereof, or (iv) biosynthesis using metabolically engineered cell-factories (ref. 2), have opened the way to significant progress in the fields of glycobiology and glycotherapeutics (ref. 3-5). A number of efficient and elegant synthetic methods such as one-pot oligosaccharide synthesis (ref. 6) or automated synthetic tools (ref. 7) have been developed to provide more straightforward access to structurally-defined carbohydrates. The use of lightly protected precursors and the regioselective one-pot protection of monosaccharides were recently emphasized (ref. 8).

Nevertheless, despite accomplished advancements, chemical approaches towards specific usable microbial oligosaccharides still need considerable effort. They remain, for the most part, highly dependent on the design of appropriate combinations of multiple protection, deprotection, and efficient glycosylation steps, which often involve numerous tedious chromatographic separations (ref. 9). Avoiding the need for protecting groups, organisms engineered to express several glycosyltransferase genes have been used to produce a nice range of biologically active complex
carbohydrates, but they remain to date limited to the synthesis of short oligosaccharides which can passively cross the cell membrane (ref. 2, 10).

Following the early success of polysaccharide vaccines in the second half of the 20th century, polysaccharide-protein conjugate vaccines were seen as a major progress in antibacterial vaccination (ref. 11, 12). Indeed, made from bacterial polysaccharides purified from pathogen cell cultures, eventually shortened following partial-chemical hydrolysis or enzymatic depolymerisation of the native antigen, and subsequently covalently coupled to a protein carrier, these second generation carbohydrate vaccines are suitable for use in human (ref. 12). Potential extrapolations are numerous since for a large number of pathogens, surface carbohydrates behave as key "protective antigens". Interestingly, this long known property of a range of bacterial polysaccharides was extended in recent years to other microbial carbohydrates of fungal (ref. 13) and parasitic origin (ref. 14). Besides, the disclosure that cancer cells could among other features be differentiated from healthy ones by the presence of surface glycoconjugates, often termed tumor-associated carbohydrate antigens, contributed to additional interest in carbohydrate antigens (ref. 15). Overall, interest in synthetic carbohydrate-based vaccines has emerged as one amongst the many exploding fields of carbohydrate medical applications (ref. 15, 16). The development of synthetic microbial carbohydrate-protein conjugates, thus termed third generation carbohydrate vaccines, was proposed as an alternative to conventional polysaccharide vaccines, compatible with the increasingly demanding requirements in terms of safety, efficiency, product definition, and needs for multivalent vaccines (ref. 17). Most interestingly, use of the natural antigen, and consequently risks associated to materials of biological origin, are avoided. However, chemical synthesis of carbohydrate-protein conjugates, more precisely of appropriate carbohydrate haptens can be seen as a drawback. By way of example, Figure 2B shows the first steps of a known chemical synthetic route to *S. flexneri* Ia O-antigen (ref. 66). In order to control the 1,2-cis glycosidic linkage involved in the key disaccharide motif α-D-glucopyranosyl-1(→4)-β-D-2-/V-acetyl-2-deoxyglucopyranosyl, the synthesis relies on the recently developed method of intramolecular glycosylation through prearranged unsymmetrically tethered glycosides. This attractive strategy involves a thioethyl glucopyranoside bearing an hydroxyl group at position 2 differentiated from the others in order to allow introduction of the tether, next linked at position 3 of the glucosamine acceptor. Subsequent deblocking at position 4 of the
glucosamine residue provided the donor/acceptor bis(glycoside). Intramolecular glycosylation gave the (1→4)-α-D-glucosidic linkage only, providing a suitable disaccharide donor in 45% yield over two steps. Next, transesterification at position 3 of the glucosamine residue provided a disaccharide acceptor in 34% yield over three steps.

As expected, the glycosylation step was both highly stereoselective and high yielding (72%, α-anomer only). A limitation of the strategy, however, is the 8 step-synthesis of the required tethered donor from glucose, added to the 5 steps synthesis of the acceptor. More recently, an alternative was proposed on a model disaccharide (K. Descroix & L. Mulard, unpublished). In this case (Figure 2E), the construction of the α-D-glucopyranosyl-(1→2)-α-D-2-iV-acetyl-2-deoxy-glucopyranosyl motif involves a more conventional glycosylation step between the phenyl tetra-O-benzyl-l-thio-β-D-glucopyranoside donor (XX14) and allyl 2-acetamido-3-O-acetyl-6-0-benzyl-2-deoxy-α-D-glucopyranoside (XX15) as acceptor. Interestingly, the disaccharide acceptor XX15, bearing the required 1,2-cis stereochemistry was obtained in 69% yield over two steps indicating a good stereoselectivity of the glycosylation step despite the absence of any participating group at position 2 of the donor. Nevertheless, preparation of disaccharide XX25 from the free monosaccharide precursors required a total of 10 synthetic steps combined to 3 purifications.

The use of enzymes as catalysts has then emerged as a practical alternative to a number of limitations encountered in chemical synthesis (ref. 18, 19). Leloir-type glycosyltransferases and transglycosidases constitute the two major classes of enzymes that can be used for the synthesis of glycosidic linkages. Both are enzymes transferring a glycosyl group from a donor to an acceptor. Glycosyltransferases require nucleotide sugar as donor substrate whereas transglycosidases usually employ mono- and/or oligosaccharides as donor substrates.

The term "donor" refers to a molecule that provides a glycosyl moiety which will be transferred to an acceptor molecule.

The term "acceptor" refers to a molecule that will receive the glycosyl moiety through the formation of a chemical bond, preferentially C-O-linkage.

However, despite the increasing number of available transglycosidases and glycosyltransferases (Leloir type), the lack of appropriate enzymatic tools with requisite substrate specificity has prevented extensive exploration of the chemo-
enzymatic strategies when dealing with complex bacterial carbohydrate antigens. The use of multiple overexpressed native glycosyltransferases was shown to be highly rewarding for the synthesis of the upstream pentasaccharide terminus of the Neisseria meningitidis lipo-oligosaccharide (ref. 20), but examples of enzymatic synthesis of complex carbohydrate remain scarce (ref. 21). Indeed, certain membrane Leloir-type glycosyltransferases are not easily available. Their nucleotide activated sugar substrates are expensive. They may be generated in situ, but the process necessitates additional enzymes (ref. 21).

Replacement of glycosyltransferases by transglycosidases has been proposed to proceed from different types of glycosyl donors, and to be compatible with a larger variety of acceptors (ref. 22). Interestingly, modified donors were occasionally used successfully (ref. 23). Nonetheless, the availability of these enzymes is often critical, especially when considering the appropriate regio- and stereospecificity required for a given target (ref. 22, 24 and 25).

Protein engineering based on rational, semi-rational or fully combinatorial approaches (directed evolution) has also proven to be extremely useful to generate catalysts with improved natural properties but also to create new substrate specificities (ref. 26, 27). In the field of carbohydrate-enzymes, glycosyltransferase substrate specificity has been successfully modified by site-directed mutagenesis assisted by computational modelling or directed evolution for the synthesis of biologically relevant carbohydrate structure (ref. 27). Promiscuous β-glycosidases showing altered and new specificities towards the donor or the acceptor sugar have been generated (ref. 28, 29, 31). Engineering of new glycosynthases (i.e., enzymes catalyzing the condensation of sugar residues for synthesizing a glycoside) from β-glycosidases (classified into EC 3.2.1) also emerged as a powerful way to generate modified transferases, even if they use only fluoride donors (ref. 32). In addition, this methodology has never been shown to be successful for α-retaining enzyme. However, one case of active glycosynthase derived from invertase enzymes has been described up to now (ref. 33).

Interestingly, the use of intermediates issued from enzymatic glycosylation in the subsequent generation of glycoconjugates of higher complexity has also been reported (ref. 36-38). In all cases, the building blocks, conceived by action of
native glycosyltransferases, were converted to donors and used as such, following peracylation.

The term "building block" refers to a suitably protected carbohydrate intermediate occurring in the chemical pathway of synthesis of complex oligosaccharides, e.g., said carbohydrate can be a disaccharide.

The term "intermediate" refers to a compound, protected or not, issued from an enzymatic and/or synthetic step, and involved in the multi-step synthesis of a specific target, e.g., said compound can be a disaccharide.

The 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.

The Inventors have thus investigated the applicability of enzymatic glycosylation for the synthesis of building blocks compatible with chemical chain extension both at the reducing and non-reducing ends, which is compatible with subsequent conversion into donors as well as into acceptors. In the course of their investigation, the Inventors have surprisingly demonstrated the applicability of engineered amylosucrases (AS) - which are glycoside hydrolases - to the synthesis of disaccharide intermediates to synthetic oligosaccharide fragments of *Shigellaflexneri* Ia and/or Ib lipopolysaccharide by *in vitro* chemo-enzymatic synthetic methodologies, implicating an enzymatic step at an early stage in the synthesis.

Amylosucrases (EC 2.4.1.4) as well as sucrose hydrolases (EC 3.2.1.-) belong to the family 13 of the glycoside hydrolases (GHI 3), and more particularly the subfamily 4 (GHI 3.4) as defined per the CAZY nomenclature (ref. 62-65).

Amylosucrases and sucrose hydrolases operate on the same substrate (sucrose) with the same molecular mechanism (ref. 67). The difference between the amylosucrases and the sucrose hydrolases resides mostly in their transglycosylation abilities (ref. 65).

The structure of amylosucrase from *Neisseriapolyasacchararea* is the only known structure of enzymes from family GHI 3.4 (ref. 68). The single polypeptide chain (628 amino acid residues) of amylosucrase from *Neisseriapolyasacchararea* is folded into a tertiary structure consisting of five domains named N (residues 1-90), A (residues 98-184; 261-395; 461-550), B (residues 185-260), B’ (residues 395-460) and C (residues 555-628). Domains A, B and C are common domains found in family GH13. Domains N
and B’ are specific to family GHI 3.4. Domain N is the N-terminal domain composed of 6 \( \alpha \)-helices. Domain A is made up of eight alternating \( \beta \)-sheets (\( \beta_1 - \beta_8 \)) and \( \alpha \)-helices (\( \alpha_1-\alpha_8 \)) building up the catalytic core: the \( (\beta_8\alpha_8) \) barrel common to family GHI 3. It contains also eight loops connecting helices to strands (labeled loop 1 to loop 8). Domain B, or loop 3, is an extension of domain A, containing two short antiparallel \( \beta \)-sheets flanked by two \( \alpha \)-helices. Domain B’, or loop 7, is another extension of domain A, composed of two \( \alpha \)-helices followed by a \( \beta \)-sheet and another short \( \alpha \)-helice. Domain C is an eight-stranded \( \beta \)-sandwich found C-terminal to the \( (\beta_8\alpha_8) \) barrel.

Interestingly, at least 4 different patterns of \( \alpha \)-D-glucosylation, have been characterized.

Unexpectedly, the Inventors have now found eleven consensus amino acid sequences to characterize glycoside hydrolases: eight consensus motifs defined hereafter (SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 6; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 10; SEQ ID NO: 11) are localised in said \( \beta \)-sheets (6) or said loops (2) constituting domain A, two consensus motifs (SEQ ID NO: 4; SEQ ID NO: 5) are found in said domain B and one consensus motif (SEQ ID NO: 9) is found in said domain B’.

**Shigella** is the causal agent of shigellosis, or bacillary dysentery. In developing countries, it induces about 1 million deaths per year, most of which involve children under five years of age (ref. 39). In countries where disease is endemic, a number of *S. flexneri* serotypes and to a lesser extent *S. sonnei* are isolated, emphasizing the need for a multivalent vaccine. Noteworthy, despite numerous clinical trials (ref. 40), no vaccine is available so far. Epidemiological as well as experimental data point to the polysaccharide part, or O-antigen (O-Ag), of the bacterial lipopolysaccharide as an important virulence factor (ref. 41) and the major target of protective humoral response against reinfection (ref. 42). *S. flexneri* is divided into at least 14 serotypes based on known O-Ag structures. Interestingly, protein-conjugates of short synthetic oligosaccharides mimicking *S. flexneri* 2a O-Ag induced in mice a potent anti-0-Ag humoral immune response, which was shown to be protective against homologous challenge (ref. 43). The diversity, associated to a close resemblance in composition, of the known *S. flexneri* O-Ag repeating units was found of utmost interest to challenge the investigation. Indeed, except for serotype 6, all *S. flexneri* O-Ag repeating units share a linear tetrasaccharide backbone (ref. 41). Diversity resides in the branching pattern, which involves \( O \)-acetyl and/or \( \alpha \)-D-glucopyranosyl decorations (ref. 41, 44). Interestingly, at least 4 different patterns of \( \alpha \)-D-glucosylation, have been characterized.
for this family of bacterial polysaccharides. N-acetyl-D-glucopyranosamine residue can be implicated as branching acceptor. By way of example, serotypes Ia and Ib of *S. flexneri* share the α-D-glucopyranosyl-(1→4)-N-acetyl-β-D-glucopyranosaminyll (ED) branching pattern.

Within the framework of research that has led to the present invention, the Inventors have demonstrated the chemo-enzymatic synthesis of disaccharide building blocks to *S. flexneri* Ia and Ib serotype-specific oligosaccharides by selecting a 2-acetamido-2-deoxy-D-glucopyranoside residue as substrate acceptor, and using as enzyme a recombinant amylosucrase (an α-retaining transglucosidase from family 13 of glycoside-hydrolases that uses sucrose as glucosyl donor, (ref. 45, 46)), at an earlier stage of a multi-step synthesis. New amylosucrase specificities were then surprisingly generated to glucosylate efficiently and regiospecifically allyl 2-acetamido-2-deoxy- α-D-glucopyranoside to provide building blocks compatible with chemical chain elongation as exemplified (Figures 2A, 2C, and 2D, Examples 2, 3, and 4).

Here-under are examples of repeating units and/or cores of bacterial surface polysaccharides containing the disaccharide motives synthesized by glucansucrases (ref. 39, 41):

<table>
<thead>
<tr>
<th>Target disaccharide</th>
<th>Organism</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-D-GlcP-(1→4)-D-GlcPNac</td>
<td><em>E. coli</em> O18B1</td>
<td>→6)-α-D-GlcP-(1→4)-[β-D-GlcP-(1→3)]-α-D-GalP(1→3)-[α-D-GlcP-(1→4)]-α-D-GlcPNac-(1→2)-α-L-Rhap-(1→)</td>
</tr>
<tr>
<td></td>
<td><em>H. pantelleriensis</em></td>
<td>→6)-α-D-GlcP-(1→4)-α-D-GlcPNac-(1→) as core fragment</td>
</tr>
</tbody>
</table>

Accordingly, the present invention provides a method for preparing a building block corresponding to the disaccharide α-D-glucopyranosyl-(l→4)-2-N-acetyl-2-deoxy- α-D-glucopyranoside of formula (I):

(I)
said method being characterized in that it comprises the step of using a
mutant of a wild-type glycoside hydrolase, wherein said wild type glycoside hydrolase has 450 to 850 amino acids, preferably 580 to 735 amino acids, and comprises, preferably from the N- to C-terminus, eleven motifs defined by the following consensus motifs:

1. the amino acid sequence LGVNYLHLMP (SEQ ID NO: 1), which is located in the β-strand 2 of said wild type glycoside hydrolase;
2. the amino acid sequence DGGYAV (SEQ ID NO: 2), which is located in the loop 2 of the βααs-barrel of said wild type glycoside hydrolase;
3. the amino acid sequence DFVFNH (SEQ ID NO: 3) which is located in the β-strand 3 of said wild type glycoside hydrolase;
4. the amino acid sequence LREIFPDTAPGNF (SEQ ID NO: 4), which is located in the domain B of said wild type glycoside hydrolase;
5. the amino acid sequence FNSYQWDLN (SEQ ID NO: 5), which is located in the C-terminal part of the domain B of said wild type glycoside hydrolase;
6. the amino acid sequence ILRLDAVAFLWK (SEQ ID NO: 6), which is located in the β-strand 4 of said wild type glycoside hydrolase;
7. the amino acid sequence EAIV (SEQ ID NO: 7), which is located in the β-strand 5 of said wild type glycoside hydrolase;
8. the amino acid sequence YVRCHDDI (SEQ ID NO: 8), which is located in the β-strand 7 of said wild type glycoside hydrolase;
9. the amino acid sequence RISGTLASLAG (SEQ ID NO: 9), which is located in the domain B’ of said wild type glycoside hydrolase;
10. the amino acid sequence GIPLIYLGDE (SEQ ID NO: 10), which is located in the β-strand 8 of said wild type glycoside hydrolase;
the amino acid sequence RWVHRP (SEQ ID NO: 11), which is located in the loop 8 of the (β/α)₈-barrel, and the sequence formed by said eleven motifs joined end-to-end from motif (1) to motif (11) of said wild type glycoside hydrolase has at least 65%, preferably at least 70%, and by order of increasing preference, at least 75%, 80%, 85%, 90%, 95%, 95%, 97%, 98%, and 99%, or 100% sequence identity or at least 80%, preferably at least 85%, and by order of increasing preference, at least 90%, 95%, 95%, 97%, 98%, and 99%, or 100% sequence similarity with the amino acid sequence SEQ ID NO: 12, which is formed by the concatenation of the eleven consensus motifs ordered from (1) to (11); wherein said mutant has one or two mutation(s) consisting of, when said mutant has only one mutation:

- the substitution of the amino acid residue at position 4 in said motif (4) with any amino acid selected from the group consisting of alanine (A), cysteine (C), glutamic acid (E), glycine (G), histidine (H), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), valine (V) with the proviso that said wild type glycoside hydrolase does not contain a valine at this position, tryptophan (W) and tyrosine (Y), or

- the substitution of the amino acid residue at position 5 in said motif (4) with any amino acid selected from the group consisting of leucine (L), methionine (M) and valine (V), or

- the substitution of the amino acid residue at position 8 in said motif (6) with any amino acid selected from the group consisting of glutamic acid (E), phenylalanine (F), glycine (G), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R) and valine (V), or

- the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V) and tryptophan (W), or

- the substitution of the amino acid residue at position 4 in said motif (7) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), serine (S), threonine (T) and tyrosine (Y), or
- the substitution of the amino acid residue at position 7 in said motif (8) with any amino acid selected from the group consisting of alanine (A) and valine (V), or

- the substitution of the amino acid residue at position 1 in said motif (9) with any amino acid selected from the group consisting of alanine (A), cysteine (C), phenylalanine (F), glycine (G) with the proviso that said wild type glycoside hydrolase does not contain a glycine at this position, lysine (K), asparagine (N), glutamine (Q), serine (S) with the proviso that said wild type glycoside hydrolase does not contain a serine at this position, threonine (T) and tryptophan (W), or

when said mutant has two mutations:

- the substitution of the amino acid residue at position 4 in said motif (4) with an alanine (A) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a cysteine (C) and the substitution of the amino acid residue at position 5 in said motif (4) with a leucine (L), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a lysine (K) and the substitution of the amino acid residue at position 5 in said motif (6) with any amino acid selected from the group consisting of leucine (L) and tryptophan (W), or

- the substitution of the amino acid residues at positions 4 and 5 in said motif (4) respectively with a leucine (L), or

- the substitution of the amino acid residues at positions 4 and 5 in said motif (4) respectively with a methionine (M), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with a cysteine (C), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of histidine (H) and lysine (K), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a valine (V) and the substitution of the amino acid residue at position 5 in said
motif (4) with any amino acid selected from the group consisting of leucine (L) and methionine (M), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a valine (V) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of histidine (H), lysine (K), arginine (R), and valine (V), or
- the substitution of the amino acid residue at position 8 in said motif (6) with a histidine (H) and the substitution of the amino acid residue at position 9 in said motif (6) with a serine (S), or
- the substitution of the amino acid residue at position 8 in said motif (6) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of cysteine (C), isoleucine (I) and leucine (L), these mutants being preferred, or
- the substitution of the amino acid residue at position 8 in said motif (6) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H).

According to a preferred embodiment of said mutant of a wild type glycoside hydrolase having only one mutation, the amino acid residue at position 9 in said motif (6) is substituted with any amino acid selected from the group consisting of cysteine (C), aspartic acid (D), isoleucine (I), lysine (K) and glutamine (Q), and more preferably with any amino acid selected from the group consisting of aspartic acid (D) and lysine (K).

A "wild type glycoside hydrolase" refers to an amylosucrase (EC 2.4.1.4) or a sucrose hydrolase (EC 3.2.1.-), preferably an amylosucrase. A wild type glycoside hydrolase belongs to the family 13, subfamily 4, of the glycoside hydrolases (GH13.4) as defined per the CAZY nomenclature (ref. 66-69, http://www.cazy.org).

The eleven consensus motifs of the wild type glycoside hydrolases have been found by the Inventors by aligning 34 wild type glycoside hydrolases as shown in Figures 8 and 9.

By way of example, the glycoside hydrolase 1G5A (gi|16974797, SEQ ID NO: 13) comprises, from the N- to C-terminus, the eleven following motifs: (1) 125-134, (2) 144-149, (3) 182-187, (4) 225-237, (5) 250-258, (6) 282-293, (7) 328-331, (8) 388-395, (9) 446-456, (10) 480-489 and (11) 509-514 of SEQ ID NO: 13. These
eleven motifs joined end-to-end form motif (1) to motif (11) form an amino acid sequence which has 83% sequence identity and 92% sequence similarity with the sequence SEQ ID NO: 12.

In order to identify the eleven motifs from a wild type glycoside hydrolase, one of skilled in the art can align the amino acid sequence of this wild type glycoside hydrolase against the amino acid sequence of the glycoside hydrolase 1G5A (SEQ ID NO: 13) for example, and therefore identify the eleven motifs thereof matching the eleven consensus motifs as described above.

Unless otherwise specified, sequence alignments are performed using the well-known MUSCLE program under default parameters (http://phylogenomics.berkeley.edu/cgi-bin/muscle/input_muscle.py). Jalview software can be used for visualizing the alignment and generating the eleven motifs joined end-to-end. The sequence identity and similarity values provided herein are calculated using the Vector NTI AlignX program (V9.1.0, Invitrogen, USA) on a comparison window including the whole set of eleven consensus motifs ordered from (1) to (11) as defined above.

In a particular embodiment of said wild type glycoside hydrolase, it is an amylosucrase selected from the group consisting of the proteins available in the GENBANK database under the following accession number: gi|16974797 (named 1G5A and also reproduced herein as SEQ ID NO: 13), gi|99031739 (named 1ZS2), gi|27574003 (IMVY), gi|27574004 (named IMWO), gi|47169012 (named 1S46), gi|16974938 (named UGI), gi|27574006 (named 1MW2), gi|27574007 (named 1MW3), gi|27574005 (named IMWI), gi|16974937 (named 1JG9), gi|27728142 (named Q84HD6), gi|16670577, gi|32473567, gi|149179129, gi|76260974, gi|77163753, gi|158336602, gi|87310603, gi|149187214, gi|19944090, gi|1099001 19, gi|88795755, gi|152994364, gi|88800970, gi|14778050, gi|17926788, gi|78486138, gi|87300744, gi|91776960, gi|8880471 1, gi|158438431, gi|153811783, gi|153810451, gi|15805957, gi|94984679, gi|19715503, gi|13941581, gi|16125387, gi|19477809 and gi|89092061.

In a more preferred embodiment of said wild type glycoside hydrolase, it is an amylosucrase from Neisseria polysaccharea, and is preferably selected from the group consisting of 1G5A, 1ZS2, IMVY, IMWO, 1S46, 1JGI, 1MW2, 1MW3, IMWI and 1JG9 proteins.
In another preferred embodiment of said wild type glycoside hydrolase, it is a sucrose hydrolase from *Xanthomonas*, and is preferably selected from the group consisting of the proteins available in the GENBANK database under the following accession number: gi|78049174, gi|21244215, gi|58580721, gi|84622653, gi|21232788.

The Table I below shows the sequence identity and similarity percent of the eleven motifs joined end-to-end for each of 34 glycoside hydrolases as described above with the sequence SEQ ID NO: 12.
According to a preferred embodiment of said wild type glycoside hydrolase, it contains an isoleucine (I) or valine (V) residue at position 4 in said motif (4), preferably an isoleucine.

<table>
<thead>
<tr>
<th>GI number in the GenBank Database</th>
<th>Organisms</th>
<th>% identity</th>
<th>% similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>16974797</td>
<td>Neisseria polysaccharea</td>
<td>83</td>
</tr>
<tr>
<td>gi</td>
<td>27728142</td>
<td>Neisseria meningitidis</td>
<td>84</td>
</tr>
<tr>
<td>gi</td>
<td>116670577</td>
<td>Arthrobacter sp. FB24</td>
<td>83</td>
</tr>
<tr>
<td>gi</td>
<td>76260974</td>
<td>Chloroflexus aurantiacus J-10-fl</td>
<td>85</td>
</tr>
<tr>
<td>gi</td>
<td>32473567</td>
<td>Rhodopirellula baltica SH 1</td>
<td>86</td>
</tr>
<tr>
<td>gi</td>
<td>77163753</td>
<td>Nitrosococcus oceani ATCC 19707</td>
<td>85</td>
</tr>
<tr>
<td>gi</td>
<td>149179129</td>
<td>Planctomyces maris DSM 8797</td>
<td>85</td>
</tr>
<tr>
<td>gi</td>
<td>158336602</td>
<td>Acaryochloris marina MBIC11017</td>
<td>81</td>
</tr>
<tr>
<td>gi</td>
<td>109900119</td>
<td>Pseudoalteromonas atlantica T6c</td>
<td>81</td>
</tr>
<tr>
<td>gi</td>
<td>19944090</td>
<td>Psychromonas ingrahamii 37</td>
<td>81</td>
</tr>
<tr>
<td>gi</td>
<td>88795755</td>
<td>Alteromonas macleodii 'Deep ecotype'</td>
<td>79</td>
</tr>
<tr>
<td>gi</td>
<td>149187214</td>
<td>Vibrio shilonii AK1</td>
<td>79</td>
</tr>
<tr>
<td>gi</td>
<td>152994364</td>
<td>Marinomonas sp. MWTL1</td>
<td>79</td>
</tr>
<tr>
<td>gi</td>
<td>87310603</td>
<td>Blastopirellula marina DSM 3645</td>
<td>80</td>
</tr>
<tr>
<td>gi</td>
<td>88800970</td>
<td>Reinekeia sp. MED297</td>
<td>80</td>
</tr>
<tr>
<td>gi</td>
<td>113941581</td>
<td>Herpetosiphon aurantiacus ATCC 23779</td>
<td>82</td>
</tr>
<tr>
<td>gi</td>
<td>15805957</td>
<td>Deinococcus radiodurans R1</td>
<td>83</td>
</tr>
<tr>
<td>gi</td>
<td>94984679</td>
<td>Deinococcus geothermalis DSM 11300</td>
<td>83</td>
</tr>
<tr>
<td>gi</td>
<td>78486138</td>
<td>Thiocapspira crunobena XCL-2</td>
<td>83</td>
</tr>
<tr>
<td>gi</td>
<td>88804711</td>
<td>Robiginitalea bifor mata HTCC2501</td>
<td>82</td>
</tr>
<tr>
<td>gi</td>
<td>119715503</td>
<td>Nocardioidea sp. JS614</td>
<td>81</td>
</tr>
<tr>
<td>gi</td>
<td>114778050</td>
<td>Maripropus ferroxysdans PV-1</td>
<td>82</td>
</tr>
<tr>
<td>gi</td>
<td>117926788</td>
<td>Magnetococcus sp. MC-1</td>
<td>84</td>
</tr>
<tr>
<td>gi</td>
<td>21232788</td>
<td>Xanthomonas campestris pv. campestris</td>
<td>73</td>
</tr>
<tr>
<td>gi</td>
<td>21244215</td>
<td>Xanthomonas axonopodis pv. citri str. 306</td>
<td>74</td>
</tr>
<tr>
<td>gi</td>
<td>78049174</td>
<td>Xanthomonas campestris pv. vesicatoria</td>
<td>74</td>
</tr>
<tr>
<td>gi</td>
<td>84622653</td>
<td>Xanthomonas oryzae pv. oryzae MAFF 311018</td>
<td>74</td>
</tr>
<tr>
<td>gi</td>
<td>87300744</td>
<td>Synechococcus sp. WH 5701</td>
<td>79</td>
</tr>
<tr>
<td>gi</td>
<td>91776960</td>
<td>Methylcarccus flagellatus KT</td>
<td>82</td>
</tr>
<tr>
<td>gi</td>
<td>58580721</td>
<td>Xanthomonas oryzae pv. oryzae KACC10331</td>
<td>74</td>
</tr>
<tr>
<td>gi</td>
<td>153810451</td>
<td>Ruminococcus obeum ATCC 29174</td>
<td>73</td>
</tr>
<tr>
<td>gi</td>
<td>158438431</td>
<td>Clostridium botale ATCC BAA-613</td>
<td>68</td>
</tr>
<tr>
<td>gi</td>
<td>153811783</td>
<td>Ruminococcus obeum ATCC 29174</td>
<td>67</td>
</tr>
<tr>
<td>gi</td>
<td>16125387</td>
<td>Caulobacter crescentus CB15</td>
<td>68</td>
</tr>
</tbody>
</table>
According to another preferred embodiment of said wild type glycoside hydrolase, it contains a phenylalanine (F) residue at position 5 in said motif (4).

According to another preferred embodiment of said wild type glycoside hydrolase, it contains an alanine (A) or proline (P) residue at position 8 in said motif (6), preferably an alanine.

According to another preferred embodiment of said wild type glycoside hydrolase, it contains a phenylalanine (F) or a tyrosine (Y) residue at position 9 in said motif (6), preferably a phenylalanine.

According to another preferred embodiment of the said wild type glycoside hydrolase, it contains a valine (V), a methionine (M) or a glutamic acid (E) residue at position 4 in said motif (7), preferably a valine.

According to another preferred embodiment of said wild-type glycoside hydrolase, it contains an aspartic acid (D) residue at position 7 in said motif (8).

According to another preferred embodiment of said wild-type glycoside hydrolase, it contains a glycine (G), an arginine (R) or a serine (S) residue at position 1 in said motif (9), preferably a glycine.

In a preferred embodiment of said mutant of a glycoside hydrolase, it is a mutant of the 1G5A amylosucrase from Neisseriapoly saccharea having the amino acid sequence SEQ ID NO: 13, wherein said mutant, has in reference to SEQ ID NO: 13, a mutation consisting of,

when said mutant has only one mutation:

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with any amino acid selected from the group consisting of alanine (A), cysteine (C), glutamic acid (E), glycine (G), histidine (H), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), valine (V) with the proviso that said wild type glycoside hydrolase does not contain a valine at this position, tryptophan (W) and tyrosine (Y), or

- the substitution of the phenylalanine (F) at position 229 (F229) corresponding to position 5 in said motif (4), with any amino acid selected from the group consisting of leucine (L), methionine (M) and valine (V), or

- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with any amino acid selected from the
group consisting of glutamic acid (E), phenylalanine (F), glycine (G), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R) and valine (V), or

- the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V) and tryptophan (W), preferably with any amino acid selected from the group consisting of cysteine (C), aspartic acid (D), isoleucine (I), lysine (K) and glutamine (Q), and more preferably with any amino acid selected from the group consisting of aspartic acid (D) and lysine (K), or

- the substitution of the valine (V) residue at position 331 (V331), corresponding to position 4 in said motif (7), with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), serine (S), threonine (T) and tyrosine (Y), or

- the substitution of the aspartic acid (D) residue at position 394 (D394), corresponding to position 7 in said motif (8), with any amino acid selected from the group consisting of alanine (A) and valine (V), or

- the substitution of the arginine (R) residue at position 446 (R446), corresponding to position 1 in said motif (9), with any amino acid selected from the group consisting of cysteine (C), phenylalanine (F), glycine (G), lysine (K), asparagine (N), glutamine (Q), serine (S), threonine (T) and tryptophan (W), or

when said mutant has two mutations:

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with an alanine (A) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a histidine (H), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding position 4 in said motif (4), with a cysteine (C) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with a leucine (L), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding position 4 in said motif (4), with a lysine (K) and the substitution of the
phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of leucine (L) and tryptophan (W), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a leucine (L) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with a leucine (L), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a methionine (M) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with a methionine (M), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with proline (P) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a cysteine (C), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a threonine (T) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with any amino acid selected from the group consisting of histidine (H) and lysine (K), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a valine (V) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with any amino acid selected from the group consisting of leucine (L) and methionine (M), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a valine (V) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of histidine (H), lysine (K), arginine (R), and valine (V), or

- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with a histidine (H) and the substitution of
the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a serine (S), or
- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with a proline (P) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of cysteine (C), isoleucine (I) and leucine (L), or
- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with a threonine (T) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a histidine (H). Unexpectedly, the mutants of a glycoside hydrolase according to the present invention present a specific activity toward D-GlcpNHTCA (NHTCA is iV-trichloroacetyl) or/and substantially improve the glucosylation rate of D-GlcpNAc and α-D-GlcpNAc-OAll. The use of an appropriate combination of a mutant of a glycoside hydrolase of the present invention with a donor and an acceptor as defined in the present invention at an earlier stage of a multi-step synthesis leads to the synthesis of complex oligosaccharides, such as S.flexneri Ia and Ib O-antigens.

Especially, the invention is directed to a method for preparing the building block corresponding to a disaccharide \( \alpha\)-D-glucopyranosyl-(1→4)-2-amino-2-deoxy-\( \alpha\)-D-glucopyranoside of formula and/or \( \alpha\)-D-glucopyranosyl-(1→4)-2-\( \text{NHR} \)-acyl-2-deoxy-\( \alpha\)-D-glucopyranoside of formula (Ia):

\[
(Ia)
\]

advantageously the disaccharide allyl \( \alpha\)-D-glucopyranosyl-(1→4)-2-iV-acetyl-2-deoxy-\( \alpha\)-D-glucopyranoside of formula (I):
said method being characterized in that it comprises the step of reacting a mutant of a glycoside hydrolase as above disclosed, with the acceptor of formula (II), preferably of formula (Ha):

\[
\text{(II)}
\]

wherein \(Y\) is selected from -O- and -S- and \(R\) is selected from the group consisting of: \(C_{1-6}\) alkyl, \(C_{1-6}\) alkenyl, aryl, allyl, -CO-alkyl (\(C_{1-6}\)), -CO-alkenyl (\(C_{1-6}\)), -CO-aryl,

\(R'\) designates a group selected from: acetyl, trichloroacetyl (TCA), trifluoroacetyl (TFA),

wherein aryl designates an aromatic group like phenyl, benzyl, possibly substituted by one or several of the following groups: \(C_{1-4}\) alkyl, -NO\(_2\), a halogen atom, -O-alkyl (\(C_{1-6}\)),

with a donor of formula (IIia):

\[
\text{(IIia)}
\]

wherein \(R_1\) represents a group selected from:

- F,
- \(\text{Ph-NO}_2\),
- \(\text{Py-H-Br}\).
and preferably with the donor of formula (III), sucrose:

(III)
advantageously of formula (I)

Some of the molecules obtained by the method of the invention are new and as such are another object of the invention. Their list is given here-under:

- **AHyI** $\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)$\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)-2-acetamido-2-deoxy-$\alpha$-D-glucopyranoside (XX$_2$),
- **AUyI** $\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)-2-acetamido-2-deoxy-$\alpha$-D-glucopyranoside (XX$_3$),
- Allyl 2,3,4,6-tetra-O-acetyl-$\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)-2-acetamido-3,6-di-O-acetyl-2-deoxy-$\alpha$-D-glucopyranoside (XX$_4$),
- Allyl $\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)-2-amino-2-deoxy-$\alpha$-D-glucopyranoside (XX$_5$),
- Allyl $\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)-2-amino-2-7V,3-O-carbonyl-2-deoxy-$\alpha$-D-glucopyranoside (XX$_6$),
- Allyl 2,4,6-tri-O-benzyl-$\alpha$-D-glucopyranosyl-(l-$\rightarrow$4)-6-O-benzyl-2-benzylamino-2-IV,3-0-carbonyl-2-deoxy-$\alpha$-D-glucopyranoside (XX$_7$).
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-O-benzyl-2-benzylamino-2-α-D-glucopyranoside (XX₉),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-(9-benzyl-2-amino-2-N,3-O-carbonyl-2-deoxy-α-D-glucopyranoside (XX₈A),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-O-benzyl-2-benzylamino-2-deoxy-α-D-glucopyranoside (XX₉₁),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-0-benzyl-2-benzylacetamido-2-deoxy-α-D-glucopyranoside (XX₉,)
- Allyl 2-0-benzoyl-4-0-benzyl-3-(9-chloroacetyl-α-L-rhamnopyranosyl-(1→3)[2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)]-6-O-benzyl-2-benzylacetamido-2-deoxy-α-D-glucopyranoside (XX₁₁),
- Allyl 2-0-acetyl-3,4-di-0-benzyl-L-rhamnopyranosyl-(1→3)[2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)]-6-O-benzyl-2-benzylacetamido-2-deoxy-α-D-glucopyranoside (XX₁₁₁),
- Allyl α-D-glucopyranosyl-(1→4)-2-deoxy-2-trichloroacetamido-α-D-glucopyranoside (XX₁₃),
- Allyl 2,3,4,6-tetra-0-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-0-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranoside (XX₁₄),
- 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranose (XX₁₅),
- 2,3,4,6-Tetra-0-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranosyl trichloroacetimidate (XX₁₆),
- Allyl 2,3,4,6-tetra-0-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranosyl-(1→2)-3,4-di-O-benzyl-α-L-rhamnopyranoside (XX₁₇),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-2-acetamido-3-0-acetyl-6-0-benzyl-2-deoxy-α-D-glucopyranoside (XX₂₄),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-2-acetamido-6-O-benzyl-2-deoxy-α-D-glucopyranoside (XX₂₃).
- the substitution of the amino acid residue at position 4 in said motif (4) with any amino acid selected from the group consisting of alanine (A), cysteine (C), glutamic acid (E), glycine (G), histidine (H), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), valine (V) with the provision that said wild type glycoside hydrolase does not contain a valine at this position, tryptophan (W) and tyrosine (Y), or

- the substitution of the amino acid residue at position 5 in said motif (4) with any amino acid selected from the group consisting of leucine (L), methionine (M) and valine (V), or

- the substitution of the amino acid residue at position 8 in said motif (6) with any amino acid selected from the group consisting of glutamic acid (E), phenylalanine (F), glycine (G), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R) and valine (V), or

- the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V) and tryptophan (W), preferably selected from the group consisting of cysteine (C), aspartic acid (D), isoleucine (I), lysine (K) and glutamine (Q), and more preferably with any amino acid selected from the group consisting of aspartic acid (D) and lysine (K), or

- the substitution of the amino acid residue at position 4 in said motif (7) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), serine (S), threonine (T) and tyrosine (Y), or

- the substitution of the amino acid residue at position 7 in said motif (8) with a valine (V), or

- the substitution of the amino acid residue at position 1 in said motif (9) with any amino acid selected from the group consisting of alanine (A), cysteine (C), phenylalanine (F), glycine (G) with the provision that said wild type glycoside hydrolase does not contain a glycine at this position, lysine (K), asparagine (N), glutamine (Q), serine (S) with the provision that said wild type glycoside hydrolase does not contain a serine at this position, threonine (T) and tryptophan (W), or
when said mutant has two mutations:
- the substitution of the amino acid residue at position 4 in said motif (4) with an alanine (A) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a cysteine (C) and the substitution of the amino acid residue at position 5 in said motif (4) with a leucine (L), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a lysine (K) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of leucine (L) and tryptophan (W), or
- the substitution of the amino acid residues at positions 4 and 5 in said motif (4) respectively with a leucine (L), or
- the substitution of the amino acid residues at positions 4 and 5 in said motif (4) respectively with a methionine (M), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with a cysteine (C), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of histidine (H) and lysine (K), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a valine (V) and the substitution of the amino acid residue at position 5 in said motif (4) with any amino acid selected from the group consisting of leucine (L) and methionine (M), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a valine (V) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of histidine (H), lysine (K), arginine (R), and valine (V), or
- the substitution of the amino acid residue at position 8 in said motif (6) with a histidine (H) and the substitution of the amino acid residue at position 9 in said motif (6) with a serine (S), or
- the substitution of the amino acid residue at position 8 in said motif (6) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of cysteine (C), isoleucine (I) and leucine (L), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H).

Advantageously, the mutants of a glycoside hydrolase according to the present invention having a mutation consisting of:

- the substitution of the amino acid residue at position 4 in said motif (4) with any amino acid selected from the group consisting of alanine (A), cysteine (C), glycine (G), histidine (H), asparagine (N), serine (S), threonine (T), tryptophan (W) and tyrosine (Y), or

- the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of cysteine (C), aspartic acid (D), glutamic acid (E), isoleucine (I) and valine (V),

present also a specific activity toward L-Rha\(^+\), \(\alpha\)-L-Rha\(^+\)-OMe and \(\alpha\)-L-Rha\(^+\)-OAllyl. These mutants also catalyze the glucosylation of \(\alpha\)-L-Rhap-OMe to give the disaccharide \(\alpha\)-D-Glc\(\alpha\)(1\(\rightarrow\)3)\(-\alpha\)-L-Rhap-OMe.

In a preferred embodiment of said mutant of a wild-type glycoside hydrolase, it is a mutant of the 1G5A amylaseucrase from \textit{Neisseria polysaccharea} having the amino acid sequence SEQ ID NO: 13, wherein said mutant, has in reference to SEQ ID NO: 13, a mutation consisting of

- the substitution of the isoleucine (I) residue at position 228 (1228) with any amino acid selected from the group consisting of alanine (A), cysteine (C), glutamic acid (E), glycine (G), histidine (H), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), valine (V), tryptophan (W) and tyrosine (Y), or

- the substitution of the phenylalanine (F) at position 229 (F229) corresponding to position 5 in said motif (4), with any amino acid selected from the group consisting of leucine (L), methionine (M) and valine (V), or
- the substitution of the alanine (A) residue at position 289 (A289) with any amino acid selected from the group consisting of glutamic acid (E), phenylalanine (F), glycine (G), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R) and valine (V), or

- the substitution of the phenylalanine (F) residue at position 290 (F290) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V) and tryptophan (W), or

- the substitution of the valine (V) residue at position 331 (V331) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), serine (S), threonine (T) and tyrosine (Y), or

- the substitution of the aspartic acid (D) residue at position 394 (D394) with a valine (V), or

- the substitution of the arginine (R) residue at position 446 (R446) with any amino acid selected from the group consisting of cysteine (C), phenylalanine (F), glycine (G), lysine (K), asparagine (N), glutamine (Q), serine (S), threonine (T) and tryptophan (W),

when said mutant has two mutations:

- the substitution of the isoleucine (I) residue at position 228 (I228), corresponding to position 4 in said motif (4), with an alanine (A) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a histidine (H), or

- the substitution of the isoleucine (I) residue at position 228 (I228), corresponding position 4 in said motif (4), with a cysteine (C) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with a leucine (L), or

- the substitution of the isoleucine (I) residue at position 228 (I228), corresponding position 4 in said motif (4), with a lysine (K) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of leucine (L) and tryptophan (W), or
- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a leucine (L) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4) with a leucine (L), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a methionine (M) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with a methionine (M), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with proline (P) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a cysteine (C), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a threonine (T) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of histidine (H) and lysine (K), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a valine (V) and the substitution of the phenylalanine (F) residue at position 229 (F229), corresponding to position 5 in said motif (4), with any amino acid selected from the group consisting of leucine (L) and methionine (M), or

- the substitution of the isoleucine (I) residue at position 228 (1228), corresponding to position 4 in said motif (4), with a valine (V) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of histidine (H), lysine (K), arginine (R), and valine (V), or

- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with a histidine (H) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a serine (S), or

- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with a proline (P) and the substitution of the
phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with any amino acid selected from the group consisting of cysteine (C), isoleucine (I) and leucine (L); these mutants being preferred because they present a glucosylation rate of D-GlcNAc of 100%, or

- the substitution of the alanine (A) residue at position 289 (A289), corresponding to position 8 in said motif (6), with a threonine (T) and the substitution of the phenylalanine (F) residue at position 290 (F290), corresponding to position 9 in said motif (6), with a histidine (H).

According to a preferred embodiment of said mutant of a glycoside hydrolase, it is a mutant of the 1G5A amyloucrose from Neisseria polysaccharea which has in reference to SEQ ID NO: 13, a mutation consisting of the substitution of the phenylalanine (F) residue at position 290 (F290) with any amino acid selected from the group consisting of cysteine (C), aspartic acid (D), isoleucine (I), lysine (K) and glutamine (Q), and more preferably with any amino acid selected from the group consisting of aspartic acid (D) and lysine (K). These mutants present a glucosylation rate of D-GlcNAc of higher than 50%.

The present invention also provides polynucleotides encoding a mutant of a glycoside hydrolase according to the present invention.

Polynucleotides of the invention may be obtained by the well-known methods of recombinant DNA technology and/or of chemical DNA synthesis. These methods also allow introducing the desired mutations in a naturally occurring DNA sequence.

The invention also provides recombinant DNA constructs comprising a polynucleotide of the invention, such as expression cassettes wherein said polynucleotide is linked to appropriate control sequences allowing the regulation of its transcription and translation in a host cell and optionally to a sequence encoding a GST tag allowing a rapid purification of the mutant enzymes and recombinant vectors comprising a polynucleotide or an expression cassette of the invention.

Another object of the invention is a method for determining whether a wild type protein is a wild type glycoside hydrolase, said method comprising the steps of:

a) determining the amino acid sequence of said protein,

b) identifying in the amino acid sequence of said protein, preferably from the N- to C-terminus, eleven motifs defined by the following consensus motifs:
(1) the amino acid sequence LGVNYLHLMPL (SEQ ID NO: 1);
(2) the amino acid sequence DGGYAV (SEQ ID NO: 2);
(3) the amino acid sequence DFVFNH (SEQ ID NO: 3);
(4) the amino acid sequence LREIFPDTPGNF (SEQ ID NO: 4);
(5) the amino acid sequence FNSYQWDLN (SEQ ID NO: 5);
(6) the amino acid sequence ILRLDAVAFWKL (SEQ ID NO: 6);
(7) the amino acid sequence EAIYVHDD (SEQ ID NO: 7);
(8) the amino acid sequence YVRCHDDI (SEQ ID NO: 8);
(9) the amino acid sequence RISGTLASLAG (SEQ ID NO: 9);
(10) the amino acid sequence GIPLIYLGDE (SEQ ID NO: 10);
(11) the amino acid sequence RWVHRP (SEQ ID NO: 11);

c) determining the sequence identity percent or sequence similarity percent between the sequence formed by said eleven motifs joined end-to-end from motif (1) to motif (11) with the amino acid sequence SEQ ID NO: 12, and if the sequence identity percent is at least 65%, preferably at least 70%, and by order of increasing preference, at least 75%, 80%, 85%, 90%, 95%, 95%, 97%, 98%, and 99%, or 100% or if the sequence similarity is at least 80%, preferably at least 85%, and by order of increasing preference, at least 90%, 95%, 95%, 97%, 98%, and 99%, or 100%, then the wild type protein is a wild type glycoside hydrolase.

In addition to the preceding features, the invention further comprises other features which will emerge from the following description, which refers to examples illustrating the present invention, as well as to the appended figures.

Figure 1 shows the repeating unit of S. flexneri serotypes Ia and Ib O-Antigen. Knap = rhamnopyranosyl - GlcpNAc = 2-JV-acetyl-2-deoxy-glucopyranosyl - Glcp = glucopyranosyl - Ac = acetyl.

Figure 2 shows the first steps of chemo-enzymatic routes (A, C, D) to potential synthetic intermediates to oligosaccharide fragments of S. flexneri Ib and/or 1a O-antigens, a chemical synthetic route (B) to S. flexneri 1a pentasaccharides (ref. 66), and the chemical synthesis (E) of a model disaccharide intermediate to oligosaccharide fragments of S. flexneri Ib and/or 1a O-antigens. In Figure 2A: Chemo-enzymatic synthesis of disaccharide XX₃, a. AlI(OH), BF₃OEt₂; b. non-purified F290K extract, sucrose; c. Ac₂O, Pyridine; d. MeONa, MeOH; e. Amyloglucosidase from Aspergillus niger, acetate buffer (pH 4.8). In Figure 2C: a. MeONa, MeOH; b. Ba(OH)₂, 8H₂O, H₂O,
90 °C; c. PNO$_2$C$_6$H$_4$OCOCl, MeONa, MeOH; d. (i) CCl$_3$CH$_2$OCOCl, NaOMe, (ii) BnBr, NaH; e. (i) BnBr (6 eq), (ii) NaH; f. (i) NaH, (ii) BnBr (5.5 eq); g. IM aq. NaOH, 1,4-dioxane, 70 °C; h. Ac$_2$O, Pyridine; i. TMSOTf, MS 4A, Toluene, 0°C; j. TMSOTf, MS 4A, Toluene, -100°C. In Figure 2D: a. MeONa, MeOH; b. Ba(OH)$_2$-SH$_2$O, H$_2$O, 90 5 °C; c. (Cl$_3$Ac)$_2$O, NaOMe, O 0°C; d. Ac$_2$O, Pyridine; e. H$_2$, cat. [Ir(COD)\{PCH$_2$(C$_6$H$_5$)$_2$\}]$_2$^+$PF$_6$^-$, THF then I$_2$, H$_2$O; f. CCl$_3$CN, DBU, -5 °C; g. TMSOTf, MS 4A, Toluene, -60 °C.

Figure 3 shows the reaction catalyzed by glucansucrases. 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 to an acceptor which depends on the conditions of reaction may be (i) water to give glucose (H) fructose to form sucrose isomers (Hi) glucose released from hydrolysis to form soluble oligosaccharides, or (iv) an exogeneous hydroxylated acceptor.

Figure 4 shows the architecture of the active site in complex with maltoheptaose (G7).

Figure 5 shows the comparison of docking modes: (A) Maltose moiety from the crystallographic maltoheptaose (PDB: IMWO) occupying binding subsites (-1) and (+1) of amylosucrase from Neisseria polysaccharea and (B) α-D-Glc-/\(\text{GlcpNAc}\) in the active site of AS. The seven amino acid residues (1228, A289, F290, 1330, V331, D394 and R446) selected for mutagenesis are shown on the figures. Hydrogen atoms have been omitted on the figures for clarity purpose.

Figure 6 shows the screening of the library for their ability to synthesize the desired disaccharide: α-D-Glc-/\(\text{GlcpNAc}\)OAll using the most improved 1G5A variant F290K and F290D and corresponding HPLC chromatogramm (with UV$_{220nm}$ detection) comparing F290K, F290D and ASNPwt. Initial reaction conditions: Sucrose=Acceptor=146mM. At final time, sucrose was fully consumed. Conversion Rate= (Q(Acceptor)$_0$-Q(Acceptor)$_t$)/Q(Acceptor)$_0$ where Q(X)= Quantity of X in moles. % Glc transferred onto acceptor derivatives = [Q(Glucosyl units transferred onto acceptor derivatives)/Q(Glucosyl units transferrable from initial
sucrose) x 100. % Monoglucosylated acceptor = [Q(Monoglucosylated acceptor)/Q (acceptor derivatives)] x 100. % Diglucosylated acceptor = [Q(Diglucosylated acceptor)/ Q (acceptor derivatives)] x 100. G=Glucose; F=Fructose; GF=Sucrose; DP<sub>1</sub>= Acceptor; DP<sub>2</sub>= monoglucosylated Acceptor; DP<sub>3</sub>= diglucosylated Acceptor (to=initial time of the reaction - the medium contains only the donor and the acceptor, no products have been formed yet).

Figure 7B shows the comparison of Dionex HPAEC product profiles obtained at the end of the reaction \( t = 24 \) h with wild-type AS and the variant F290K using 146 mM sucrose.

Figure 7C shows the comparison of Dionex HPAEC product profiles obtained at the end of the reaction \( t = 24 \) h with wild-type AS and the variant F290K using 146 mM sucrose supplemented with 146 mM acceptor (\( \alpha \)-D-GlcpNAc-OAll).

Figure 7D shows the comparison of kinetic parameters between ASNPwt and the variant F290K.

Figure 7E shows the determination of kinetic parameters for the variant F290K and ASNPwt catalyzed reactions: (a) varied acceptor (b) varied donor. Figure 7E(a), varied acceptor: Allyl 2-/N/-acetyl-2-deoxy-\( \alpha \)-D-glucopyranoside, constant donor: 250 mM. Figure 7E(b), varied donor, constant acceptor: Allyl 2-/N/-acetyl-2-deoxy-\( \alpha \)-D-glucopyranoside.

Figures 8.1 to 8.8 show the sequence alignment of 34 wild type glycoside hydrolases using the CLUSTALW program under default parameters.

Figure 9 shows the alignment of the eleven different motifs found in 34 wild type glycoside hydrolases.

Figure 10 shows transglucosylation rates of D-GlcpNAc derivatives using the most improved variant F290K and ASNPwt. Initial reaction conditions: Sucrose=Acceptor=146mM. At final time, sucrose was fully consumed. Conversion Rate= \( (Q(\text{Acceptor})_{t=24\text{h}}-Q(\text{Acceptor})_{t=0})/Q(\text{Acceptor})_{t=0} \) 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})) \times 100. \%

Monoglucosylated acceptor = \( [Q(\text{Monoglucosylated acceptor})/Q(\text{acceptor derivatives})] \times 100. \%

Diglucosylated acceptor = \( [Q(\text{Diglucosylated acceptor})/Q(\text{acceptor derivatives})] \times 100. \%

Triglucosylated acceptor = \( [Q(\text{Triglucosylated acceptor})/Q(\text{acceptor derivatives})] \times 100. \%\)
Figure 11 shows the structure of disaccharide \( a-D-Glcp-(1\rightarrow4)-D-GlcpNAc \) (P2) obtained by AS-mediated glucosylation of D-GlcpNAc.

Figure 12 shows the strategy adopted for the construction of the four double-mutant libraries; (A): libraries 1, 2 and 3; (B) library 4.

Figure 13 shows the comparison of amyllose synthesis by wtAS, variant F290K and double-mutants A289P-F290C, A289P-F290I, A289P-F290L from 250 mM sucrose. In Figure 13A: superposition of the HPAEC-PAD profiles obtained at the end of the reaction (final time = 24h) with wtAS, variant F290K and double-mutants A289P-F290C, A289P-F290I, A289P-F290L. In Figure 13B: yields of glucosyl units incorporated into the various products synthesized in the total reaction medium by wtAS, variant F290K and double-mutants A289P-F290C, A289P-F290I, A289P-F290L.

Figure 14 shows the comparison of \( \alpha-D-GlcpNAc-OAll \) (D') transglucosylation with wtAS, variant F290K and double-mutants A289P-F290C, A289P-F290I, A289P-F290L from 250 mM sucrose supplemented with 250 mM \( \alpha-D-GlcpNAc-OAll \). In Figure 14A: Superposition of the HPAEC-PAD profiles obtained at the end of the reaction (final time = 24h) with wtAS, variant F290K and double-mutants A289P-F290C, A289P-F290I, A289P-F290L. In Figure 14B: Yields of mono- and di-glucosylated \( \alpha-D-GlcpNAc-OAH \) obtained with wtAS, variant F290K and double-mutants A289P-F290C, A289P-F290I, A289P-F290L. At final time, >90% of sucrose was consumed.

EXAMPLE 1: Engineering transglucosidase for the synthesis the \( \alpha-D-glucopyranosyl-(1 \rightarrow 4)\)ViV-acetyl-\( \alpha/\beta\)-D-glucopyranosaminyl disaccharide

1) Materials and Methods

Bacterial strains, plasmids and chemicals

Plasmid pGST-AS, derived from the pGEX-6P-3 (GE Healthcare Biosciences) and containing the \( N. polysaccharae \) amylosucrase encoding gene (ref. 45) was used for the construction of the AS single mutant library.

\( E. coli \) JM1 09 was used as host for the plasmid library transformation, gene expression and large-scale production of the selected mutants.

Sucrose, iV-acetyl-D-glucosamine and glycogen were purchased from Sigma-Aldrich (Saint-Louis, MO, USA). Known allyl 2-acetamido-2-deoxy- \( \alpha-D-glucopyranoside \) (\( \alpha-D-GlcpNAc-OAll \)) (ref. 69), allyl 2-acetamido-2-deoxy- \( \beta-D-glucopyranoside \) (ref. 70) (\( \beta-D-GlcpNAc-OAH \)), methyl 2-acetamido-2-deoxy- \( \alpha-D-glucopyranoside \) (\( \alpha-D-GlcpNAc-OAll \)), and \( \beta-D-glucopyranoside \) (\( \beta-D-GlcpNAc-OAH \)).
glucopyranoside (ref. 71) (α-D-GlcNAc-OMe), methyl 2-acetamido-2-deoxy-β-D-glucopyranoside (ref. 70) (β-D-GlcNAc-OMe), and 2-deoxy-2-trichloroacetonio -β-D-glucopyranose (ref. 72) (D-GlcNHTCA) were synthesized chemically.

The reference disaccharides α-D-Glc-(1→4)-D-GlcNAc and Glc-(1→3)-D-GlcNAc were enzymatically synthesised and characterized (see Example 7).

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 synthetised by Eurogenetec (Liege, Belgium).

DNA extraction (QIAspin) and purification (QIAQuick) columns were purchased from Qiagen (Chatsworth, CA).

Wild type glycoside hydrolase: amylosucrase (ASNPwt) 1G5A of sequence SEQ ID NO: 13.

Selection of mutation position by molecular modelling

Starting models for the disaccharide and for AS: The disaccharide α-D-Glc-(1→4)-D-GlcNAc was constructed with the monosaccharide obtained from a database of carbohydrate three-dimensional structures. All molecular modelling calculations were performed using the SYBYL 7.3 software. The coordinates of amylosucrase were taken from the 2.0 A resolution crystal structures of amylosucrase from N. polysaccharea in complex with sucrose (PDB: IIGI) and maltoheptaose, a reaction product (PDB: IMWO). All hydrogen atoms were added to the enzyme and their position optimized with the Tripos force field.

Systematic conformational search for the disaccharide: Both anomeric forms of the α-D-Glc-(1→4)-D-GlcNAc, 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: Φ = O5'-C1'-O4'-C4 and Ψ = C1'-O4-C4-C3 for α-D-Glc-(1→4)-D-GlcNAc. The MM3 force field implemented in SYBYL 7.3 software was used for this purpose together with the energy parameters appropriate for carbohydrates. Different maps were constructed with the dielectric constant set to 4.0 and 78.0 (to mimic gas phase and water environment, respectively). 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 obtained for all disaccharides were then used to locate the different energy minima that were subsequently fully relaxed.

**Docking of disaccharide in the binding site of AS:** The lowest energy conformations identified on the disaccharide potential energy maps were used as starting structures to be docked in the binding site of amylase. This was performed by superimposing the glucosyl unit of the disaccharides at (-1) subsite 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.

**Construction of mutant library**

Single mutagenesis, focused on +1 subsite amino acids retained 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 II below). XXX codon indicates the bases which were used to obtain the replacement by the desired amino acids and are listed in Table III below.
Table II: Primers used to generate the 19 monomutants for the positions 228, 289, 290, 330, 331, 394 and 446.

<table>
<thead>
<tr>
<th>SEQ ID</th>
<th>Primer</th>
<th>Nucleotide</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>NO:</td>
<td>Name</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>I228for</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>I228rev</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>A289for</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>A289rev</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>F290for</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>F290rev</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>I330for</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>I330rev</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>V331for</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>V331rev</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>D394for</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>D394rev</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>R446for</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>R446rev</td>
</tr>
<tr>
<td>Ala</td>
<td>Cys</td>
<td>Asp</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>for</td>
<td>GCC</td>
<td>TGC</td>
</tr>
<tr>
<td>rev</td>
<td>GGC</td>
<td>GCA</td>
</tr>
<tr>
<td>for</td>
<td>wt</td>
<td>TGC</td>
</tr>
<tr>
<td>rev</td>
<td>GCA</td>
<td>GTC</td>
</tr>
<tr>
<td>for</td>
<td>GCT</td>
<td>TGT</td>
</tr>
<tr>
<td>rev</td>
<td>AGC</td>
<td>ACA</td>
</tr>
<tr>
<td>for</td>
<td>GCC</td>
<td>TGC</td>
</tr>
<tr>
<td>rev</td>
<td>GGC</td>
<td>GCA</td>
</tr>
<tr>
<td>for</td>
<td>GCC</td>
<td>TGC</td>
</tr>
<tr>
<td>rev</td>
<td>GGC</td>
<td>GCA</td>
</tr>
<tr>
<td>for</td>
<td>GCC</td>
<td>TGC</td>
</tr>
<tr>
<td>rev</td>
<td>GGC</td>
<td>GCA</td>
</tr>
<tr>
<td>for</td>
<td>GCT</td>
<td>TGT</td>
</tr>
<tr>
<td>rev</td>
<td>AGC</td>
<td>ACA</td>
</tr>
</tbody>
</table>

Table III: Sequence of XXX codon used to replace each selected amino acids by the 19 other ones
PCR amplification was carried out with Pfu DNA polymerase (2.5 U) for 16 cycles (95°C, 30s; 55°C, 30s; 72°C 2min). The DNA was digested with Dpn/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 (1228Ai, 1228Vp, I228Y1, A289D1, F290Di, F290K1, F290Qi, 1330A1, 1330D1, 1330Ei, 1330Fi, 1330Ti, 1330Wi, V331Ai, V331Si, 0394V1 and R446K0 were sequenced on the entire gene and showed no other mutations by Millegen (Labege, France) or Cogenics (Meylan, France).

**Expression of mutant library**

The protocol was established to enable the rapid identification of clones for which D-GlcNAc glucosylation was improved. To obtain higher amounts of enzymes and facilitate detection of glucosylated compounds upon HPLC screening, monomutants were produced in 96-DeepWell Format plates. Storage microplates containing 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 (100µg/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 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, 3000g, 4°C) and 300µL of the supernatant was transferred to a filter micro-plate (PVDF 0.2 µm) 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.
Development of the BBT test in liquid medium

Mono-mutant library was cultured in 96-well microplates as previously described (ref. 81). 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 1µL 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 (dinitrosalicylic) assay (ref. 57) for comparison. A volume of 100µL of reaction mixture was added to 100µL of dinitrosalicylic acid 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 assay (ref. 82), amylose formation was detected by adding 1µL 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.

Development of the BBT test at colony level

On day 1, libraries were transformed into electrocompetent E. coli TOPIO 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 300°C, they were stored at -20°C and -80°C. Inactive clones (blue) were scraped, cultivated and stored in the same conditions.

**Screening of mutant library**

Efficiency of the glucosylation reaction was evaluated by HPLC analysis of the acceptor reaction product synthesized when using D-GlcpNAc as acceptor 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 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 nL/min. HPX-87K column was used to determine sucrose consumption by RI detection and concomitant α-D-Glcp-(1→4)-D-GlcpNAc formation by UV=220nm detection.

**Production and Purification of the Selected Variants: F290D and F290K**

Production and purification of AS variants were performed as previously described (ref. 45). 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 standard (ref. 56).

**Biochemical Characterization of the Selected 1G5A Variants: F290D and F290K**

All assays were performed at 30°C in 50 mM Tris buffer, pH=7.0. For the acceptor reactions, F290D, F290K and ASNPwt (1G5A) were tested on α-D-GlcpNAc-OAll.

*Standard activity determination.* Specific activity of the purified enzymes was determined by measuring the initial rate of released fructose under standard conditions (146 mM sucrose). Fructose concentration was determined using the dinitrosalicylic acid (DNS) method (ref. 57). One unit of AS variant corresponds to
the amount of enzyme that catalyses the production of 1 µmole 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 18 000 g for 10 min, filtered on 0.22 µm membrane and analyzed by
HPLC, as previously described.

Different carbohydrate analyses were performed to compare the
product profiles synthesized by ASNPwt and AS variant (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 (150 mM NaOH) was set at 1 mL/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 α-D-Glc/?NAc-OAll (acceptor) and its derivatives are
not oxidable products and thus are not detectable by HPAEC. Sucrose, glucose,
fructose, α-D-GlcNAc-OAll (acceptor) and its derivatives (glucosylation products)
were quantified by HPLC, as previously described

Kinetics studies. (k_cat,K_v)

Enzyme assays were carried out in a total volume of 2 mL containing
pure enzyme (0.115 mg and 0.106 mg when using ASNPwt and F290K, respectively).
Catalytic efficiency (Eff= k_cat/K_v) of ASNPwt and F290K variant was determined with
both sucrose (D=Donor) and α-D-GlcNAc-OAll (A=Acceptor) as variable substrates.
For the determination of the catalytic efficiency for D, A was held constant at 250 mM
and D was varied between 0 and 600 mM. For the determination of the catalytic
efficiency for A, D was held constant at 250 mM and A was varied between 0 and 250
mM. Experiments were performed until A and D solubility limits were reached. For
each experiment, the reaction velocity corresponding to the acceptor glucosylation was
determined by the formation of α-D-Glc-(1 →4)-α-D-GlcNAc-OAll (called Vi(Dp2)),
corresponding to the kinetics of the reaction of interest.
Initial velocities were fitted to the Michaelis-Menten equation using Sigma-Plot. As saturation was not achieved with the mutant, efficiency was calculated by linear regression analysis of the velocity versus substrate concentration plot. Aliquots (200 µL) 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 µm membrane and analyzed using HPLC material previously described. HPLC analyses were performed using 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 to determine the released fructose by RI detection and to detect the formation of α-D-Glcj9-(1→4)-α-D-GlcpNAc-OAll by UV=220nm detection.

Preparative synthesis of the acceptor reaction product

In order to characterize the products of D-GlcpNAc glucosylation, a synthesis was carried out at the preparative scale. 10 ml mixture containing 146 mM sucrose, 146 mM D-GlcpNAc 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 0.22 µm membrane.

In the same way, oligosaccharides produced from α-D-GlcpNAc-OAll glucosylation were produced in a 100 mL mixture reaction (146 mM sucrose + 146 mM α-D-GlcpNAc-OAll), using 4 LVMl of non-purified F290K extract (sonication supernatant).

Purification of the oligosaccharides

Concerning D-GlcpNAc glucosylation, monoglucosylated product was 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 verify the compound’s purity.

Structural analysis of the acceptor reaction products

The structure of Dp2, which was synthesized using sucrose as donor and D-GlcpNAc as acceptor with AS mutant F290K was analyzed by HRMS and NMR.
It corresponds to \( \alpha\)-D-glucopyranosyl-(1\( \rightarrow \)4)-Af-acetyl-D-glucosamine and was found identical to Dp2 formed by ASNPwt.

HRMS (FAB): Anal. Calcd for \( C_{14}H_{25}NO_{11}Na \): 406.1325 [MNa\(^+\)]. Found: 406.1356; \( ^1\)H and \( ^{13}\)C listed in the Table IV below.

Table IV. \( ^1\)H and \( ^{13}\)C chemical shifts (ppm, \( D_2\)O) and \( ^3\)J coupling constants (Hz) of P2 resulting from F290K-mediated glucosylation ofD-GlcpNAc

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Residue</th>
<th>( \delta ) (ppm)</th>
<th>Carbohydrate Ring</th>
<th>N-Acetyl group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>( \alpha)-D-Glcp</td>
<td>( ^1)H</td>
<td>5.39</td>
<td>3.56</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>( ^{13})C</td>
<td>99.87</td>
<td>72.06</td>
<td>73.27</td>
</tr>
<tr>
<td>( \beta)-D-GlcpNAc</td>
<td>( ^1)H</td>
<td>4.70</td>
<td>3.68</td>
<td>3.79</td>
</tr>
<tr>
<td>P2 (( \beta) anomer)</td>
<td>( ^{12})C</td>
<td>95.25</td>
<td>57.09</td>
<td>74.79</td>
</tr>
<tr>
<td>( \alpha)-D-Glcp</td>
<td>( ^1)H</td>
<td>5.40</td>
<td>3.56</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>( ^{13})C</td>
<td>109.15</td>
<td>72.17</td>
<td>73.30</td>
</tr>
<tr>
<td>( \alpha)-D-Glcp</td>
<td>( ^1)H</td>
<td>5.18</td>
<td>3.89</td>
<td>4.00</td>
</tr>
<tr>
<td>P2 (( \alpha) anomer)</td>
<td>( ^{12})C</td>
<td>91.13</td>
<td>54.36</td>
<td>71.62</td>
</tr>
</tbody>
</table>

2) Results

10 Choice of acceptor substrates

The choice of a suitable precursor to residue D (Figure 1) to be used as acceptors in the enzymatic steps to \( S. flexneri \) serotype 1a and 1b O-antigen was crucial. Three major features had to be taken into account: (i) light protecting pattern compatible with the limited ability of selected glucansucrases to modulate their acceptor binding site, (ii) easy synthetic access, and (iii) the possible conversion of the glucosylation product into a protected disaccharide building block known to be compatible with additional chemical elongation. The selection of an appropriate precursor to residue D (D') was made according to the same criteria. Allyl 2-iV-acetyl-2-deoxy-\( \alpha\)-D-glucopyranoside (XXi) was selected based on the assumption that the \( 3_{\text{D}}\)-OH group would be easily differentiated at the disaccharide level providing that a 2,3-oxazolidinone moiety could be introduced following iV-deacetylation. Regioselective
differentiation of the 3D-0H is indeed a pre-requisite to any specific chain elongation at this position as required in the synthesis of S. flexneri Ia and Ib oligosaccharides.

Thus, isolation of ED' (XX4) was best performed following rough chromatography of the crude enzymatic glucosylation mixture issued from D' (XXi), peracetylation into pure XX4, then transesterification of intermediate XX4 (30% from XXi) (Scheme 2A). To access S. flexneri oligosaccharides bearing the ED branching pattern, the product of XXi enzymatic glucosylation, disaccharide XX3 had to be turned into a donor allowing chain elongation at the reducing end (XXi6) (Scheme 2D) and/or an acceptor allowing chain extension at position 3D XX10 (Scheme 2C). Therefore, selective iV-trichloroacetylation and per-O-acetylation of XX5, resulting from extensive deacetylation of XX4, gave XXi4 (58%), which was converted, via hemiacetal XX15, to trichloroacetimidate XXi6 (26%) bearing a participating group at position 2D as required (Scheme 2D). Alternatively, XX4 was saponified into XX5 (83%) (Scheme 2C), which was turned into the benzylated 2D,3D-oxazolidinone XX8 via XX6 (34%). Oxazolidinone cleavage and subsequent JV-acetylation gave acceptor XX10 (80%) bearing a free hydroxyl group at position 3 and a participating group at position 2. Both disaccharide donor XXi6 and disaccharide acceptor XX10 were converted to trisaccharides by reaction with a rhamnopyranoside acceptor or a rhamnopyranosyl donor.

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 can 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, the method was applied to the library of 133 mono mutants constructed. 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 dinitosalicylic assay (DNS) (ref. 57) which enables the detection of the reducing power released from sucrose cleavage. 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. 54% of active clones were detected in this library. By comparison, only 31% of the recombinant clones were detected as active with the DNS assay. Notably, all the clones detected with the DNS assay were also detected using the new colorimetric assay

**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\(_{\text{amp}}\) medium supplemented with sucrose, IPTG and Tris-Buffer at pH = 7.5 to maintain medium staining (blue color). The BBT concentration was first optimized from 0.05 g/L to 0.5 g/L. It was found that 0.1 g/L of BBT offered the best contrast between blue and yellow colonies after incubation at 30°C during 48h. This protocol was first applied to freshly transformed cells and it was 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\(_{\text{amp}}\) 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\(_{\text{amp}}\) 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.
Screening for native transglucosidases able to synthesize the starting building block

Selected α-retaining transglucosidases are glucansucrases found in families 13 and 70 of glycoside-hydrolases (ref. 51). They catalyze the synthesis of α-glucan polymers by successive transfers of α-D-glucopyranosyl units from sucrose without any mediation of sugar nucleotides. Using the high energy of the sucrose bond to catalyze condensation reaction, they stand among the most efficient transglucosidases in the glycoside-hydrolase family. Depending on regiospecificity of the enzyme, distinct types of glucosidic linkage are found in the polymer formed. Notably, polymerization reaction can be redirected toward the glucosylation of exogenous acceptors, when the latter are well recognized (Figure 3). In addition, glucansucrases generally possess a broad acceptor spectrum, what indicates a certain plasticity of the acceptor recognition at the acceptor binding site. However, none of them had yet been tested for the glucosylation of the starting acceptor of interest. Glucosylation of D-Glc/?NAc was thus attempted with four recombinant glucansucrases, which were selected for their very distinct specificities. For this first screening it has been preferred to start with commercially available D-GlcpNAc instead of the allyl-derivative, assuming that the modification with an allyl group at the anomic position would not significantly affect the acceptor recognition as preliminary studies performed on D-Glcp and α-D-Glcp-OAlkyl derivatives with glucansucrases previously showed negligible differences. Enzymes specific for α-1,6 and α-1,3, α-1,2 or α-1,4 glucosidic bond formation were tested in the presence of the target acceptor. None of them was able to glucosylate this acceptor with the requisite regiospecificity and good yields.

Consequently, to overcome the limited substrate recognition by the enzymes, it was opted for the engineering of novel transglucosidases with altered regio and stereospecificities. The combination of combinatorial and rational enzyme engineering in the form of focused small size libraries was used. Further, the three-dimensional structure of AS (ref. 52) was available in complex with either the substrate or the natural reaction product.

**Engineering of** amylosucrases **able to glucosylate the target building block**

To modify enzyme specificity, an approach based on site-directed evolution targeted at the binding pocket was followed. The catalytic site pocket is
defined by the subsites (-1) and (+1) according to the nomenclature earlier described for glycoside hydrolases (Figure 4). The subsite (-1) is responsible for the specificity towards sucrose and is occupied by the glucosyl unit which will be transferred whereas the subsite (+1) ensures a correct positioning of the acceptor and is also responsible for specificity of synthesis of the ($\alpha$-1→4) glucan linkage (ref. 46).

It was combined both the rational selection of mutation targets at the acceptor binding site (noted +1 subsite in Figure 4) and for each of the identified positions, a systematic modification of the residue by all 19 remaining possible amino acids. As (+1) subsite is responsible for the enzyme specificity toward acceptors, the approach consisted in (i) mapping the binding site residues important for functional plasticity and (ii) identifying the most promising positions to be modified to favour acceptor recognition. Starting from the crystallographic structure of AS in complex with sucrose (PDB: UGI) (ref. 54) and maltoheptaose (PDB: IMWO) (ref. 55), the target disaccharide $\alpha$-D-Glc-(l→4)-D-GlcNAc was docked in the AS active site (Figure 5).

**Structural analysis of $\alpha$-D-Glc-(l→4)-D-GlcNAc [EDJ: AS complex**

The desired disaccharide $\alpha$-D-Glc-(l→4)-D-GlcNAc was docked into the AS active site using the crystallographic maltose glucosyl units (*i.e.* $\alpha$-D-Glc-(1→4)-D-Glc: native product) bound at (-1) and (+1) subsites (PDB: IMWO) as a template for the starting location. As amylosucrase synthesizes naturally an $\alpha$-(1-4) osidic linkage, the docking mode adopted by $\alpha$-D-Glc-(l→4)-D-GlcNAc (target product) was close to the one observed for maltose glucosyl units at subsites (-1) and (+1) by crystallography. The unique difference between both disaccharides is the JV-acetyl group at the C2 carbone of the D-GlcNAc unit that substitutes the hydroxyl group carried by the D-Glc in maltose. Modeling results indicated that to accommodate the D-GlcNAc moiety at the (+1) subsite, several hydrophobic residues (Ala289, Phe290, Ile330 and Val331) surrounding the JV-Acetyl group had to move away to provide enough fitting space. Monomutants for position Ile228, Asp394 and Arg446 have been also considered in the screen.

Out of the 18 residues identified as surrounding the (+1) subsite, 7 positions that were presumed to be not critical for sucrose binding but beneficial for target acceptor glucosylation have been selected for mutagenesis: Ile228, Ala289, Phe290, Ile330, Val331, Asp394 and Arg446. It was systematically mutated the selected
amino acids by the 19 other possible residues to create a small size library focused on 
+1 subsite.

**Detection of mutants able to glucosylate the target acceptors**

For each of the 7 selected positions, 19 single mutants were generated (corresponding to each possible amino acid change). Site-directed mutagenesis has been preferred to saturation mutagenesis, as each mono mutant generated is directly identified and can be easily isolated and characterized. A first library of 133 monomutants (7x19) was thus obtained and stored in cryotubes and 96-wells microplates. The mutants were tested for the glucosylation of the target acceptor D-GlcNAc in microtiter format experiments. HPLC screening was performed to identify those able to form the desired disaccharide. Both sucrose consumption and disaccharide formation were determined in order to calculate the glucosylation rate defined as the molar ratio of monoglucosylated acceptor versus/sucrose consumed.

D-GlcNAc was poorly recognized by the wild type AS (glucosylation rate-2%). The remodelling of the +1 subsite led to performant results. Indeed, 17 mutants catalyzed the formation of the \(\alpha\)-D-Glc-(1\(\rightarrow\)4)-D-Glc/\(\beta\)NAc with a glucosylation rate comprised between 10 and 50% and 5 mutants with a glucosylation rate higher than 50% (Figure 6). Position 290 is clearly a key position to improve the formation of \(\alpha\)-D-Glc-(1\(\rightarrow\)4)-D-GlcNAc. Of the 19 mutants, five synthesize the desired product with the correct regiospecificity and a glucosylation rate higher than 50%. Notably, F290D and F290K yielded the desired disaccharide with glucosylation rate of more than 90%, which represent a 45 fold increase compared to the wild type. Mutations at positions 228, 289, 331 and 446 also led to the improvement of the \(\alpha\)-D-Glc-(1\(\rightarrow\)4)-D-GlcNAc synthesis.

In overall, two mutants of interest for the chemo-enzymatic pathway were retained and further characterized F290D and F290K which are specific for the production of \(\alpha\)-D-Glc-(1\(\rightarrow\)4)-D-GlcNAc.

**Characterization of F290K and F290D 1G5A mutants**

F290D and F290K were produced in a larger amount and purified to homogeneity for further characterization. Glucosylation reactions were performed using \(\alpha\)-D-GlcNAc-OAll as acceptor with F290D and F290K. The distribution of the acceptor reaction products is shown in Figure 7A. Regarding the variant F290K and
48

F290D, both are highly specific for the glucosylation of α-D-GlcNAc-OAll. F290K and F290D yielded comparable amount of α-D-Glc-(1→4)-α-D-GlcNAc-OAll and are 10 times more efficient for α-D-GlcNAc-OAll glucosylation; 20 times more efficient to form α-D-Glc-(1→4)-D-GlcNAc-OAll. The best catalytic efficiency is however obtained with the F290K mutant, which was thus selected to carry out the production of α-D-Glc-(1→4)-D-GlcNAc-OAll.

Regarding the characterization of the products formed by F290K mutant, it was observed that addition of acceptor totally suppressed maltooligosaccharide formation (Figures 7B and 7C). Note that in the absence of acceptor, this mutant mainly catalyzed hydrolysis reaction but was still able to synthesize longer maltooligosaccharides (up to DP 20). Our assumption regarding the influence of the allyl aglycone in α-D-GlcNAc-OAU on the glucosylation rate and regiospecificity of glucosaminyl acceptors was thus valid.

Kinetic measurements showed that 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 (Figure 7D). In the presence of sucrose alone, $k_{cat}/K_m$ values decreased by 40 fold for F290K mutant. In the presence of the acceptor, it was observed for F290K mutant a formidable increase of catalytic efficiency towards both sucrose and the acceptor (Figures 7D and 7E). Such an improvement can be attributed to the high specificity of the mutant F290K for the α-D-Glc-(1→4)-D-GlcNAc-OAll formation. The adaptation of the acceptor binding site to the target acceptor enhanced the rate of the de-glucosylation step, which is no more the limiting step of the reaction. A remarkable 130 fold increase of the efficiency of the mutant compared to ASNPwt was observed and the expected regioselectivity was achieved.

**EXEMPLE 2: Enzymatic glucosylation of D-GlcNAc derivatives**

The mutant F290K and the ASNPwt were tested on other D-GlcNAc derivatives (α-D-GlcNAc-OMe, β-D-GlcNAc-OMe, β-D-GlcNAc-Oallyl and D-GlcNHTCA ) used as potential acceptors. The enzymatic reaction was carried out using sucrose as donor and D-GlcNAc derivative as acceptor in a 1:1 ratio (146mM).

The distribution of the acceptor reaction products is shown in Figure 10. Comparing to ASNPwt, the variant F290K is highly specific for the glucosylation of all the other D-
GlcNAc derivatives. (>95% for α-D-GlcpNAc-OMe, β-D-GlcpNAc-OAll and D-GlcNHTCA and 65-70% for β-D-GlcpNAc-OMe).

Efficiency of the glucosylation reaction was evaluated by HPLC analysis of the acceptor reaction product synthesized a Biorad HPLC Carbohydrate Analysis column (HPX-87K column (300x7.8 mm))

Percentages of glucosylation rates were estimated according to peak areas as no reference was available.

**EXAMPLE 3:** Chemo-enzymatic synthesis of potential building blocks to oligosaccharide fragments of *S. flexneri* 1b and/or 1a O-antigens

**General methods.** All moisture sensitive reactions were carried out under an atmosphere of argon in oven-dried glassware. Anhydrous solvents sold on molecular sieves were used as such. 4 Å powder molecular sieves was kept under vacuum and activated before use by heating at 250 °C under vacuum. 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). Signal assignments marked with a * are interchangeable assignments.

AUyI 2-acetamido-2-deoxy- α-D-glucopyranoside (XXi) (ref. 69). A mixture of 2-acetamido-2-deoxy-D-glucose (50 g, 226 mmol), allyl alcohol (400 mL), and BF3·OEt2 (7.5 mL, 61 mmol) was heated at 90 °C. After 24 h, the reaction mixture was concentrated under reduced pressure, and the resulty gummy residue was purified by column chromatography (9:1 CH2Cl2-MeOH) to obtain compound XX1 as a white solid (49.6 g, 84%). Compound XX1 had Rf = 0.5 (85:15 CH2Cl2-MeOH); 1H NMR (D2O, 400 MHz) δ (ppm): 5.87 (m, IH, CH=), 5.25 (m, IH, =CH2), 5.17 (m, IH, =CH2), 4.83 (d, IH, J1,2 = 3.6 Hz, H1), 4.13 (m, IH, OCH2), 3.94 (m, IH, OCH2), 3.82 (dd, IH, J2,3 = 10.7 Hz, H2), 3.78 (dd, IH, J5,6a = 2.2 Hz, J6a,6b = 12.0 Hz, H6a), 3.69 (dd, IH, H6b), 3.67 (dd, IH, J3,4 = 9.7 Hz, H3), 3.63 (ddd, IH, J4,5 = 9.9 Hz, J5,6b = 5.6 Hz, H5),
3.39 (t, IH, H4), 1.94 (s, 3H, COCH3); 13C NMR (D2O, 100 MHz) δ (ppm): 174.5 (C=O), 133.7 (CH2), 117.9 (=CH2), 96.2 (Cl), 72.0 (C5), 71.1 (C3), 70.1 (C4), 68.5 (OCH3), 60.6 (C6), 53.7 (C2), 21.9 (COCH3). HRMS (ESI+) of C11H16NO6Na ([M+Na]+, 284.1110) m/z 284.1112.

5

AHyI α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy- α-D-glucopyranoside (XX2). The crude mixture issued from the enzymatic glucosylation of XX1 (3.80 g, 14.54 mmol) was purified by silica gel column chromatography (9:1 CH3CN-H2O) to give a mixture of allyl α-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy- α-D-glucopyranoside (XX3) and fructose (4.16 g) and of the pure trisaccharide XX2 (2.53 g, 4.32 mmol) resulting from the enzyme-mediated transfer of glucose onto disaccharide XX3. Trisaccharide XX2 had RF = 0.3 (9:1 CH3CN-H2O). 1H NMR (D2O, 400 MHz) δ (ppm): 5.91 (m, IH, CH=), 5.35-5.32 (m, 2H, H1E, HIEO), 5.29 (m, IH, =CH2), 5.20 (m, IH, =CH2), 4.86 (d, IH, J1-2 = 3.5 Hz, H10), 4.15 (m, IH, OCH2), 4.01-3.93 (m, 2H, H3D, OCH2), 3.92-3.85 (m, 2H, H2D, 15

H3E/3.83-3.58 (m, 12H, H4D, H5D, H6aD, H6bD, H4E, H5E, H6aE, H6bE, H3E′, H5E′, H6aE′, H6bE′), 3.56 (dd overlapped, IH, J1-2 = 4.0 Hz, J2-3 = 10.0 Hz, H2E′), 3.52 (dd overlapped, IH, J1-2 = 3.8 Hz, J2-3 = 9.8 Hz, H2E′), 3.35 (pt, IH, J3-4 = J4-5 = 9.3 Hz, H4E′0, 1.97 (s, 3H, COCH3); 13C NMR (D2O, 100 MHz) δ (ppm): 174.5 (C=O), 133.7 (CH=), 118.1 C=CH2), 99.7 (C1E′, Jc-H i = 171.6 Hz), 99.9 (C1E′, Jc-Hi = 172.2 Hz), 96.0 (CID, Jc-Hi = 171.5 Hz), 77.7 (C4D), 76.9 (C4E), 73.4 (C3E), 72.9 (C3E′), 72.8 (C5E), 71.8, 71.6 (2C, C2E, C2F), 71.5 (C3D), 71.3 (C5D), 70.5 (C5D′), 69.4 (C4E), 68.6 (OCH2), 60.6 (3C, C6D, C6E, C6F0), 53.5 (C2D), 21.9 (COCH3). HRMS (ESI+) for C22H39NO16Na ([M+Na]+, 608.2167) m/z 608.2172.

Allyl 2,3,4,6-tetra-O-acetyl- α-D-glucopyranosyl-(1→4)-2-acetamido-3,6-di-O-acetyl-2-deoxy- α-D-glucopyranoside (XX4). Acetic anhydride (80 mL, 725 mmol) was added dropwise to a solution of XX3 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→4:6 then CH2Cl2/MeOH 92:2→95:5) to give XX4 (2.96 g, 30% from XX1) as a colourless amorphous solid. Compound XX4 had
RF = 0.6 (95:5 CH₂Cl₂-MeOH). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 5.90 (m, IH, CH=), 5.71 (d, IH, J = 9.7 Hz, NH), 5.45 (d, IH, J₁₂ = 4.0 Hz, H₁E), 5.35 (dd, IH, J₃⁻⁴ = 9.6 Hz, H₃E), 5.34-5.23 (m, 3H, =CH₂, H₃O), 5.04 (pt, IH, J₄⁻⁵ = 9.8 Hz, H₄E), 4.85 (dd, IH, J₂⁻₃ = 10.5 Hz, H₂E), 4.79 (d, IH, J₁₋₂ = 3.6 Hz, H₁D), 4.42 (dd, IH, J₅₋₆ₐ = 2.0 Hz, J₆a₋₇ = 12.0 Hz, H₆aD), 4.23-4.15 (m, 4H, OCH₂H₂D, H6bD, H₆aE), 4.04-3.91 (m, 5H, OCH₂H₄D, H₅D, H₅E, H₆bE), 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, 170.0, 169.4 (7C, C=O), 133.0 (CH=), 118.7 (=CH₂), 96.1 (C₁D), 95.5 (C₁E), 74.1 (C₃D), 72.4 (C₄D), 70.1 (C₂D), 69.4 (C₃E), 68.8 (OCH₂), 68.4 (C₅E), 67.9 (C₄E), 62.8 (C₆D), 61.4 (C₆E), 52.2 (C₂D), 23.1, 21.0, 20.8, 20.7, 20.6, 20.5 (7C, COCH₃). HRMS (ESI⁺) for C₉H₁₄NO₁₇Na ([M+Na]⁺, 698.2272) m/z 698.2277.

Allyl α-D-glucopyranosyl-(1 → 4)-2-acetamido-2-deoxy-α-D-glucopyranoside (XX₃). Method a: Methanolic sodium methoxide (0.5 M solution, 8 mL, 4 mmol) was added to a stirred solution of XX₄ (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 (RF= 0.3 in 9:1 CH₃CN-H₂O). Excess base was neutralized with Dowex-H⁺ resin. The suspension was filtered, and the filtrate was concentrated under reduced pressure. The residue was purified by silica gel chromatography, eluting with 10% H₂O in CH₃CN to afford compound XX₃ as a brownish foamy solid (1.65 g, 98%); ¹H NMR (D₂O, 400 MHz) δ (ppm): 5.95 (m, IH, CH=), 5.38 (d, IH, J₁₋₂ = 3.9 Hz, H₁D), 5.34 (m, IH, =CH₂), 5.25 (m, IH, =CH₂), 4.91 (d, IH, J₁₋₂ = 3.5 Hz, H₁D), 4.20 (m, IH, OCH₂), 4.02 (m overlapped, IH, OCH₂), 4.01 (dd overlapped, IH, J₃₋₄ = 8.4 Hz, H₃D), 3.94 (dd, IH, J₂₋₃ = 10.8 Hz, H₂D), 3.89-3.67 (m, 7H, H₄D, H₅D, H₆aD, H₆bD, H₅E, H₆aE, H₆bE), 3.66 (pt, IH, H₃E), 3.56 (dd, IH, J₂₋₃ = 9.9 Hz, H₂E), 3.40 (pt, IH, J₃₋₄ = J₄₋₅ = 9.5 Hz, H₄E), 2.02 (s, 3H, COCH₃); ¹³C NMR (D₂O, 100 MHz) δ (ppm): 174.5 (C-O), 133.7 (CH=), 118.1 (=CH₂), 99.9 (C₁E, JC₁Hi = 171.3 Hz), 96.0 (C₁D, Jc-iHi = 171.4 Hz), 77.6 (C₄D), 72.9 (C₃E), 72.8 (C₅E), 71.8 (C₂E), 71.5 (C₃D), 70.6 (C₅D), 69.4 (C₄D), 68.6 (OCH₂), 60.6 (2C, C₆D, C₆E), 53.5 (C₂D), 21.9 (COCH₃). HRMS (ESI⁺) for C₁₇H₃₀NO₁₇ ([M+H]⁺, 424.1819) m/z 424.1821.

Method b: Trisaccharide XX₂ (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 with 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 XX₃ and d-glucose (1.43 g, 62%) as seen by NMR.

**Method c:** To a solution of the crude mixture issued from the enzymatic glucosylation of XX₁ (6.43 g, 24.61 mmol) in anhydrous pyridine (170 mL) was added dropwise anhydride acetic (150 mL, 1.59 mol). 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 then CH₂Cl₂-MeOH, 97:3) to give a mixture (12.88 g) of peracetylated starting material, intermediate XX₄ and peracetylated trisaccharide resulting from the enzyme-mediated transfer of glucose onto disaccharide XX₂. To a stirred solution of the above mixture (12.88 g) in anhydrous methanol (140 mL) was added sodium methoxide (25% wt solution, 3.2 mL, 11.4 mmol). The reaction mixture was stirred for 23 h at room temperature by which time all the starting material had been consumed (Rf= 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 XX₃ as a brownish foamy solid (4.44 g, 43%).

**EXAMPLE 4:** Chemo-enzymatic synthesis of potential acceptor building blocks to oligosaccharide fragments of S. flexneri 1b and/or 1a O-antigens

\[ \text{AUyI} \rightarrow \alpha-D\text{-glucopyranosyl}(l\rightarrow4)-2\text{-amino-2-deoxy-} \ \alpha-D\text{-glucopyranoside} \ (XX₃). \]

**Method a:** Disaccharide XX₃ (784 mg, 1.85 mmol) was treated with Ba(OH)₂·8 H₂O (10 g) in water (50 mL), and the mixture was stirred at 90 °C overnight. When N-deacetylation was completed (Rf = 0.5 in 4:1:0.5 zPrOH-H₂O-conc. NH₄OH), the cooled mixture was saturated with CO₂, the volume was diminished to ~5 mL, EtOH (30 mL) was added, and solids were sedimented by centrifugation (0 °C, 666 x g, 2000 r.p.m., 10 min.). The supernatant was decanted, the residue was re-extracted with EtOH, and supernatants were pooled. The residue remaining after solvent removal was taken up in 13:7:2 CH₂Cl₂-MeOH-COnC. NH₄OH, and passed through a layer (8 x 3) of silica
gel wet with the same solvent. Pooled fractions containing the product were 
concentrated and freeze-dried to give XX₅ as a solid (587 mg; 83%). ¹H NMR (D₂O, 400 MHz) δ (ppm): 5.88 (m, IH, CH-), 5.30 (d overlapped, IH, H1ₐ), 5.27 (m, IH, =CH₂), 5.18 (m, IH, =CH₂), 4.97 (d, IH, J₁₋₂ = 3.7 Hz, H1₋₂), 4.15 (m, IH, OCH₂), 3.98 (m, IH, OCH₂), 3.93 (dd, IH, J₂₋₃ = 10.5 Hz, J₃₋₄ = 8.8 Hz, H3₋₄), 3.80-3.62 (m, 5H, H₅₋₇, H₆a₋₇, H₆b₋₇), 3.62-3.57 (m, 2H, H₄₋₅, H₅₋₆), 3.57 (dd overlapped, IH, J₃₋₄ = 9.1 Hz, H₃₋₄), 3.48 (dd, IH, J₁₋₂ = 4.0 Hz, J₂₋₃ = 9.9 Hz, H₂₋₃), 3.31 (pt, IH, J₄₋₅ = 9.6 Hz, H₄₋₅), 3.03 (dd, IH, H₂₋₃); ¹³C NMR (D₂O, 100 MHz) δ (ppm): 133.4 (CH═), 118.5 C=CH₂, 99.7 (C₁₋₇), 95.8 (C₁₋₇), 76.7 (C₄₋₅), 72.9 (C₅₋₆), 72.8 (C₃₋₄), 72.1 (C₃₋₄), 71.7 (C₂₋₄), 70.7 (C₅₋₆), 69.4 (C₄₋₅), 68.8 (OCH₂), 60.5, 60.4 (2C, C₆₋₇, C₆₋₇), 54.2 (C₂₋₃).

HRMS (ESI⁺) for C₁₅H₂₈NO₁₀ [(M+H)⁺, 382.1713] m/z 382.1719.

Method b: One-pot removal of O-acetyl and N-acetamide groups in 
disaccharide XX₄ followed the above mentioned procedure. Briefly, XX₄ (5.05 g, 7.47 mmol) was treated overnight with Ba(OH)₂·8 H₂O (25 g) in water (50 mL) at 90 °C. After neutralization with dry ice and repeated centrifugations (0 °C, 666 x g, 2000 r.p.m., 10 min.), compound XX₅ was purified by silica gel chromatography eluting with 13:7:2 CH₂Cl₂-MeOH-conc. NH₄OH (2.39 g, 84%).

AUyI α-D-glucopyranosyl-(1 →4)-2-amino-2-JV,3-0-carbonyl-2-deoxy-α-D-glucopyranoside (XX₆). An ice-cooled solution of α-nitrophenoxycarbonyl chloride (2.50 g, 12.6 mmol) in acetonitrile (25 mL) was added over several minutes to a stirred, ice bath-cooled, mixture of XX₅ (1.60 g, 4.2 mmol) and methanolic NaOMe (25% w/w solution, 3.5 mL, 12.6 mmol) in 50 mL methanol. The mixture was vigorously stirred with ice-bath cooling for 30 min, then at room temperature for 1 h. Water (100 mL) was added and the resulting aqueous mixture was extracted with diethyl ether. The aqueous layer was made acidic (pH ~3) by dropwise addition of 10% hydrochloric acid solution. The acidified aqueous phase was extracted with diethyl ether to remove α?-nitrophenol until the etheral layer failed to turn yellow upon NaHCO₃ addition. The aqueous solution was then neutralized with Amberlyst-A26 (OH⁻ form) resin, filtered, and evaporated to give crude oxazolidinone XX₆ (1.70 g). ¹H NMR (CD₃OD, 400 MHz) δ (ppm): 5.98 (m, IH, CH=), 5.37 (m, IH, =CH₂), 5.29 (d, IH, J₁₋₂ = 3.8 Hz, H₁₋₂), 5.22 (m, IH, =CH₂), 5.17 (d, IH, J₁₋₂ = 2.9 Hz, H₁₋₂), 4.78 (dd, IH, J₂₋₃ = 12.0 Hz, J₃₋₄ = 9.8 Hz, H₃₋₄), 4.30 (m, IH, OCH₂), 4.29 (pt overlapped, IH, H₄₋₅).
4.13 (m, IH, OCH₂), 3.86 (dd, IH, J₅-6a = 2.3 Hz, J₆a-6b = 11.8 Hz, H₆a₁), 3.83 (dd, 2H, J₁₂ = 3.1 Hz, H₆a₁, H₆b₁), 3.77 (dt, IH, J₄-5 = 9.3 Hz, H₅₁), 3.72 (dd, IH, J₅-6b = 5.4 Hz, H₆b₁), 3.68 (dd, IH, H₂D), 3.65-3.59 (m overlapped, 2H, H₃E, H₅₁), 3.46 (dd, IH, J₂-₃ = 9.8 Hz, H₂₁), 3.35 (dd, IH, J₃-₄ = 9.0 Hz, J₄-₅ = 9.8 Hz, H₄₁); ¹³C NMR (CD₃OD, 100 MHz) δ (ppm): 160.6 (C=O), 133.7 (CH=), 116.3 (=CH₂), 97.4 (Cl₁), 94.8 (Cl₀), 79.5 (C₃₁), 73.3 (C₃₂), 73.1 (C₂₁), 71.8 (C₂₂), 71.0 (C₄₁), 70.2 (C₄₂), 68.5 (OCH₂₁), 61.1 (C₆₁), 60.2 (C₆₂), 58.4 (C₅₁). HRMS (ESI⁺) for C₁₆H₂₅NO₁Na ([M+Na]+, 430.1325) m/z 430.1331.

Allyl 2,4,6-tri-(?)-benzyl- α-D-glucopyranosyl-(1 →4)-6-0-benzyl-2-benzylamino-2-iV,3,0-carbonyl-2-deoxy-a-D-glucopyranoside (XX₆). 60% NaH in oil (57 mg, 1.42 mmol) was added portionwise to an ice-cold mixture of the crude oxazolidinone XX₆ (108 mg) and benzyl bromide (0.19 mL, 1.60 mmol) in DMF (5 mL). After stirring for 30 min at 0 °C, the reaction mixture was warmed up and stirred for an additional 5.5 h at room temperature. The reaction was quenched by addition of Et₃N, diluted with EtOAc, and the mixture was washed with brine, dried (Na₂SO₄), filtered, and concentrated. The residue was purified by flash-chromatography (4:1 Cyclohexane-EtOAc) to afford XXg (48 mg, 19% over 2 steps) and the corresponding 3-hydroxyl compound XX₇ (39 mg, 17% over two steps); Rf 0.2 (3:1 Cyclohexane-EtOAc). IH NMR (CDCl₃, 400 MHz) δ (ppm): 7.45-7.22 (m, 25H, H arom.), 5.76 (m, IH, CH=), 5.54 (d, IH, J₁₂ = 3.6 Hz, H₁₁), 5.30-5.20 (m, 2H, =CH₂), 4.86 (d overlapped, IH, CH₂Ph), 4.84 (d overlapped, IH, CH₂Ph), 4.76 (d, IH, J₁₂ = 2.8 Hz, H₁₁), 4.65 (d overlapped, IH, H₃D), 4.64 (d overlapped, IH, CH₂Ph), 4.54 (d overlapped, IH, CH₂Ph), 4.52 (d overlapped, IH, CH₂Ph), 4.48 (m, 2H, CH₂Ph), 4.41 (m, 2H, CH₂Ph), 4.33 (d overlapped, IH, CH₂Ph), 4.30 (pt overlapped, IH, J₃₄ = J₄₅ = 9.6 Hz, H₄p₁), 4.02 (m overlapped, IH, OCH₂), 4.00 (pt overlapped, IH, H₃₁), 3.81 (ddd, IH, H₅₁), 3.76 (dd, IH, J₅-₆a = 3.7 Hz, J₆a-₆b = 11.0 Hz, H₆a₁), 3.70-3.57 (m, 4H, OCH₂₁, H₆b₁, H₄E₁, H₅₁), 3.54 (dd, IH, J₅-₆b = 2.4 Hz, J₆a-₆b = 10.7 Hz, H₆a₁), 3.65 (dd IH, J₂-₃ = 9.7 Hz, H₂₁), 3.41 (d overlapped, IH, J₅-₆b = 1.2 Hz, H₆b₁), 3.34 (dd, IH, J₂-₃ = 11.8 Hz, H₂₁); ¹³C NMR (CDCl₃, 100 MHz) δ (ppm): 158.6 (C=O), 138.6, 138.0, 137.8, 137.6, 135.3 (5C, C quat. arom.), 133.1 (CH=), 128.8, 128.7, 128.6, 128.5, 128.3, 128.2, 128.1, 127.9, 127.8, 127.6, 127.3 (25C, C arom.), 118.4 (=CH₂), 94.8 (Cl₁), 94.3 (Cl₂), 78.7 (C₂₁), 77.2 (C₄₁), 76.9 (C₃₁), 74.6, 73.5 (3C, OCH₂Ph), 73.2 (C₃₂), 72.5
55
(OCH₂Ph), 72.1 (C₅D), 71.0 (C₄D), 70.9 (C₅E), 69.1 (OCH₂), 68.3 (C₆D), 68.1 (C₆E),
61.1 (C₂D), 48.7 (NCH₂Ph). HRMS (ESI⁺) for C₅H₅NO₁⁺Na ([M+Naf, 880.3673] m/z
880.3704.

AUyI 2,3,4,6-tetra-0-benzyl- ⍺-D-glucopyranosyl-(1 →4)-6-0-
5
benzyl-l-benzylamino-l-IV ⍺-D-glucopyranoside (XXs)-
Methanolic sodium methoxide (0.5 M solution, 1.6 mL, 0.79 mmol) and
trichloroethylchloroformate (225 µL, 1.56 mmol) were added to a solution of compound
XX₅ (100 mg, 0.26 mmol) in methanol (3.4 mL) at 0 °C. After 4 h, the reaction mixture
was neutralized by addition of Dowex-H⁺ resin, filtered, and concentrated to dryness to
afford the crude trichloroethyl carbamate. This material (Rf= 0.7, 85:15 CH₃CN-H₂O)
was used directly without further purification.

NaH (60% in oil, 63 mg, 2.60 mmol) was added to an ice-cold mixture
of the crude trichloroethyl carbamate and benzyl bromide (0.31 mL, 2.60 mmol) in
DMF (5 mL). The reaction mixture was stirred for 30 min at 0 °C, then stirred for an
additional 4.5 h at room temperature. The reaction was quenched by addition of Et₃N,
diluted with EtOAc, poured into saturated aqueous NH₄Cl, and extracted with EtOAc.
The combined organic extracts were washed with water and brine, dried (Na₂SO₄),
filtered, and concentrated. The residue was purified by flash-chromatography (17:3
Cyclohexane-EtOAc) to afford XX₈ (81 mg, 34% over 2 steps); Rf = 0.3 (4:1
Cyclohexane-EtOAc). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 7.46-7.16 (m, 30H, H
arom.), 5.79 (m, IH, CH=), 5.56 (d, IH, J₁₂ = 3.5 Hz, H₁ₑ), 5.26 (m, 2H, =CH₂), 5.02
(d overlapped, IH, CH₂Ph), 4.87-4.80 (m, 3H, CH₂Ph), 4.79-4.72 (m overlapped, 3H,
H₁D, H₃D₂, CH₂Ph), 4.56-4.49 (m, 4H, CH₂Ph), 4.47 (m, 2H, CH₂Ph), 4.35 (pt
overlapped, IH, H₄D), 4.34 (d overlapped, IH, CH₂Ph), 4.05 (m, IH, OCH₂), 3.92-3.88
(m, 2H, H₅D, H₃E), 3.81 (dd, IH, J₅₋₆a = 3.5 Hz, J₆a₋₆b = 11.0 Hz, H₆aD), 3.73-3.66 (m,
4H, OCH₂, H₆bD, H₄E, H₅E), 3.65 (dd overlapped, IH, H₂E), 3.57 (dd, IH, J₅₋₆a = 2.7
Hz, J₆a₋₆b = 10.8 Hz, H₆aD), 3.43 (dd, IH, J₅₋₆b = 1.4 Hz, H₆bE), 3.37 (dd, IH, J₁₋₂ = 2.7
Hz, J₁₋₂ = 11.8 Hz, H₂D); ¹³C NMR (CDCl₃, 100 MHz) δ (ppm): 158.7 (C=O), 138.9,
138.5, 138.1, 137.9, 137.7, 135.3 (6C, C quat. arom.), 133.1 (CH=), 128.8, 128.7, 128.5,
128.4, 128.3, 128.2, 127.9, 127.8, 127.6, 127.5, 127.3 (3OC, C arom.), 118.3 (=CH₂),
95.4 (CIB), 94.4 (C₁D), 81.8 (C₃E), 79.5 (C₂D), 77.4 (C₄E), 76.9 (C₃D), 75.6, 75.1, 73.5,
73.1 (5C, OCH₂Ph), 72.1 (C₅D), 71.3 (C₅E), 70.9 (C₄D), 69.1 (OCH₂), 68.4 (C₆D), 68.2
(C6E)5 61.2 (C2D), 48.7 (NCH2Ph). HRMS (ESI+) for C58H46FNO11Na [(M-HNa)+, 970.4142] m/z 970.4107.

AUyI 2,3,4,6-tetra-0-benzyl- α-D-glucopyranosyl-(1→4)-6-0-benzyl-2-amino-2- N3-O-carbonyl-2-deoxy-a-D-glucopyranoside (XX 9) A.

The crude oxazolidinone XX6 (500 mg, 1.23 mmol) in dry DMF (10 mL) at 0°C. After 1 h, the mixture was treated with benzyl bromide (800 µL, 6.75 mmol), stirred overnight, treated with MeOH, and concentration to dryness. The residue was purified by flash-chromatography (85:15 Toluene-EtOAc) to afford XX8 (74 mg, 6% over 2 steps), and the corresponding penta-O-benzyl derivative XX8a (158 mg, 15%); Rf = 0.3 (85:15 Toluene-EtOAc). 1H NMR (CDCl3, 400 MHz) δ (ppm): 3.75-7.17 (m, 25H, H arom.), 5.80 (m, IH, CH=), 5.57 (d, IH, J1-2 = 3.5 Hz, H1E), 5.21 (m, IH, =CH2), 5.13 (m, 2H, =CH2), 4.95 (d, IH, CH2Ph), 4.86 (d overlapped, IH, CH2Ph), 4.83 (d overlapped, IH, CH2Ph), 4.78 (d overlapped, IH, CH2Ph), 4.77 (d, IH, J1-2 = 3.6 Hz, H1p), 4.71-4.49 (m, 5H, CH2Ph), 4.37 (d overlapped, IH, CH2Ph), 4.29 (ddd, IH, J2-3 = 10.1 Hz, H2D), 4.19 (pt, IH, J3-4 = J4-5 = 9.2 Hz, H4D), 4.06 (m, IH, OCH2), 3.97 (pt overlapped, IH, H3E), 3.95 (dd overlapped, IH, H6aD), 3.91-3.81 (m, 3H, H3D, H5D, H5E), 3.76-3.64 (m, 3H, OCH2, H6bD, H4E), 3.57 (dd, IH, J5-6a = 3.2 Hz, J6a-6b = 10.8 Hz, H6aE), 3.52 (dd, IH, J2-3 = 9.8 Hz, H2E), 3.48 (dd, IH, J5-6b = 1.4 Hz, H6bE); 13C NMR (CDCl3, 100 MHz) δ (ppm): 156.4 (C=O), 138.8, 138.6, 138.4, 138.1, 135.3 (5C, C quat. arom.), 133.8 (CH=), 128.3, 128.1, 128.0, 127.9, 127.8, 127.6, 127.5, 127.4 (25C, C arom.), 117.6 C=CH2, 97.4 (C1D), 96.7 (C1E), 82.0 (C3E), 80.5 (C3D), 79.5 (C2E), 77.7 (C4E), 75.6, 75.0, 73.2, 72.9 (4C, OCH2Ph), 72.6 (C4D), 71.7 (OCH2Ph), 71.1 (C5D), 70.7 (C5D), 69.0 (C6D), 68.4 (C6E), 68.4 (OCH2), 52.8 (C2p); HRMS (ESI+) for C51H49F5N11O4Na [(M+Na)+, 880.3673] m/z 880.3710.

AHyI 2,3,4,6-tetra-O-benzyl- α-D-glucopyranosyl-(1→4)-6-O-benzyl-2-benzylamino-2-deoxy- α-D-glucopyranoside (XX9). Perbenzylated XX8 (36 mg, 0.04 mmol) was treated with 1 M aqueous NaOH/1,4-dioxane (1:1, v/v, 10 mL) at 65 °C for 2 days. The mixture was diluted with EtOAc and washed with water. The separated aqueous layer was extracted with EtOAc. The combined organic extracts were washed with water and brine, dried (Na2SO4), filtered and concentrated. The residue was purified by flash-chromatography (4:1 Cyclohexane-EtOAc) to afford target XX9.
(32 mg, 91%) as an oil; Rf = 0.1 (4:1 Cyclohexane-EtOAc). 1H NMR (CDCl₃, 400 MHz) δ (ppm): 7.41-7.16 (m, 3OH, H arom.), 5.95 (m, IH, CH-), 5.33 (m, IH, =CH₂), 5.24 (d overlapped, IH, J₁₂ = 3.4 Hz, H₁₂), 5.23 (m, IH, =CH₂), 4.96 (d, IH, CH₂Ph), 4.90-4.76 (m, 4H, CH₂Ph, H₁₆), 4.57 (d overlapped, IH, CH₂Ph), 4.55 (d overlapped, IH, CH₂Ph), 4.50 (d overlapped, IH, CH₂Ph), 4.46 (d overlapped, IH, CH₂Ph), 4.40 (d overlapped, IH, CH₂Ph), 4.16 (m, IH, OCH₂), 4.00 (pt overlapped, IH, H₃₄), 3.99-3.95 (m, 4H, OCH₂, CH₂Ph, H₃₅), 3.85-3.77 (m, 2H, H₅₄, H₅₅), 3.76-3.63 (m, 4H, H₄₅, H₆₆), 3.62-3.56 (m, 2H, H₂₄, H₆₆), 3.47 (dd, IH, J₅₆ = 1.8 Hz, J₆₅ = 10.6 Hz, H₆₅), 2.80 (dd, IH, J₁₂ = 3.5 Hz, J₂₃ = 10.2 Hz, H₂₃); 13C NMR (CDCl₃, 100 MHz) δ (ppm): 138.7, 138.4, 138.3, 137.9, 135.4 (6C, C quat. arom.), 134.0 (CH=), 128.6, 128.5, 128.4, 128.3, 128.1, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 127.1 (30C, C arom.), 117.7 (=CH₂), 99.4 (C₁₄), 96.3 (C₁₃), 82.1 (C₃₉), 80.3 (C₄₅), 79.7 (C₂₂), 77.7 (C₁₄), 75.7, 75.0, 73.7, 73.5, 73.2 (5C, OCH₂Ph), 73.2 (C₃₉), 71.2 (C₅₉), 69.5 (C₅₉), 69.1 (C₅₉), 68.5 (C₅₉), 68.4 (OCH₂), 61.1 (C₂₅), 51.9 (NCH₂Ph). HRMS (ESI⁺) for C₅₇H₆₄NOI₀ ([M+H]+, 922.4530) m/z 922.4540.

AHyI  2,3,4,6-tetra-0-benzyl- α-D-glucopyranosyl-(1 →4)-6-0-benzyl-2-benzylacetamido-2-deoxy- α-D-glucopyranoside (XXio).

Method a: To a solution of disaccharide XX₉ (424 mg, 0.46 mmol) in anhydrous pyridine (5 mL) was added dropwise anhydride acetic (250 µL, 2.65 mmol). The resulting mixture was stirred at room temperature under argon. After 24 h, 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, 2:1) to give XX₁₀ (391 mg, 88%) as a colourless oil. Compound XXio had Rf = 0.2 (2:1 Cyclohexane-EtOAc). NMR analysis indicated an equilibrium between two rotamers.

Method b: Perbenzylated XX₈ (125 mg, 0.13 mmol) was treated with 1 M aqueous NaOH/1,4-dioxane (1:1, v/v, 15 mL) at 65 °C for 2 days. The mixture was diluted with EtOAc and washed with water. The separated aqueous layer was extracted with EtOAc. The combined organic extracts were washed with water and brine, dried (Na₂SO₄), filtered, and concentrated to give crude XX₉ (109 mg, 0.12 mmol) as an oil.
Anhydride acetic (56 µL, 0.59 mmol) was added to a solution of crude disaccharide \(XX_9\) (109 mg, 0.12 mmol) in anhydrous pyridine (1 mL). The mixture was stirred at room temperature under argon. After 5 h, TLC showed the complete disappearance of the starting materials. Volatiles were eliminated by repeated coevaporation with toluene under reduced pressure. The residue was purified by column chromatography (Cyclohexane-EtOAc, 2:1) to give \(XX_1\)O (98 mg, 77%) as a colourless oil. \(^1\)H NMR (CDCl\(_3\), 400 MHz) \(\delta\) (ppm): 7.37-7.15 (m, 30H, H arom.), 5.50-5.40 (m, IH, CH=), 5.10-4.76 (m, 7H, H\(1_D\), H\(1_E\), H\(2_D\), =CH\(_2\), OCH\(_2\)Ph), 4.75-4.65 (m, 3H, OCH\(_2\)Ph, NCH\(_2\)Ph), 4.60-4.38 (m, 5H, OCH\(_2\)Ph), 4.27-4.12 (m, 0.9H, H\(3_D\), OCH\(_2\)).

4.04-3.91 (m, 2.1H, H\(3_{D'}\), OCH\(_2\)), 3.91-3.81 (m, IH, H\(5_E\)), 3.80-3.50 (m, 5H, H\(2_{E'}\), H\(4_{D'}\), H\(4_E\), H\(6a_{D'}\), H\(6b_{D'}\), H\(6a_{E}\)), 3.50-3.38 (m, IH, H\(6b_E\)), 2.31 (s, 0.9H, COCH\(_3\)), 1.99 (s, 2.1H, COCH\(_3\)). \(^{13}\)C NMR (CDCl\(_3\), 100 MHz) \(\delta\) (ppm): 173.9, 172.5 (C=O), 139.8, 139.4, 138.5, 138.2, 137.8, 136.9 (C quat. arom.), 133.5 (CH=), 128.8, 128.7, 128.6, 128.4, 128.3, 127.9, 127.7, 127.6, 127.5, 127.4, 126.6, 126.4, 125.9 (C arom.).

117.8, 117.3 (=CH\(_2\)), 100.8 (C\(1_E\)), 98.0, 97.7 (C\(1_p\)), 83.8 (C\(4_D\)), 82.2 (C\(3_{E'}\)), 79.4 (C\(2_{E'}\)), 77.7 (C\(4_E\)), 75.7, 75.6, 75.0, 74.2, 73.5, 73.3, 73.1 (OCH\(_2\)Ph), 71.3 (C\(5_E\)), 70.0 (C\(5_{D'}\)), 69.1 (C\(6_{D'}\)), 69.0 (C\(3_{D'}\)), 68.9 (OCH\(_2\)), 68.3 (C\(6_E\)), 56.4 (C\(2_{D'}\)), 48.5 (NCH\(_2\)Ph), 22.6 (COCH\(_3\)). HRMS (ESI\(^+\)) for C\(_{59}\)H\(_{88}\)ClNO\(_{17}\)Na ([M+Na]\(^+\), 964.4636) \(m/z\) 964.4813.

AHyI 2-O-benzoyl-4-(-benzyl-3-O-chloroacetyl- \(\alpha-L\)-rhamnopyranosyl-(1 →3)-[2,3,4,6-tetra-O-benzyl- \(\alpha-D\)-glucopyranosyl-(1 →4)]-6-(-benzyl-2-benzylacetamido-2-deoxy- \(\alpha-D\)-glucopyranoside (XXn). TMSOTf (5.5 µL, 0.030 mmol) was added to a solution of known 2-0-benzoyl-4-0-benzyl-3-O-chloroacetyl-\(\alpha-L\)-rhamnopyranosyl trichloroacetimidate (ref. 73) (73 mg, 0.126 mmol) and acceptor \(XX_1\)O (98 mg, 0.102 mmol) in toluene (2.75 mL) containing activated MS4A under argon at -10 °C. The mixture was stirred at room temperature and, when TLC monitoring indicated complete consumption of the donor, the reaction was quenched by addition of triethylamine. After filtration through a bed of celite and concentration under vacuum, flash chromatography (85:15 Toluene-EtOAc) gave a mixture of two products (RF = 0.4, 4:1 Toluene-EtOAc), among which the target trisaccharide \(XX_{11}\) as indicated by mass spectrometry analysis. Compound \(XX_{11}\) had HRMS (ESI\(^+\)) for C\(_{81}\)H\(_{86}\)ClNO\(_{17}\)Na ([M+Na]\(^+\), 1402.5482) \(m/z\) 1402.5525.
Ahyl 2-0-acetyl-3,4-di-0-benzyl- α-L-rhamnopyranosyl-(1 →3)-
[2,3,4,6-tetra-0-benzyl- α-D-glucopyranosyl-(1 →4)]-6-0-benzyl-2-benzylacetamido-
2-deoxy- α-D-glucopyranoside (XX12). TMSOTf (4 µL, 0.026 mmol) was added to a
solution of known 2-O-acetyl-3,4-di-O-benzyl- α-L-rhamnopyranosyl
trichloroacetimidate (ref. 74) (132 mg, 0.249 mmol) and acceptor XX10 (76 mg, 0.079
mmol) in toluene (4 mL) containing activated MS4A, stirred under argon at 0 °C.
Following additional stirring for 3 h at 60 °C, TLC monitoring indicated complete
consumption of the donor. The reaction was quenched by addition of triethylamine.
After filtration through a bed of celite and concentration under vacuum, flash
chromatography (9:1 Toluene-EtOAc) gave the expected trisaccharide XX12 (Rf= 0.4,
4:1 Toluene-EtOAc). 1H NMR (CDCl3, 400 MHz) δ (ppm): 7.28-7.12 (m, 41H, NH, H
arom.), 5.61 (d, IH, J1-2 = 3.6 Hz, H1E), 5.22 (m, IH, H2p), 5.14 (pt, IH, J1-2 = J2-3 =
2.5 Hz, H2c), 5.08-4.97 (m, IH, CH=), 4.83 (d overlapped, IH, H1c), 4.83-4.57 (m,
HH, H1d), 2 =CH2, 6 OCH2Ph, 2 NCH2Ph), 4.50-4.32 (m, 7H, 7 OCH2Ph), 4.30-4.20
(2H, H3d, OCH2Ph) 4.09 (dd, IH, J3-4 = 7.2 Hz, J4-5 = 9.0 Hz, H4D), 4.01-3.90 (m,
2H, H5c, OCH2), 3.85 (dd overlapped, IH, H3c), 3.83-3.60 (m, 6H, H5d, H3e, H4e,
H5e, H6a, OCH2), 3.56-3.42 (m, 2H, H6b, H6a), 3.40 (dd, IH, J2-3 = 9.6 Hz, H2e),
3.33 (pt, IH, J3-4 = J4-5 = 9.3 Hz, J5-6 = 5.7 Hz, J6a-6b = 11.8 Hz,
H6b), 2.04, 1.98 (2s, 6H5 COCH3), 1.20 (d, 3H, J5-6 = 6.1 Hz, H6c); 13C NMR (CDCl3,
100 MHz) δ (ppm): 173.8, 170.4 (C=O)5 139.0, 138.9, 138.7, 138.6, 138.4, 138.3 (C
quat. arom.), 133.1 (CH=), 128.5, 128.4, 128.3, 128.2, 128.1, 127.9, 127.8, 127.7,
127.6, 127.4, 127.3, 126.6, 125.3 (C arom.), 117.5 (=CH2), 98.1 (Cl), 97.5 (Cl), 94.2
(Cl1) 81.8 (Cl2), 80.2 (Cl2), 79.8 (Cl4), 79.7 (Cl3), 77.7 (Cl4), 77.6 (Cl4), 75.4,
75.0, 74.8, 73.4 (4C, OCH2Ph), 73.2 (Cl4), 73.2, 73.1, 71.6 (3C, OCH2Ph), 71.3
25 (Cl5, 70.2 (Cl2), 70.1 (Cl5), 69.4 (Cl6), 68.8 (OCH2), 68.7 (Cl5), 68.6 (Cl6), 54.9
(C1), 48.3 (NCH2Ph), 22.7, 21.1 (2C, COCH3), 18.1 (Cl6). HRMS (ESI+) for
C81H89NO16Na ([M+Na]+) 1354.6079 (m/z 1354.5994).
EXAMPLE 5: Chemo-enzymatic synthesis of potential donor building blocks to
oligosaccharide fragments of S. flexneri Ib and/or Ia O-antigens.
Allyl α-D-glucopyranosyl-(1 →4)-2-deoxy-2-trichloroacetamido-
α-D-glucopyranoside (XX13). Methanolic sodium methoxide (25% wt solution in MeOH,
1.72 mL, 6.10 mmol) and trichloracetic anhydride (560 µL, 3.07 mmol) were added to a
solution of compound $XX_5$ (390 mg, 1.02 mmol) in methanol (8 mL) at 0 °C. After 1 h the reaction mixture was neutralized by addition of Dowex-H+ resin, filtered, and concentrated to dryness to obtain crude $XX_0$. This material was used directly without further purification. Compound $XX_{13}$ had $R_f = 0.7$ (85:15 $CH_3CN-H_2O$). HRMS (ESI+) of $C_{17}H_{26}Cl_3NO_8Na$ ([M+Naf, 548.0469] m/z 548.0524.

AIIyI 2,3,4,6-tetra-O-acetyl- $\alpha$-D-glucopyranosyl-(1 $\rightarrow$4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido- $\alpha$-D-glucopyranoside ($XX_{4}$). Acetic anhydride (5 mL, 52.9 mmol) was added dropwise to a solution of crude $XX_{n}$ (1.02 mmol) in anhydrous pyridine (5 mL) and the resulting mixture was stirred overnight at room temperature. The reaction was quenched at 0 °C by addition of methanol, and the mixture was evaporated to dryness. The residue was purified by flash-chromatography (3:1 Cyclohexane-Acetone) to yield to $XX_{14}$ (458 mg, 58% over 2 steps) as an oil. Compound $XX_{14}$ had $R_f = 0.5$ (3:1 Cyclohexane-Acetone). $^1$H NMR (CDCl$_3$, 400 MHz) δ (ppm): 6.93 (d, IH, $J = 9.1$ Hz, NH), 5.92 (m, IH, CH=), 5.51 (d, IH, $J_{i-2} = 4.0$ Hz, H1E), 5.47-5.27 (m, 4H, =CH$_2$, H3D, H3E), 5.09 (pt, IH, $J_{A-5} = 10.0$ Hz, H4E), 4.95 (d, IH, $J_{i-2} = 3.6$ Hz, H1D), 4.89 (dd, IH, $J_{2-3} = 10.5$ Hz, H2E), 4.48 (dd, IH, $J_{5-63} = 1.3$ Hz, J$_{6a-6b}$ = 12.0 Hz, H6aD), 4.31-4.23 (m, 3H, OCH$_2$, H6bD, H6aE), 4.14 (ddd overlapped, IH, $J_{2-3} = 10.8$ Hz, H2D), 4.11-4.05 (m, 4H, OCH$_2$, H4D, H5D, H6bE), 3.99 (pdt, IH, $J_{6a}$ = 3.1 Hz, $J_{5-6b} = 10.0$ Hz, H5E), 2.18, 2.13, 2.06, 2.05, 2.02 (6s, 18H, COCH$_3$); $^{13}$C NMR (CDCl$_3$, 100 MHz) δ (ppm): 171.4, 170.5, 169.2, 169.4 (6C, C=O), 161.9 (C=O), 132.7 (CH=), 119.2 (=CH$_2$), 95.6 (C1E), 95.2 (C1D), 92.0 (CC1$_3$), 73.1 (C3D), 72.3 (C4D), 70.1 (C2E), 69.4 (C3E), 69.1 (OCH$_2$), 68.6 (C5E), 68.3 (C5D), 68.0 (C4E), 62.7 (C6D), 61.4 (C6E), 54.4 (C2D), 21.0, 20.8, 20.7, 20.6, 20.5 (6C, COCH$_3$). HRMS (ESI+) of $C_{29}H_{38}Cl_3NO_7$Na ([M+Na]$^+$, 800.1 103) m/z 800.1 112.

2,3,4,6-tetra-O-acetyl- $\alpha$-D-glucopyranosyl-(1 $\rightarrow$4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido- $\alpha$-D-glucopyranose (XXis). A catalytic amount of 1,5-cyclooctadiene-bis[methylidiphenylphosphine]-iridium hexafluorophosphate (50 mg) was dissolved in dry THF (20 mL). The stirred solution was degassed, placed under hydrogen for 10 minutes until the orange colour turned yellow, degassed and placed under nitrogen. A solution of $XX_{14}$ (444 mg, 0.57 mmol) in THF/H$_2$O mixture (5:2, v/v, 7mL), was then poured into the solution of iridium complex. The reaction mixture was stirred for 3 h and I$_2$ (290 mg, 1.14 mmol) was added. After stirring for another 3 h, the
mixture was diluted with dichloromethane, washed with aqueous NaHSO₃ and water, dried and concentrated. The crude material was used as such. Compound XXishad Rf= 0.2 (3:2 Cyclohexane-EtOAc). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 7.03 (d, IH, J = 9.2 Hz, NH), 5.53 (d overlapped, 1H₃-1₈), 5.50 (dd overlapped, IH, J₂⁻₃ = 10.8 Hz, J₅-₆ = 9.0 Hz, H₃₈), 5.39 (pt, IH, J₂⁻₃ = J₅⁻₆ = 10.0 Hz, H₃₉), 5.31 (si, IH, H₁₈), 5.10 (pt, IH, J₄⁻₅ = 10.0 Hz, H₄₈), 4.89 (dd, IH, J₁⁻₂ = 4.0 Hz, H₂₈), 4.52 (dd, IH, J₅₋₆₈ = 3.7 Hz, J₆₈₋₆₉ = 13.4 Hz, H₆₈₋₆₉), 4.32-4.23 (m, 3H, H₅₈, H₆₈, H₆₉), 4.15-3.94 (m, 4H, H₂₈, H₄₈, H₅₈, H₆₈), 2.18, 2.13, 2.06, 2.05, 2.02 (6s, 18H, COCH₃). ¹³C NMR (CDCl₃, 100 MHz) δ (ppm): 171.4, 170.6, 170.0, 169.4 (6C, C=O), 162.0 (C=O), 95.5 (C₁₈), 92.0 (CCl₃), 90.6 (C₁₈), 72.9 (C₃₈), 72.3 (C₄₈), 69.9 (C₂₈), 69.4 (C₃₈), 68.5 (C₅₈), 68.0 (2C, C₅₈, C₄₈), 62.8 (C₆₈), 61.4 (C₆₈), 54.7 (C₂₈), 21.0, 20.8, 20.7, 20.6, 20.5 (6C, COCH₃). HRMS (ESI⁺) OIC 2₅H₃₄Cl₂N₄O₁₇N₈ ([M+Na]⁺, 760.0709) m/z 760.0804.

2,3,4,6-Tetra-O-acetyl- α-D-glucopyranosyl-(l →4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-a-D-glucopyranosyl trichloroacetimide (XX₁₈).

DBU (26 µL, 0.17 mmol) was added to a stirred solution of crude XX₁₈ (0.57 mmol) and CCl₃CN (70 µL, 0.68 mmol) in dichloroethane (6 mL) at -5 °C. After 3 h, the solution was concentrated. Column chromatography (2:1 Cyclohexane-EtOAc) of the residue yielded XX₁₆ (131 mg, 26%) as a syrup. Compound XX₁₆ had Rf = 0.4 (3:2 Cyclohexane-EtOAc). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 6.98 (d, IH, J = 8.7 Hz, NH), 6.46 (d, IH, J₁₋₂ = 3.6 Hz, H₁₈), 5.56 (d, IH, J₁₋₂ = 4.1 Hz, H₁₈), 5.51 (dd overlapped, IH, H₃₈), 5.41 (dd overlapped, IH, H₃₈), 5.10 (pt, IH, J₄₋₅ = 10.0 Hz, H₄₈), 4.89 (dd, IH, J₂₋₃ = 10.4 Hz, H₂₈), 4.50 (dd, IH, J₅₋₆₈ = 1.8 Hz, J₆₈₋₆₉ = 12.4 Hz, H₆₈₋₆₉), 4.38-4.24 (m, 4H, H₂₈, H₅₈, H₆₈, H₆₉), 4.20-3.94 (m, 3H, H₄₈, H₅₈, H₆₈), 2.18, 2.13, 2.06, 2.05, 2.02 (6s, 18H, COCH₃).

AUyl 2,3,4,6-tetra-O-acetyl- α-D-glucopyranosyl-(l →4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido- α-D-glucopyranosyl-(l →2)-3,4-di-O-benzyl- α-L-rhamnopyranoside (XX₁₇). TMSOTf (8 µL, 0.04 mmol) was added to a solution of the trichloroacetimidate donor XX₁₆ (131 mg, 0.15 mmol) and known allyl 3,4-di-O-benzyl- α-L-rhamnopyranoside (ref. 76) (86 mg, 0.22 mmol) in toluene (4 mL) containing activated MS4A under argon. The mixture was stirred at -60 °C for 60 min and, when TLC monitoring indicated complete consumption of the donor, the reaction was quenched by addition of triethylamine. After filtration through a bed of celite and
concentration under vacuum, flash chromatography (2:1 Cyclohexane-EtOAc) gave slightly contaminated trisaccharide XX17 (Rf = 0.2 in 2:1 Cyclohexane-EtOAc). 1H NMR (CDCl3, 400 MHz) δ (ppm): 7.41-7.28 (m, 10H, H arom.), 6.71 (d, IH, J = 8.7 Hz, NH), 5.91 (m, IH, CH3=), 5.45 (d, IH, J = 4.0 Hz, H1e), 5.43-5.37 (m, 2H, H4D, H3e), 5.27 (m, IH, =CH2), 5.19 (m, IH, =CH2), 5.11-5.07 (m, 2H, H3D, H4e), 4.89 (dd, IH, J = 10.1 Hz, H2e) 4.86 (d, IH, J = 1.4 Hz, H5HU) 4.78 (d, IH, OCH2Ph) 4.62 (d, IH, OCH2Ph) 4.59 (d, IH, OCH2Ph), 4.56 (d, IH, J = 8.5 Hz, H5H1d), 4.50 (m, IH, H6a), 4.31-4.19 (m, 2H, H6bD, H6aE 4.17-3.95 (m, 6H, 2 OCH2H2D, H5E, H6bE), 3.93 (dd, IH, J = 5.1 Hz, H2A), 3.89 (dd, IH, J = 9.2 Hz, H3A, 3.76-3.68 (m, 1H, H5A), 3.38 (pt, IH, J = 9.2 Hz, H4A) 2.17, 2.12, 2.10, 2.06, 2.05, 2.04 (6S, 18H, 6 COCH3), 1.30 (d, IH, J = 5.1 Hz, H6A); 13C NMR (CDCl3, 100 MHz) δ (ppm): 171.4, 170.5, 170.3, 170.3, 169.4, 161.8 (C5, C=O), 138.5, 138.3 (2C, C quart. arom.), 133.8 (CH=), 128.9, 128.5, 128.4, 128.2, 127.7, 127.6 (10C, C arom.), 117.2 (=CH2), 101.4 (Cl), 98.2 (C1A), 95.5 (C1E), 92.1 (CCl3), 81.0 (C4A), 80.0 (C3A), 76.8 (C2A), 75.6 (OCH2Ph), 75.0 (C3D), 73.6 (OCH2Ph), 73.2 (C4D), 70.0 (C2E), 69.4 (C3E), 68.5 (C5E), 68.0 (3C, C5A, C5D, C4E), 67.8 (OCH2), 62.5 (C6D), 61.5 (C6E), 55.9 (C2D), 22.5, 21.1, 21.0, 20.8, 20.7 (6C, COCH3), 17.9 (C6A). HRMS (ESI+) for C49H60C13NO22Na ([M+Na]+, 1126.2621) m/z 1126.2623.

EXAMPLE 6: Chemical synthesis of potential acceptor building blocks to oligosaccharide fragments of S. flexneri 1b and/or 1a O-antigens

AHyI 2-acetamido-3-0-acetyl-4,6-0-benzylidene-2-deoxy-α-D-glucopyranoside (ref. 69) (XX19). Benzaldehyde dimethacetal (2.6 mL, 17.55 mmol) and CSA (2.72 g, 11.70 mmol) were added to a solution of compound XX1 (1.53 g, 5.85 mmol) in acetonitrile (50 mL). After 45 min, the methanol formed during the reaction was removed under reduced pressure and another 2 mL of benzaldehyde dimethacetal were added to the mixture. After 60 min, the reaction mixture was cooled to 0 °C, neutralized by addition of triethylamine and concentrated to dryness to afford crude XX18 (ref. 75) (2.03 g) as a white solid (Rf = 0.6, 1:1 Cyclohexane-Acetone). This material was used directly without further purification.

Ac2O (20 mL, 211.6 mmol) was added dropwise to a solution of crude XX18 (2.03 g) in anhydrous pyridine (50 mL) and the resulting mixture was stirred for one night at room temperature. The reaction was then quenched at 0 °C by addition of
methanol, and the mixture was evaporated to dryness. The residue was purified by flash-chromatography (Rf = 0.2, 7:3 Cyclohexane-Acetone) to obtain XX$_{19}$ (2.27 g, 99% over 2 steps) as a white solid. $^1$H NMR (CDCl$_3$, 400 MHz) $\delta$ (ppm): 7.49-7.45 (m, 2H, H arom.), 7.41-7.34 (m, 3H, H arom.), 5.91 (m, IH, CH=), 5.85 (d, IH, $J = 9.6$ Hz, NH), 5.55 (s, IH, H7), 5.35 (dd overlapped, IH, H3), 5.32 (m overlapped, IH, =CH$_2$), 5.27 (m, IH, O H$_2$), 4.90 (d, IH, $J_{1-2} = 3.7$ Hz, Hl), 4.37 (ddd, IH, $J_{2-3} = 10.6$ Hz, H2)$_5$ 4.30 (dd, IH, $J_{5-6a} = 4.8$ Hz, $J_{6a-6b} = 10.3$ Hz, H6a), 4.22 (m, IH, OCH$_2$), 4.02 (m, IH, OCH$_3$), 3.95 (ddd, IH,$ J_{5-6b} = 9.5$ Hz, $J_{5-6b} = 9.8$ Hz, H5), 3.80 (pt, IH, H6b), 3.74 (pt, IH, $J_{3-4} = 9.5$ Hz, H4), 2.08, 1.97 (2s, 6H, COCH$_3$); $^{13}$C NMR (CDCl$_3$, 100 MHz) $\delta$ (ppm): 171.4, 170.0 (2C, C=O), 137.0, 133.2, 129.1, 128.2, 128.1, 126.2 (7C, CH=), C arom.), 118.3 (=CH$_2$), 101.6 (CT), 97.2 (Cl), 79.1 (C4), 70.3 (C3), 68.9, 68.7 (2C, C6, OCH$_2$), 63.0 (C5), 52.6 (C2), 23.2, 20.9 (2C, COCH$_3$). HRMS (ESI$^+$) for C$_{20}$H$_{25}$NO$_7$Na ([M+Na]$^+$, 414.1529) m/z 414.1525.

Allyl 2-acetamido-3-0-acetyl-6-0-benzyl-2-deoxy- $\alpha$-D-glucopyranoside (ref. 69) (XX$_2$). Trifluoroacetic acid (1.1 mL, 14.81 mmol) was slowly added to an ice-cold mixture of XX$_{19}$ (1.30 g, 3.29 mmol) and triethylsilane (2.6 mL, 16.45 mmol) in CH$_2$Cl$_2$ (35 mL). After stirring the mixture for 2 h at 0 °C, then for 4.5 h at room temperature, the reaction was quenched by addition of Et$_3$N and concentrated. Purification of the residue by flash column chromatography on silica gel (7:3 → 3:2, Cyclohexane-Acetone) gave $\alpha$-glycoside XX$_{20}$ (0.84 g, 65%) and the de-O-acetylated analogue XX$_{20}^*$ (ref. 77) (284 mg, 21%).

Compound XX$_{20}$ had Rf = 0.2 (7:3 Cyclohexane-Acetone): $^1$H NMR (CDCl$_3$, 400 MHz) $\delta$ (ppm): 7.38-7.28 (m, 5H, H arom.), 5.89 (m, IH, CH=), 5.84 (d, IH, $J = 9.6$ Hz, NH), 5.29 (m, IH, =CH$_2$), 5.22 (m, IH, =CH$_2$), 5.12 (dd, IH, $J_{2-3} = 10.8$ Hz, $J_{3-4} = 8.6$ Hz, H3), 4.86 (d, IH, $J_{1-2} = 3.6$ Hz, Hl), 4.63 (d, IH, OCH$_2$Ph), 4.57 (d, IH, OCH$_2$Ph), 4.26 (ddd, IH, H2), 4.19 (m, IH, OCH$_2$), 4.00 (m, IH, OCH$_2$), 3.85-3.76 (m, 3H, H4, H5, H6a), 3.73 (dd, IH, $J_{5-6b} = 3.9$ Hz, $J_{6a-6b} = 10.2$ Hz, H6b), 2.09, 1.95 (2s, 6H, COCH$_3$); $^{13}$C NMR (CDCl$_3$, 100 MHz) $\delta$ (ppm): 172.2, 170.1 (2C, C=O), 137.7 (C quat. arom.), 133.4 (CH=), 128.5, 127.8, 127.6 (5C, C arom.), 118.0 (=CH$_2$), 96.5 (Cl), 74.1 (C3), 73.7 (OCH$_2$Ph), 70.3 (C5*), 70.0 (C4*), 69.9 (C6), 68.4 (OCH$_2$), 51.8 (C2), 23.2, 21.0 (2C, COCH$_3$). HRMS (ESI$^+$) for C$_{20}$H$_{27}$NO$_7$Na ([M+Na]$^+$, 416.1685) m/z 416.1670.
Compound XX_{20A} had Rf = 0.2 (1:1 Cyclohexane- Acetone). ^1H NMR (CDCl₃, 400 MHz) δ (ppm): 6.03 (d, IH, J = 9.0 Hz, NH), 5.92 (m, IH, CH=), 5.30 (m, IH=C=H₂), 5.24 (m, IH, O H₂), 4.84 (d, IH, J_{i2} = 3.8 Hz, H1), 4.60 (d, IH, OCH₂Ph), 4.57 (d, IH, OCH₂Ph), 4.20 (m overlapped, IH, OCH₂), 4.14 (ddd overlapped, IH, H2), 4.00 (m, IH, OCH₃), 3.82-3.71 (m, 4H, H3, H5, H6a, H6b), 3.60 (pt, IH, J_{3-4} = J_{4-5} = 8.8 Hz, H4), 1.97 (s, 3H, COCH₂); ^13C NMR (CDCl₃, 100 MHz) δ (ppm): 172.5 (C=O), 138.1 (C quat. arom.), 133.4 (CH=), 128.4, 127.8, 127.7 (5C, C arom.), 118.1 (=CH₂), 96.5 (Cl), 74.0 (C3), 73.6 (OCH₂Ph), 71.7 (C4), 70.6 (C5), 69.6 (C6), 68.3 (OCH₂), 53.3 (C2), 23.3 (COCH₂).

Phenyl 2,3,4,6-tetra-O-acetyl-l-thio- β-D-glucopyranoside (ref. 78) (XX₂₁). Thiophenol (7.9 ml, 76.9 mmol) and boron trifluoride etherate (6.3 ml, 51.2 mmol) were successively added at 0°C to a solution of β-D-glucose pentaacetate (10 g, 25.6 mmol) in CH₂Cl₂ (100 mL). After 4 h stirring at this temperature, the organic solution was washed with saturated aqueous NaHCO₃ and water, then dried (Na₂SO₄) and concentrated to give XX₂₁ as a white foam (10.28 g, 91%). Compound XX₂₁ had Rf = 0.5 (4:1 Cyclohexane-EtOAc). ^1H NMR (CDCl₃, 400 MHz) δ (ppm): 7.53-7.51 (m, 2H, H arom.), 7.36-7.32 (m, 3H, H arom.), 5.25 (pt, IH, J_{2-3} = J_{3-4} = 9.4 Hz, H3), 5.06 (pt, IH, J_{4-5} = 9.4 Hz, H4), 5.00 (dd, IH, J_{1-2} = 10.1 Hz, H2), 4.73 (d, IH, H1), 4.25 (dd, IH, J_{5-6a} = 5.0 Hz, J_{6a-6b} = 12.3 Hz, H6a), 4.20 (dd, IH, J_{5-6b} = 2.7 Hz, H6b), 3.75 (ddd, IH, H5), 2.11, 2.04, 2.01 (4s, 12H, COCH₃); ^13C NMR (CDCl₃, 100 MHz) δ (ppm): 170.9, 170.5, 169.7, 169.6 (4C, C=O), 133.5, 132.1, 129.3, 128.8 (6C, C arom.), 86.1 (Cl), 76.2 (C5), 74.4 (C3), 70.4 (C2), 68.7 (C4), 62.6 (C6), 21.1, 21.0, 20.9 (4C, COCH₃).

Phenyl 2,3,4,6-tetra-O-benzyl-l-thio- β-D-glucopyranoside (ref. 79) (XX₂₃). Methanolic sodium methoxide (0.5 M solution, 20 mL, 10.0 mmol) was added to a solution of peracetylated XX₂₁ (10.27 g, 23.3 mmol) in methanol (50 mL). After 7 h at room temperature, the reaction was quenched with Dowex-H⁺ resin, filtered, and concentrated. The crude tetraol (ref. 80) XX₂₂ was dissolved in N,N-dimethylformamide (100 mL) and a 60% suspension of NaH in oil (5.71 g, 143 mmol), and benzyl bromide (28.3 mL, 238 mmol), were added at 0°C. The mixture was allowed to reach room temperature and was stirred overnight. Excess NaH was neutralized with MeOH. After concentration to dryness, the perbenzylated product XX₂₃ (10.94 g, 72% over 2 steps)
was purified by flash chromatography (9:1 Cyclohexane-EtOAc). Compound XX₃ had RF = 0.5 (9:1 Cyclohexane-EtOAc). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 7.64-7.59 (m, 2H₅ H arom.), 7.43-7.21 (m, 23H, H arom.), 4.94-4.84 (m, 4H, OCH₂Ph), 4.76 (d, IH, OCH₂Ph), 4.70 (d overlapped, IH, Hl), 4.65-4.56 (m, 3H, OCH₂Ph), 3.82 (dd, d, IH, J₅₋₆a = 2.0 Hz, J₆a₋₆b = 10.9 Hz, H₆a), 3.76 (dd overlapped, IH₂ H₆b), 3.74 (pt overlapped, IH₅ H₃), 3.68 (pt, IH₅ J₃₋₄ = J₅ = 9.4 Hz H₄, 3.54 (dd, IH₅ J₁₋₂ = 9.8 Hz, J₂₋₃ = 8.6 Hz H₂), 3.53 (ddd, IH, H₅); ¹³C NMR (CDCl₃, 100 MHz) δ (ppm): 138.8, 138.5, 134.3 (4C₅ C quat. arom.), 132.3, 129.3, 128.8, 128.7, 128.6, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8 (2OC, C arom.), 87.9 (Cl) 87.2 (C3), 81.3 (C2), 79.5 (C5), 78.3 (C4), 76.2, 75.8, 75.4, 73.8 (4C₅ OCH₂Ph) 69.5 (C6). HRMS (ESI⁺) for C₄₀H₄₀O₅SNa ([M+Naf, 655.2494] m/z 655.2475.

AHyI 2,3,4,6-tetra-0-benzyl- α-D-glucopyranosyl-(1 →4)-2-acetamido-S-O-acetyl- 6-O-benzyl-1-deoxy- α-D-glucopyranoside (XX₄₄). NIS (353 mg, 1.57 mmol) and TMSOTf (65 μL, 0.36 mmol) were successively added to a solution of thioglycoside donor XX₃ (917 mg, 1.45 mmol) and glycosyl acceptor XX20 (475 mg, 1.21 mmol) in 5:2 Et₂O-CH₂Cl₂ (14 mL) containing activated MS4A under argon. The mixture was stirred at 0°C for 60 min. When TLC monitoring indicated reaction completion (RF = 0.3, 1:1 Cyclohexane-EtOAc), triethylamine was added. After filtration through a bed of celite and concentration under vacuum, the residue was purified by flash chromatography (7:3 Cyclohexane- Acetone) to give slightly contaminated XX₄₄ (780 mg). ¹H NMR (CDCl₃, 400 MHz) δ (ppm): 7.38-7.13 (m, 25H, H arom.), 5.93 (m, IH₅ CH=), 5.86 (d, IH, J = 9.7 Hz NH) 5.44 (dd, IH, J₃₋₄ = 10.7 Hz, J₃₋₄ = 9.1 Hz, H₃d), 5.32 (m, IH₅ CH=), 5.25 (m, IH =CH₂), 5.17 (d, IH, J₁₋₂ = 3.3 Hz, H₁b), 4.91 (d, IH, OCH₂Ph), 4.90 (d, IH, J₁₋₂ = 4.0 Hz, H₁a), 4.83 (d, IH, OCH₂Ph), 4.80 (d, IH, OCH₂Ph), 4.67 (s, 2H, OCH₂Ph), 4.6-4.53 (m, 3H, OCH₂Ph), 4.45 (d, IH₅ OCH₂Ph) 4.37 (ddd, IH₅ H₂D), 4.35 (d, IH₅ OCH₂Ph) 4.20 (m, IH, OCH₂), 4.17 (pt, IH, J₄₋₅ = 9.4 Hz H₄D), 4.05-3.97 (m, 2H, OCH₂, H₆aD), 3.90 (pt, IH, J₃₋₄ = 9.3 Hz, H₃b), 3.89 (m overlapped, IH, H₅b), 3.89 (ddd, IH₅ J₄₋₅ = 9.9 Hz H₅b), 3.70 (dd, IH₅ J₆a₋₆b = 10.8 Hz H₆bD) 3.65 (dd, IH, J₃₋₄ = 9.2 Hz J₄₋₅ = 9.9 Hz H₄b), 3.54 (dd, IH, J₅₋₆a = 3.3 Hz, H₆aD), 3.52 (dd, IH, J₂₋₃ = 9.7 Hz, H₂E), 3.44 (dd, IH, J₆b₋₆a = 1.8 Hz, J₆a₋₆b = 10.6 Hz H₆bD), 1.99, 1.96 (2s, 6H, 2 COCH₃); ¹³C NMR (CDCl₃, 100 MHz) δ (ppm): 171.7, 169.8 (2C, C=O), 138.4, 138.2, 137.9, 137.5 (5C₅ Cquat. arom.),
133.1 (CH=), 128.2, 128.1, 128.0, 127.9, 127.8, 127.6, 127.5, 127.4, 127.3, 127.2, 127.1, 127.0 (25C, C arom.), 117.6 (=CH\_2), 97.0 (C\_1\_E), J\_C=H = 170 Hz), 96.1 (C\_1\_D), 81.3 (C\_3\_E), 79.5 (C\_2\_E), 77.2 (C\_4\_E), 75.3, 74.6 (2C, OCH\_2Ph), 73.1 (C\_3p), 73.0, 72.9, 72.8 (3C, OCH\_3Ph), 72.7 (C\_4\_D), 70.8 (C\_5\_E), 70.2 (C\_5\_D), 68.1 (2C, OCH\_2, C\_6D), 67.9 (C\_6\_E), 51.9 (C\_2D), 22.9, 20.9 (2C, COCH\_3).

AUyI 2,3,4,6-tetra-0-benzyl-\(\alpha\)-D-glucopyranosyl-(l \rightarrow 4)-2-acetamido-6-O-benzyl-2-deoxy-\(\alpha\)-D-glucopyranoside (XX\_s). Methanolic sodium methoxide (0.5 M solution, 0.4 mL, 0.2 mmol) was added to a stirred solution of XX\_2\_4 (742 mg, 0.92 mmol) in anhydrous methanol (7.5 mL). The reaction mixture was stirred for 1 day at room temperature by which time all the starting material had been consumed (Rf = 0.4, 1:1 Cyclohexane-EtOAc). Excess base was neutralized with Dowex-H\_+ resin. After removal of the resin by filtration, the filtrate was concentrated, and the residue was purified by silica gel chromatography, eluting with 1:1 Cyclohexane-EtOAc to obtain compound XX\_2\_s as a colourless oil (555 mg, 69%); \(^1\)H NMR (CDCl\_3, 400 MHz) \(\delta\) (ppm): 7.39-7.18 (m, 25H, H arom.), 5.93 (m, IH, CH=), 5.87 (d, IH, J = 8.4 Hz, NH), 5.32 (m, IH, =CH\_2), 5.24 (m, IH, =CH\_2), 5.03 (d, IH, J\_1\_2 = 3.5 Hz, H\_1\_E), 5.00 (d, IH, J\_1\_2 = 3.7 Hz, H\_1\_D), 4.96-4.83 (m, 5H, OCH\_2Ph), 4.77 (d, IH, OCH\_3Ph), 4.62-4.43 (m, 4H, OCH\_2Ph), 4.25-4.18 (m, 2H, 1H OCH\_2, H\_2D), 4.05-3.99 (m, \^2\^2\_H, 1H OCH\_2, H\_3\_E), 3.96 (dd, IH, J\_2\_3 = 10.7 Hz, J\_3\_4 = 8.6 Hz, H\_3\_D), 3.88 (ddd, IH, J\_4\_5 = 10.0 Hz, J\_5\_6\_a = 3.5 Hz, H\_5\_E), 3.81-3.60 (m, 6H, H\_4\_D, H\_5\_D, H\_6\_a\_D, H\_6b\_D, H\_4\_E, H\_6a\_E), 3.59 (dd, IH, J\_2\_3 = 10.0 Hz, H\_2\_E), 3.50 (dd, IH, J\_5\_6\_b = 1.7 Hz, J\_6\_a\_b = 10.6 Hz, H\_6b\_E), 2.06 (s, 3H, COCH\_3); \(^1\)C NMR (CDCl\_3, 100 MHz) \(\delta\) (ppm): 170.5 (C=O), 138.6, 138.5, 138.2, 137.9, 137.0 (7C, C quat. arom.), 133.8 (CH=), 128.7, 128.6, 128.5, 128.4, 128.3, 127.9, 127.8 (25C, C arom.), 117.6 (=CH\_2), 100.6 (C\_1\_E, J\_C=H = 172 Hz), 96.5 (C\_1\_D), 82.2 (C\_3\_E), 82.0 (C\_4\_D), 79.6 (C\_2\_E), 77.8 (C\_4\_E), 75.7, 75.0, 74.1, 73.5, 73.2 (5C, OCH\_2Ph), 72.0 (C\_3\_D), 71.5 (C\_5\_E), 70.2 (C\_5\_D), 69.0 (C\_6\_D), 68.5 (C\_6\_E), 68.4 (OCH\_2), 53.3 (C\_2\_D), 23.4 (COCH\_3). HRMS (ESI\+)

\(\text{for C}_{32}\text{H}_{59}\text{NO}_{11}\text{Na ([M+Na] \^+}, 896.3986 \text{ m/z 896.4020).}

**EXAMPLE 7: Synthesis and characterization of \(\alpha\)-D-glucopyranosyl\(\rightarrow 4\)-7V-acetyl-D-glucosamine (P2)**

Glucosylation of JV-acetyl-D-glucosamine (acceptor) \(\alpha\)-D-glucopyranosyl-(l \rightarrow 4)-N-acetyl-D-glucosamine (P2) is the main acceptor reaction product obtained by action
of 1G5A amylosucrase (AS) (Recombinant form of *Neisseria polysaccharea* amylosucrase EC2.4.1.4 (GH13)) using sucrose as donor and D-GlcpNAc as acceptor in a molar ratio of 1.

**7.1. Materials**

Recombinant enzyme were produced in *E.coli* as reported elsewhere. Purified AS, conserved at -20°C or -80°C, served for enzymatic reaction.

JV-acetyl-D-glucosamine was purchased from Sigma-Aldrich.

**7.2. Acceptor reaction assay**

The glucosylation reaction was performed in the enzyme optimal buffer: in Tris-HCl (50 mM, pH = 7.5) for AS assay. The reaction mixture was carried out at 30°C with sucrose and acceptor in equimolar ratio (146 mM). AS were used at 1 X1ImL. Activity one unit is defined as the amount of enzyme that catalyzes the formation of 1µmol of fructose/min at 30°C, in enzyme buffer and sucrose 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 µm membrane before HPLC analysis.

**7.3. Glucosyl acceptor production**

In order to characterize glucosylated products of iV-acetyl-D-glucosamine, acceptor reactions were conducted at preparative scale.

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).

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.
7.4. Analytical methods

7.4.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 µm (elution with ultra pure water at room temperature and 1mL/min) (iv) C30: Bischoff Prontosil Eurobond, 5 µm, 250 x 4.0 mm (elution with ultra pure water at room temperature and 1mL/min) (v) C18RP: Sinergi Fusion RP Phenomenex, 4µm, 250 x 4.6 mm (elution with ultra pure water at room temperature and 1 mL/min).

7Ab. 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 µL of sample was mixed in the tip of the probe with a glycerol or dithiothreitol/dithioerythritol matrix.

NMR analyses: ¹H (400.130 MHz), ¹³C (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.

EXAMPLE 8: Engineering double mutant amylosucrases for the synthesis the α-D-glucopyranosyl-(1→4)VΛ^-acetyl-α-D-glueopyranosaminyl disaccharide

1) Materials and Methods

Plasmids, bacterial strains and chemicals

Plasmid pGST-AS (see Example 1-1) was used for the construction of the AS double-mutant libraries.
Fusion DNA-polymerase was purchased from Finnzymes (Espoo, Finland), and Dpnl restriction enzyme from New England Biolabs (Beverly, MA, USA).

Oligonucleotides were synthetised 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 D-GlcpNAc glucosylation were sequenced on the mutated region (-600 bp) using the primer pGEXInt: CCAACGAACACGAATGGGC (SEQ ID NO: 28).

Ampicillin (Amp), lysozyme, and isopropyl-β-D-thiogalactopyranoside (IPTG) were purchased from Euromedex (Souffelweyersheim, France); Bromothymol Blue sodium salt, sucrose and JV-acetyl-D-glucosamine (D-GlcpNAc) from Sigma-Aldrich (Saint-Louis, MO, USA).

Reference disaccharides α-D-Glcp-(1→4)-D-Glc/?NAc and α-D-Glcp-(1→6)-D-GlcpNAc were enzymatically synthesized and characterized (see Example 7).

**Construction of libraries containing two vicinal mutations (libraries 1-3)**

Site-saturation mutagenesis, focused on the vicinal positions (1228-F229, A289-F290, 1330-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.

Three partial overlapping primer pairs surrounding double codons were designed. Each of these codons was replaced in the primers with degenerate NNS or NNW sequence, where N=A, C, G or T; S=C or G and W=A or T (Table V below).

Such degenerate primers were designed to generate 32 codons encoding the 20 possible amino acids.
Table V: Degenerate primers used for the construction of libraries (I228-F229, A289-F290, I330-V331)

<table>
<thead>
<tr>
<th>SEQ ID</th>
<th>Primer name</th>
<th>Nucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>228-229 for</td>
<td>CTG CGC GAA NNS NNS CCC GAC CAC CCG GGC G</td>
</tr>
<tr>
<td>30</td>
<td>228-229 rev</td>
<td>CTG GTC GGG WNN WNN TTC GCG GAT GGT GCG GTC G</td>
</tr>
<tr>
<td>31</td>
<td>289-290 for</td>
<td>GCG GTT NNS NNW ATT TGG AAA CAA ATG GGG ACA AGC TGC G</td>
</tr>
<tr>
<td>32</td>
<td>289-290 rev</td>
<td>CCA AAT SNN WNN AAC CGC ATC CAT ACG CAG GAT GTC AAT GAC GCC</td>
</tr>
<tr>
<td>33</td>
<td>330-31 for</td>
<td>TCC GAA GCC NNS NNS CAC CCC GAC CAA GTC GTC C</td>
</tr>
<tr>
<td>10</td>
<td>330-31 rev</td>
<td>GGG GTG WNN WNN GGC TTC GGA TTT GAA GAA CAC GGC</td>
</tr>
</tbody>
</table>

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 TOPIO was transformed by electroporation with 4µL of each plasmid library using standard procedures.

**Construction of library containing two distant mutations** (library 4)

pGST-AS G537D was also used as vector template for library (1228-F290) construction. First, for each targeted position, the 2x19 complementary primers previously constructed by site-directed mutagenesis (see Example 1-1) were pooled in equimolar ratio, that thus formed two pairs of degenerate primers, for positions 228 and 290, respectively. Then, it was proceeded 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.

**HPLC screening of variants able to glucosylate D-GlcNAc**

The method for the expression of the mutant libraries is described in Example 1-1.

Accepter reaction products were analyzed by HPLC analysis using a C18-AQ column (Bischoff C18, 125x4 mm, 3µm) kept at room temperature and eluted with 0.6 mL/min of ultra-pure water to detect α-D-Glc-(1→4)-α-D-GlcNAc formation (analysis time: 7 min). To evaluate the efficiency of the glucosylation reaction, it was
carried out complementary HPLC analyses with a Biorad HPX-87K Carbohydrate Analysis 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 concomitant α-D-Glc-(1→4)-α-D-GlcNAc formation by UV220nm and determine mutant ability to synthesize the desired disaccharide α-D-Glc-(1→4)-α-D-GlcNAc (z.e., % Glucosyl units transferred onto acceptor derivatives = [Glucosyl units transferred onto acceptor derivatives]/[Glucosyl units transferred from initial sucrose] x 100)

Production, purification and characterization of the selected variants: A289P-F290C, A289P-F290I and A289P-F290L.

The double mutants A289P-F290C, A289P-F290I and A289P-F290L were produced and purified, as previously described in Example 1-1. The protein content was determined by the Nanodrop ND-1000 spectrophotometer. These mutants were tested on sucrose alone or supplemented with α-D-GlcNAc-OAll acceptor, and compared to WtAS and F290K. Assays were performed at 30 °C in 50 mM Tris-HCl buffer, pH = 7.0.

Standard activity determination

One unit of amylosucrase activity corresponds to the amount of enzyme that catalyzes the release of 1 µmol of reducing sugars per minute in the assay conditions. When using sucrose as sole substrate, specific activity was determined using 250 mM sucrose. In the presence of both sucrose and α-D-GlcNAc-OAll acceptor, specific activity was determined using 250 mM sucrose and 250 mM α-D-GlcNAc-OAll. The concentration of reducing sugars was determined using the dinitrosalicylic method (ref. 35) and fructose as standard.

Comparison of products synthesized by wild-type (wtAS) and AS Variants

Reactions were performed in the presence of 250 mM sucrose alone or supplemented with 250 mM α-D-GlcNAc-OAll acceptor. The purified wtAS or mutated AS were employed at 1 U/mL. The reactions were stopped after 24 h by heating at 95°C for 5 min. The soluble part of the reaction mixture was submitted to HPAEC (high-performance anion-exchange chromatography with pulsed amperometric detection). To quantify the concentration of monosaccharides and disaccharides, the soluble fraction was diluted in water and separated on a 4*250 mm Dionex Carbo-pack
PAIOO column. A gradient of sodium acetate (from 6 to 300 mM in 28 min) in 150 mM NaOH was applied at 1 ml/min flow rate. Detection was performed using a Dionex ED40 module with a gold working electrode and an Ag/AgCl pH reference. α-D-GlcpNAc-OAU, and their derivatives, are not oxidable products and thus are not detectable by HPAEC. Therefore, α-D-GlcpNAc-OAll and their glucosylation products, were quantified by HPLC with a Biorad HPX-87K (see above). Concentration of sucrose, glucose, fructose, turanose, trehalulose and maltose was determined by HPAEC.

In parallel, the reaction mixture containing soluble and insoluble malto-oligosaccharides was solubilized in 1 M aq KOH at a final total sugar concentration of 10 g/L and analysed by HPAEC using a Dionex Carbo-Pack PAIOO column at 30 °C. Mobile phase (150 mM aq NaOH) was set at 1 mL/min flow rate with a sodium acetate gradient (6 to 500 mM over 120 min).

**Determination of kinetic parameters towards sucrose donor**

Enzyme assays were carried out in a total volume of 2 mL containing pure enzyme (0.073 mg, 0.153 mg and 0.092 mg when using A289P-F290C, A289P-F290I and A289P-F290L, respectively).

Kinetic studies of AS variants were performed in the presence of sucrose (0-500 mM). At regular time intervals (2-5 min), aliquots (200 µL) were removed, heated (95 °C, 2 min) and centrifuged (18000 g, 5 min). All the samples were filtered on a 0.22 µm membrane and analyzed using HPLC material previously described. The initial rate of sucrose consumed, corresponding to the initial rate of fructose released, was expressed in µmole of fructose released per minute and per gram of enzyme. The kinetic parameters $k_{cat}$ and $K_m$ were calculated using Eadie-Hofstee plot.

**Determination of kinetic parameters towards α-D-GlcpNAc-OAll acceptor**

Enzyme assays were carried out in a total volume of 2 mL containing pure enzyme (0.037 mg, 0.076 mg and 0.046 mg when using A289P-F290C, A289P-F290I and A289P-F290L, respectively).

Catalytic efficiency ($k_{cat}/K_m$) of AS variants towards α-D-GlcpNAc-OAll acceptor was determined using sucrose (250 mM) and acceptor α-D-GlcpNAc-
OAll as variable substrate (0-25OmM) and following the same protocol as described above.

The \( k_{cat}/K_m \) (\( \alpha \)-D-GlcpNAc-OAll) value was calculated from the initial rate of formation of desired disaccharide (\( \alpha \)-D-Glcp-(l-\( \rightarrow \)4)-\( \alpha \)-D-GlcpNAc-OAll)_5 corresponding to the initial rate of \( \alpha \)-D-GlcpNAc-OAll consumed. As saturation was not achieved with the mutants, efficiency was calculated by linear regression analysis of the velocity versus substrate concentration plot.

2) Results

Library construction and pre-screening of sucrose-utilizing variants

It was showed in Example 1 that positions 228 and 290 are key positions for altering AS selectivity towards D-GlcpNAc. However, given the spatial vicinity of these two sites with positions 229, 289 and 330, and the influence they might have on each other, it was constructed libraries focused on all 5 positions 228, 289, 290, 330, and 331.

To enlarge the exploration of subsite +1 sequence space, it was generated 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 due to the genetic code degenerescence.

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 the 20 possible amino acids (see Figure 12). In this way, 400 (20x20) 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 for the 19 possible amino acids. All possible amino acid changes are considered possible in this library.

The libraries were then screened to determine both their size and the ratio of active clones using the pH indicator screen (see the method described in Example 1), as shown in Table VI below.
Table VI: Number of clones and active clones obtained for the four libraries

<table>
<thead>
<tr>
<th>Library</th>
<th>Number of clones</th>
<th>Number of active clones isolated (%)</th>
<th>Number of active clones (yellow)</th>
<th>Number of moderately active clones (green)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>228-229</td>
<td>~8000</td>
<td>180 (2.3%)</td>
<td>84</td>
</tr>
<tr>
<td>2</td>
<td>289-290</td>
<td>~2000</td>
<td>576 (29%)</td>
<td>480</td>
</tr>
<tr>
<td>3</td>
<td>330-331</td>
<td>~3200</td>
<td>576 (18%)</td>
<td>480</td>
</tr>
<tr>
<td>4</td>
<td>228-290</td>
<td>~8000</td>
<td>384 (4.8%)</td>
<td>288</td>
</tr>
</tbody>
</table>

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.

**Isolation of improved Mutants for α-D-Glcp-(1→4)-α-D-GlcNAc 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 D-GlcpNAc was then carried out in microplate format. HLPC analysis time for each clone was reduced to 7 min for detection of α-D-Glcp-(1→4)-α-D-GlcpNAc. When disaccharide formation was detected, complementary analyses were performed to determine sucrose consumption and disaccharide formation with more accuracy in order to calculate the glucosylation yield.

**Improved Mutants for α-D-Glcp-(1→4)-D-GlcpNAc synthesis**

Of the 1716 tested mutants, 30 were able to catalyze α-D-Glcp-(1→4)-D-GlcpNAc synthesis with a glucosylation rate higher than 15% (see Table VII below).
Table VII: Improved mutants for α-D-Glc/?-(1→4)-α-D-GlcpNAc (ED) synthesis

<table>
<thead>
<tr>
<th>Mutants</th>
<th>Library number</th>
<th>Single-Mutant (SM) or Double-Mutant (DM)</th>
<th>Amino acid Mutation(s)</th>
<th>Mutation(s) (codon)</th>
<th>Glucosylation Yield to synthesize ED</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtAS</td>
<td></td>
<td>DM</td>
<td>A289P-F290C</td>
<td>CCG, TGT</td>
<td>2%</td>
</tr>
<tr>
<td>DM2 P1 B8</td>
<td>2</td>
<td>DM</td>
<td>A289P-F290C</td>
<td>CCG, TGT</td>
<td>100%</td>
</tr>
<tr>
<td>DM2 P1 B6</td>
<td>2</td>
<td>DM</td>
<td>A289P-F290I</td>
<td>CCC, ATT</td>
<td>100%</td>
</tr>
<tr>
<td>DM2 P4 F7</td>
<td>2</td>
<td>DM</td>
<td>A289P-F290L</td>
<td>CCC, TTA</td>
<td>100%</td>
</tr>
<tr>
<td>DM2 P4 H5</td>
<td>2</td>
<td>DM</td>
<td>F290V</td>
<td>GTT</td>
<td>100%</td>
</tr>
<tr>
<td>DM4 P1 H10</td>
<td>4</td>
<td>SM</td>
<td>F290R</td>
<td>CGT</td>
<td>100%</td>
</tr>
<tr>
<td>DM4 P1 H11</td>
<td>4</td>
<td>SM</td>
<td>I228V-F290V</td>
<td>GTC, GTT</td>
<td>90%</td>
</tr>
<tr>
<td>DM4 P1 G7</td>
<td>4</td>
<td>DM</td>
<td>I228V-F290K</td>
<td>GTC, AAA</td>
<td>89%</td>
</tr>
<tr>
<td>DM4 P1 C7</td>
<td>4</td>
<td>DM</td>
<td>I228A-F290H</td>
<td>GCC, CAT</td>
<td>88%</td>
</tr>
<tr>
<td>DM4 P1 C1</td>
<td>4</td>
<td>SM</td>
<td>F229V</td>
<td>GTG</td>
<td>88%</td>
</tr>
<tr>
<td>DM1 P1 C1</td>
<td>1</td>
<td>SM</td>
<td>F229M</td>
<td>ATG</td>
<td>79%</td>
</tr>
<tr>
<td>DM1 P1 C12</td>
<td>1</td>
<td>SM</td>
<td>I228C-F229L</td>
<td>TGC, CTC</td>
<td>78%</td>
</tr>
<tr>
<td>DM4 P1 D11</td>
<td>1</td>
<td>DM</td>
<td>I228T-F290K</td>
<td>ACC, AAA</td>
<td>72%</td>
</tr>
<tr>
<td>DM4 P1 A6</td>
<td>4</td>
<td>DM</td>
<td>I228V-F290H</td>
<td>GTC, CAT</td>
<td>61%</td>
</tr>
<tr>
<td>DM1 P2 F7</td>
<td>1</td>
<td>DM</td>
<td>I228F-F229L</td>
<td>GTG, TTG</td>
<td>56%</td>
</tr>
<tr>
<td>DM4 P2 A2</td>
<td>4</td>
<td>SM</td>
<td>F290L</td>
<td>CTT</td>
<td>55%</td>
</tr>
<tr>
<td>DM4 P2 H10</td>
<td>4</td>
<td>SM</td>
<td>F290G</td>
<td>GTG</td>
<td>53%</td>
</tr>
<tr>
<td>DM4 P2 C2</td>
<td>4</td>
<td>DM</td>
<td>I228K-F290W</td>
<td>AAG, TGG</td>
<td>50%</td>
</tr>
<tr>
<td>DM1 P2 C8</td>
<td>1</td>
<td>SM</td>
<td>F229L</td>
<td>CTC</td>
<td>49%</td>
</tr>
<tr>
<td>DM4 P2 B10</td>
<td>4</td>
<td>DM</td>
<td>I228V-F290V</td>
<td>GTC, GTT</td>
<td>47%</td>
</tr>
<tr>
<td>DM4 P2 G4</td>
<td>4</td>
<td>DM</td>
<td>I228K-F290L</td>
<td>AAG, CTT</td>
<td>47%</td>
</tr>
<tr>
<td>DM4 P3 C8</td>
<td>4</td>
<td>SM</td>
<td>F290H</td>
<td>CAT</td>
<td>46%</td>
</tr>
<tr>
<td>DM4 P2 H4</td>
<td>4</td>
<td>SM</td>
<td>F290K</td>
<td>AAA</td>
<td>42%</td>
</tr>
<tr>
<td>DM4 P2 D12</td>
<td>4</td>
<td>SM</td>
<td>F290K</td>
<td>AAA</td>
<td>41%</td>
</tr>
<tr>
<td>DM2 P1 G10</td>
<td>2</td>
<td>DM</td>
<td>A289T-F290H</td>
<td>ACG, CAT</td>
<td>40%</td>
</tr>
<tr>
<td>DM2 P3 C3</td>
<td>2</td>
<td>DM</td>
<td>A289H-F290S</td>
<td>CAC, TCT</td>
<td>40%</td>
</tr>
<tr>
<td>DM1 P1 H2</td>
<td>1</td>
<td>DM</td>
<td>I228L-F229L</td>
<td>TTG, CTC</td>
<td>35%</td>
</tr>
<tr>
<td>DM1 P2 F2</td>
<td>1</td>
<td>DM</td>
<td>I228V-F229M</td>
<td>GTG, ATG</td>
<td>29%</td>
</tr>
<tr>
<td>DM1 P2 C3</td>
<td>1</td>
<td>DM</td>
<td>I228M-F229M</td>
<td>ATG, ATG</td>
<td>28%</td>
</tr>
<tr>
<td>DM4 P3 G9</td>
<td>4</td>
<td>DM</td>
<td>I228V-F290R</td>
<td>GTC, CGT</td>
<td>25%</td>
</tr>
<tr>
<td>DM4 P3 F7</td>
<td>4</td>
<td>DM</td>
<td>I228P-F290C</td>
<td>CCC, TGT</td>
<td>22%</td>
</tr>
<tr>
<td>DM4 P2 G9</td>
<td>4</td>
<td>SM</td>
<td>F290L</td>
<td>CTT</td>
<td>20%</td>
</tr>
<tr>
<td>DM4 P4 B12</td>
<td>4</td>
<td>SM</td>
<td>F290I</td>
<td>ATT</td>
<td>16%</td>
</tr>
</tbody>
</table>

Among these positive mutants, it was distinguished 20 double-mutants, 6 single-mutants previously identified (see Example 1) and 4 new single-
mutants (F290R, F229L, F229M and F229V). Out of these 30 mutants, 4 were identified twice (A289P-F290C, I228V-F290V, F290K and F290L). Most active double-mutants (13 out of 20) show a mutation at position 290, a position that was previously identified as critical for D-GlcpNAc recognition. The glucosylation rates determined for F290L and I228V F290V mutants differed among the two identified clones from 20 to 55% and from 47% to 90%, respectively. These variations reflect the poor reproducibility of the glucosylation assay in microtiter plates. Microtiter experiments combined with HPLC analysis are designed for high-throughput screening. Consequently, to evaluate the impact of the mutations on the catalytic efficiency for sucrose and \( \alpha\text{-D-Glc}(1\rightarrow4)\alpha\text{-D-GlcNAc-OAll} \) formation, the most improved mutants, particularly the double-mutants A289P-F290C, A289P-F290I and A289P-F290L were further characterized and compared with F290K, which was the more appropriated mono-mutant for \( \alpha\text{-D-GlcNAc-OAll} \) glucosylation.

Characterization of A289P-F290C, A289P-F290I and A289P-F290L mutants

Mutants A289P-F290C, A289P-F290I and A289P-F290L were produced and purified to homogeneity to determine their kinetic parameters and their product reaction profile. They were both compared to that obtained with F290K and wtAS.

In the presence of sucrose alone (250 mM), all 3 double-mutants kept the ability to synthesize maltooligosaccharides (up to DP 20), similarly to F290K (see Figure 13A). In addition, all three double mutants produced an increased amount of sucrose isomers (trehalulose and mostly turanose) compared to wtAS (see Figure 13B).

When \( \alpha\text{-D-GlcNAc-OAll} \) was added, the three double mutants displayed a very high specificity for the acceptor, as observed for F290K. Indeed, maltooligosaccharide production was fully suppressed in favour of \( \alpha\text{-D-GlcNAc-OAll} \) glucosylation (see Figure 14A). The quasi-totality of the sucrose glucosyl residues was transferred onto \( \alpha\text{-D-GlcNAc-OAll} \), converting \( \alpha\text{-D-GlcNAc-OAll} \) into \( \alpha\text{-D-Glc}(1\rightarrow4)\alpha\text{-D-GlcNAc-OAll} \) (46% to 64%) and a di-glucosylated form (\( \alpha\text{-D-Glc}(1\rightarrow4)\alpha\text{-D-Glc;NAc-OAll} \)) (13% to 24%) (see Figure 14B).

To compare the improved catalytic properties of the double-mutants, it was first investigated the kinetic parameters of these mutants towards sucrose donor.
Upon varying the sucrose concentration (0-250 nM), all three double-mutants showed standard saturation kinetic behavior (Table VIII).

**Table VIII:** Comparison of the kinetic parameters of WtAS, F290K and improved double-mutants (A289P-F290C, A289P-F290I and A289P-F290L) for sucrose donor.

<table>
<thead>
<tr>
<th></th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$ mM$^{-1}$)</th>
<th>$V_{250mM}$</th>
<th>behaviour (at high concentration)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WtAS</td>
<td>1.3</td>
<td>50.2</td>
<td>0.0261</td>
<td>900</td>
<td>saturation</td>
</tr>
<tr>
<td>F290K</td>
<td>n.d.</td>
<td>n.d.</td>
<td>0.0006</td>
<td>110</td>
<td>linear</td>
</tr>
<tr>
<td>A289P-F290C</td>
<td>2.8</td>
<td>18.9</td>
<td>0.1502</td>
<td>2186</td>
<td>saturation</td>
</tr>
<tr>
<td>A289P-F290I</td>
<td>1.1</td>
<td>17.2</td>
<td>0.0655</td>
<td>946</td>
<td>saturation</td>
</tr>
<tr>
<td>A289P-F290L</td>
<td>2.1</td>
<td>10.1</td>
<td>0.2094</td>
<td>1717</td>
<td>saturation</td>
</tr>
</tbody>
</table>

n.d.: not determined ; a Data from ref. 83 ; b Initial rate of sucrose consumption was determined at a concentration of 250 mM for sucrose.

In contrast, no kinetic saturation was achieved with F290K, which displayed a poor affinity for sucrose. The $k_{cat}/K_m$ of the double-mutants was improved 100 to 350-fold relative to that of F290K. In comparison to WtAS, the catalytic efficiency of the 3 screened mutants toward sucrose was increased by 2.5 to 8-fold. For all double-mutants, the $K_m$ and the $k_{cat}$ turnover values were improved (Table VIII above).

To investigate the effect of the double mutations on the acceptor specificity of the amylosucrase, kinetic parameters were determined for various acceptor concentrations using a fixed donor sugar concentration (250 mM). Results are shown in Table IX below.
Table IX: Comparison of the kinetic parameters of wtAS, F290K and improved double-mutants (A289P-F290C, A289P-F290I and A289P-F290L) for $\alpha$-D-GlcpNAc-OAll acceptor with sucrose fixed at 250 mM.

<table>
<thead>
<tr>
<th></th>
<th>$k_{cat}/K_{M}$ (s$^{-1}$ mM$^{-1}$)</th>
<th>$V_{250\text{mM}}$ (µmol/min/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtAS</td>
<td>0.002</td>
<td>290</td>
</tr>
<tr>
<td>F290K$^a$</td>
<td>0.265</td>
<td>6690</td>
</tr>
<tr>
<td>A289P-F290C</td>
<td>0.690</td>
<td>50 000</td>
</tr>
<tr>
<td>A289P-F290I</td>
<td>0.700</td>
<td>51 400</td>
</tr>
<tr>
<td>A289P-F290L</td>
<td>0.790</td>
<td>67 300</td>
</tr>
</tbody>
</table>

$^b$ Initial rate of acceptor consumption was determined at a concentration of 250 mM for both sucrose and $\alpha$-D-GlcpNAc-OAll.

The catalytic efficiency of the 3 double-mutants towards $\alpha$-D-GlcpNAc-OAll was increased by up to a remarkable 395-fold compared to wtAS and 3-fold compared to variant F290K.

REFERENCES

1. A method for preparing the building block corresponding to a disaccharide $\alpha$-D-glucopyranosyl-(1 $\rightarrow$ 4)-N-acetyl-$\alpha$-D-glucopyranosaminyl of formula (Ia):

$$\text{(Ia)}$$

said method being characterized in that it comprises the step of reacting:

A) a mutant of a glycoside hydrolase,

* wherein said wild type glycoside hydrolase has 450 to 850 amino acids and comprises eleven motifs defined by the following consensus motifs:

1. the amino acid sequence LGVNYLHLMPL (SEQ ID NO: 1), which is located in the $\beta$-strand 2 of said wild type glycoside hydrolase;

2. the amino acid sequence DGGYAV (SEQ ID NO: 2), which is located in the loop 2 of the ($\beta$$\alpha_8$)-barrel of said wild type glycoside hydrolase;

3. the amino acid sequence DFVFNH (SEQ ID NO: 3) which is located in the $\beta$-strand 3 of said wild type glycoside hydrolase;

4. the amino acid sequence LREIFPDAPGNF (SEQ ID NO: 4), which is located in the domain B of said wild type glycoside hydrolase;

5. the amino acid sequence FNSYQWDLN (SEQ ID NO: 5), which is located in the C-terminal part of the domain B of said wild type glycoside hydrolase;

6. the amino acid sequence ILRLDAVAFLWK (SEQ ID NO: 6), which is located in the $\beta$-strand 4 of said wild type glycoside hydrolase;

7. the amino acid sequence EAIV (SEQ ID NO: 7), which is located in the $\beta$-strand 5 of said wild type glycoside hydrolase;
(8) the amino acid sequence YVRCHDDI (SEQ ID NO: 8), which is located in the β-strand 7 of said wild type glycoside hydrolase;

(9) the amino acid sequence RISGTLASLAG (SEQ ID NO: 9), which is located in the domain B’ of said wild type glycoside hydrolase;

(10) the amino acid sequence GIPLIYLGDE (SEQ ID NO: 10), which is located in the β-strand 8 of said wild type glycoside hydrolase;

(11) the amino acid sequence RWVHRP (SEQ ID NO: 11), which is located in the loop 8 of the (β/α)_{8}-barrel,

and the sequence formed by said eleven motifs from said wild type glycoside hydrolase joined end-to-end from motif (1) to motif (11) has at least 65% sequence identity or at least 80% sequence similarity with the amino acid sequence SEQ ID NO: 12;

- wherein said mutant has one or two mutation(s) consisting of, when said mutant has only one mutation:
  - the substitution of the amino acid residue at position 4 in said motif (4) with any amino acid selected from the group consisting of alanine (A), cysteine (C), glutamic acid (E), glycine (G), histidine (H), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), valine (V) with the provisio that said wild type glycoside hydrolase does not contain a valine at this position, tryptophan (W) and tyrosine (Y), or
  - the substitution of the amino acid residue at position 5 in said motif (4) with any amino acid selected from the group consisting of leucine (L), methionine (M) and valine (V), or
  - the substitution of the amino acid residue at position 8 in said motif (6) with any amino acid selected from the group consisting of glutamic acid (E), phenylalanine (F), glycine (G), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R) and valine (V), or
  - the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V) and tryptophan (W), or
- the substitution of the amino acid residue at position 4 in said motif (7) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), serine (S), threonine (T) and tyrosine (Y), or

- the substitution of the amino acid residue at position 7 in said motif (8) with any amino acid selected from the group consisting of alanine (A) and valine (V), or

- the substitution of the amino acid residue at position 1 in said motif (9) with any amino acid selected from the group consisting of alanine (A), cysteine (C), phenylalanine (F), glycine (G) with the provisio that said wild type glycoside hydrolase does not contain a glycine at this position, lysine (K), asparagine (N), glutamine (Q), serine (S) with the provisio that said wild type glycoside hydrolase does not contain a serine at this position, threonine (T) and tryptophan (W), or when said mutant has two mutations:

- the substitution of the amino acid residue at position 4 in said motif (4) with an alanine (A) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a cysteine (C) and the substitution of the amino acid residue at position 5 in said motif (4) with a leucine (L), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a lysine (K) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of leucine (L) and tryptophan (W), or

- the substitution of the amino acid residues at positions 4 and 5 in said motif (4) respectively with a leucine (L), or

- the substitution of the amino acid residues at positions 4 and 5 in said motif (4) respectively with a methionine (M), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with a cysteine (C), or
- the substitution of the amino acid residue at position 4 in said motif (4) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of histidine (H) and lysine (K), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a valine (V) and the substitution of the amino acid residue at position 5 in said motif (4) with any amino acid selected from the group consisting of leucine (L) and methionine (M), or

- the substitution of the amino acid residue at position 4 in said motif (4) with a valine (V) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of histidine (H), lysine (K), arginine (R), and valine (V), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a serine (S), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of cysteine (C), isoleucine (I) and leucine (L), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H);

B) with the acceptor of formula (Ha):

\[
\text{(IIa)}
\]

wherein Y is selected from -O- and -S- and R is selected from the group consisting of: C<sub>1</sub>-C<sub>6</sub> alkyl, C<sub>1</sub>-C<sub>6</sub> alkenyl, aryl, allyl, -CO-alkyl (C<sub>1</sub>-C<sub>6</sub>), -CO-alkenyl (C<sub>1</sub>-C<sub>6</sub>), -CO-aryl,
86

R’ designates a group selected from: acetyl, trichloroacetyl, trifluoroacetyl, and

c) with a donor of formula (IIia):

(IIIa)

wherein R₁ represents a group selected from:

2. A method according to claim 1, characterized in that said wild type glycoside hydrolase is an amylosucrase (EC 2.4.1.4) or sucrose hydrolase (EC 3.2.1.-).

3. A method according to claim 2, characterized in that said wild type glycoside hydrolase is an amylosucrase from Neisseriapoly saccharea, and is preferably selected from the group consisting of 1G5A (SEQ ID NO: 13), 1ZS2, IMVY, IMWO, 1S46, IJGI, 1MW2, 1MW3, IMW1 and IJG9 proteins.

4. A method according to anyone of claims 1 to 3, for the preparation of the disaccharide α-D-glucopyranosyl-(1→4)-2-N-acetyl-2-deoxy-α-D-glucopyranoside of formula (I):

(I)

said method being characterized in that it comprises the step of reacting a mutant of a glycoside hydrolase as disclosed in anyone of claims 1 to 3, with the acceptor of formula (II):
5. A method for the preparation of the building block corresponding to the disaccharide of formula (XX₃B) in which R₂ represents a group selected from H, Bn, Ac and AcBn and R₃ represents a group selected from H and Ac.

(XX₃B)

comprising at least one step according to claim 1 or claim 4.

6. A method according to claim 5, for the preparation of the disaccharide of formula (XX₃A)
7. A mutant of a wild type glycoside hydrolase, wherein said wild type glycoside hydrolase has 450 to 850 amino acids and comprises eleven motifs defined by the following consensus motifs:

(1) the amino acid sequence LGVNYLHLMPL (SEQ ID NO: 1), which is located in the β-strand 2 of said wild type glycoside hydrolase;

(2) the amino acid sequence DGGYAV (SEQ ID NO: 2), which is located in the loop 2 of the \((\beta/\alpha)_8\)-barrel of said wild type glycoside hydrolase;

(3) the amino acid sequence DFVFNH (SEQ ID NO: 3) which is located in the β-strand 3 of said wild type glycoside hydrolase;

(4) the amino acid sequence LREIFPDATPGNF (SEQ ID NO: 4), which is located in the domain B of said wild type glycoside hydrolase;

(5) the amino acid sequence FNSYQWDLN (SEQ ID NO: 5), which is located in the C-terminal part of the domain B of said wild type glycoside hydrolase;

(6) the amino acid sequence ILRLDAVFLWK (SEQ ID NO: 6), which is located in the β-strand 4 of said wild type glycoside hydrolase;

(7) the amino acid sequence EAIV (SEQ ID NO: 7), which is located in the β-strand 5 of said wild type glycoside hydrolase;

(8) the amino acid sequence YVRCHDDI (SEQ ID NO: 8), which is located in the β-strand 7 of said wild type glycoside hydrolase;

(9) the amino acid sequence RISGTLASLAG (SEQ ID NO: 9), which is located in the domain B’ of said wild type glycoside hydrolase;

(10) the amino acid sequence GIPLIYLGDE (SEQ ID NO: 10), which is located in the β-strand 8 of said wild type glycoside hydrolase;
(11) the amino acid sequence RWVHRP (SEQ ID NO: 11), which is located in the loop 8 of the (β/α)₈-barrel,
and the sequence formed by said eleven motifs from said wild type glycoside hydrolase joined end-to-end from motif (1) to motif (11) has at least 65% sequence identity or at least 80% sequence similarity with the amino acid sequence SEQ ID NO: 12;
characterized in that said mutant has one or two mutation(s) consisting of,
when said mutant has only one mutation:
- the substitution of the amino acid residue at position 4 in said motif (4) with any amino acid selected from the group consisting of alanine (A), cysteine (C), glutamic acid (E), glycine (G), histidine (H), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), valine (V), tryptophan (W) and tyrosine (Y), or
- the substitution of the amino acid residue at position 5 in said motif (4) with any amino acid selected from the group consisting of leucine (L), methionine (M) and valine (V), or
- the substitution of the amino acid residue at position 8 in said motif (6) with any amino acid selected from the group consisting of glutamic acid (E), phenylalanine (F), glycine (G), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R) and valine (V), or
- the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V) and tryptophan (W), or
- the substitution of the amino acid residue at position 4 in said motif (7) with any amino acid selected from the group consisting of alanine (A), cysteine (C), aspartic acid (D), glycine (G), histidine (H), isoleucine (I), leucine (L), methionine (M), asparagine (N), serine (S), threonine (T) and tyrosine (Y), or
- the substitution of the amino acid residue at position 7 in said motif (8) with a valine (V), or
the substitution of the amino acid residue at position 1 in said motif 9
with any amino acid selected from the group consisting of alanine (A), cysteine (C),
phenylalanine (F), glycine (G), lysine (K), asparagine (N), glutamine (Q), serine (S),
threonine (T) and tryptophan (W), or
when said mutant has two mutations:
- the substitution of the amino acid residue at position 4 in said motif
5 (4) with an alanine (A) and the substitution of the amino acid residue at position 9 in
said motif (6) with a histidine (H), or
- the substitution of the amino acid residue at position 4 in said motif
10 (4) with a cysteine (C) and the substitution of the amino acid residue at position 5 in
said motif (4) with a leucine (L), or
- the substitution of the amino acid residue at position 4 in said motif
15 (4) with a lysine (K) and the substitution of the amino acid residue at position 9 in said
motif (6) with any amino acid selected from the group consisting of leucine (L) and
tryptophan (W), or
- the substitution of the amino acid residues at positions 4 and 5 in
20 said motif (4) respectively with a leucine (L), or
- the substitution of the amino acid residues at positions 4 and 5 in
said motif (4) respectively with a methionine (M), or
- the substitution of the amino acid residue at position 4 in said motif
25 (4) with a proline (P) and the substitution of the amino acid residue at position 9 in said
motif (6) with a cysteine (C), or
- the substitution of the amino acid residue at position 4 in said motif
(4) with a threonine (T) and the substitution of the amino acid residue at position 9 in
said motif (6) with any amino acid selected from the group consisting of histidine (H)
and lysine (K), or
- the substitution of the amino acid residue at position 4 in said motif
(4) with a valine (V) and the substitution of the amino acid residue at position 5 in said
30 motif (4) with any amino acid selected from the group consisting of leucine (L) and
methionine (M), or
- the substitution of the amino acid residue at position 4 in said motif
(4) with a valine (V) and the substitution of the amino acid residue at position 9 in said
motif (6) with any amino acid selected from the group consisting of histidine (H), lysine (K), arginine (R), and valine (V), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a histidine (H) and the substitution of the amino acid residue at position 9 in said motif (6) with a serine (S), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a proline (P) and the substitution of the amino acid residue at position 9 in said motif (6) with any amino acid selected from the group consisting of cysteine (C), isoleucine (I) and leucine (L), or

- the substitution of the amino acid residue at position 8 in said motif (6) with a threonine (T) and the substitution of the amino acid residue at position 9 in said motif (6) with a histidine (H).

8. A mutant according to claim 7, characterised in that the amino acid residue at position 9 in said motif (6) is substituted with any amino acid selected from the group consisting of cysteine (C), aspartic acid (D), isoleucine (I), lysine (K) and glutamine (Q), and more preferably with any amino acid selected from the group consisting of aspartic acid (D) and lysine (K).

9. A mutant according to claim 7 or claim 8, characterized in that said wild type glycoside hydrolase is an amylosucrase (EC 2.4.1.4) or sucrose hydrolase (EC 3.2.1.-).

10. A mutant according to claim 9, characterized in that said wild type glycoside hydrolase is an amylosucrase from *Neisseria pofysaccharea*, and is preferably selected from the group consisting of 1G5A, 1ZS2, IMVY, IMWO, 1S46, IJGI, 1MW2, 1MW3, IMWI and 1JG9 proteins.

11. A polynucleotide encoding a mutant of a glycoside hydrolase of anyone of claims 7 to 10.

12. A recombinant vector comprising a polynucleotide of claim 11.

13. A method for determining whether a wild type protein is a wild type glycoside hydrolase, said method comprising the steps of:

   a) determining the amino acid sequence of said protein,

   b) identifying in the amino acid sequence of said protein, preferably from the N- to C-terminus, eleven motifs defined by the following consensus motifs:
(1) the amino acid sequence LGVNYLHLMPL (SEQ ID NO: 1);
(2) the amino acid sequence DGGYAV (SEQ ID NO: 2);
(3) the amino acid sequence DFVFNH (SEQ ID NO: 3);
(4) the amino acid sequence LREIFPDTAPGNF (SEQ ID NO: 4);
(5) the amino acid sequence FNSYQWDLN (SEQ ID NO: 5);
(6) the amino acid sequence ILRLDAVAFWLK (SEQ ID NO: 6);
(7) the amino acid sequence EAI (SEQ ID NO: 7);
(8) the amino acid sequence YVRCHDDI (SEQ ID NO: 8);
(9) the amino acid sequence RISGTLASLAG (SEQ ID NO: 9)
(10) the amino acid sequence GIPLIYLGDE (SEQ ID NO: 10);
(11) the amino acid sequence RWVHRP (SEQ ID NO: 11);

determining the sequence identity percent or sequence similarity percent between the sequence formed by said eleven motifs joined end-to-end from
motif (1) to motif (11) with the amino acid sequence SEQ ID NO: 12, and if the
sequence identity percent is at least 65%, or if the sequence similarity percent is at least
80%, then the wild type protein is a wild type glycoside hydrolase.

14. A molecule selected from the following list:
- AHyI  α-D-glucopyranosyl-(l →4)-α-D-glucopyranosyl-(l →4)-2-
  acetamido-2-deoxy- α-D-glucopyranoside (XX2),
- Allyl  α-D-glucopyranosyl-(l →4)-2-acetamido-2-deoxy- α-D-
  glucopyranoside (XX3),
- Allyl 2,3,4,6-tetra-O-acetyl- α-D-glucopyranosyl-(l →4)-2-acetamido-
  3,6-di-0-acetyl-2-deoxy- α-D-glucopyranoside (XX4),
- Allyl  α-D-glucopyranosyl-(l →4)-2-amino-2-deoxy- α-D-
  glucopyranoside (XX5),
- Allyl  α-D-glucopyranosyl-(l →4)-2-amino-2-iV,3-O-carbonyl-2-
  deoxy-α-D-glucopyranoside (XX6),
- Allyl 2,4,6-tri-O-benzyl-α-D-glucopyranosyl-(l →4)-6-O-benzyl-2-
  benzylamino-2-iV,3-O-carbonyl-2-deoxy- α-D-glucopyranoside (XX7),
- Allyl 2,3,4,6-tetra-O-benzyl- α-D-glucopyranosyl-(l →4)-6-O-benzyl-
  2-benzylamino-2-iV,3-O-carbonyl-2-deoxy- α-D-glucopyranoside (XX8),

- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-O-benzyl-2-amino-2-N,3-O-carbonyl-2-deoxy-α-D-glucopyranoside (XX₁₈),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-O-benzyl-2-benzylamino-2-deoxy-α-D-glucopyranoside (XX₉),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-6-O-benzyl-2-benzylacetamido-2-deoxy-α-D-glucopyranoside (XX₁₀),
- Allyl 2-O-benzoyl-4-O-benzyl-3-O-chloroacetyl-α-L-rhamnopyranosyl-(1→3)[2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)]-6-O-benzyl-2-benzylacetamido-2-deoxy-α-D-glucopyranoside (XX₁₁),
- Allyl 2-<9-acetyl-3,4-di-O-benzyl-α-L-rhamnopyranosyl-(1→3)[2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)]-6-(9-benzyl-2-benzylacetamido-2-deoxy-α-D-glucopyranoside (XX₁₂),
- Allyl α-D-glucopyranosyl-(1→4)-2-deoxy-2-trichloroacetamido-α-D-glucopyranoside (XX₁₃),
- Allyl 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranoside (XX₁₄),
- 2₃,4,6-tetra-<9-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranose (XX₁₅),
- 2₃,4,6-Tetra-O-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranosyl trichloroacetimidate (XX₁₆),
- Allyl 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl-(1→4)-3,6-di-O-acetyl-2-deoxy-2-trichloroacetamido-α-D-glucopyranosyl-(1→2)-3,4-di-O-benzyl-α-L-rhamnopyranoside (XX₁₇),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-2-acetamido-3-O-acetyl-6-O-benzyl-2-deoxy-α-D-glucopyranoside (XX₂₄),
- Allyl 2,3,4,6-tetra-O-benzyl-α-D-glucopyranosyl-(1→4)-2-acetamido-6-O-benzyl-2-deoxy-α-D-glucopyranoside (XX₂₅).
$XX_3 + \text{glucose}$

$XX_2 (30\%)$

$XX_1$

$XX_3$

+$\text{fructose}$

$XX_4, R' = Ac$

$d, XX_3, R = H, R' = Ac (ED)$
FIGURE 2C
FIGURE 2E
Figure 3
### Table: Acceptor Glucosylation

<table>
<thead>
<tr>
<th></th>
<th>Conversion Rate (DP1)</th>
<th>% Glc transferred onto α-D-GlcpNAC-O allyl derivatives</th>
<th>% Monoglucosylated α-D-GlcpNAC-O allyl (DP2)</th>
<th>% Diglucosylated α-D-GlcpNAC-O allyl (DP3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASNPwt</td>
<td>6%</td>
<td>10%</td>
<td>45%</td>
<td>55%</td>
</tr>
<tr>
<td>F290D</td>
<td>87%</td>
<td>&gt;95%</td>
<td>75%</td>
<td>25%</td>
</tr>
<tr>
<td>F290K</td>
<td>80%</td>
<td>&gt;95%</td>
<td>69%</td>
<td>31%</td>
</tr>
</tbody>
</table>

### Figure 7A

Graph showing the conversion rates and glucosylation levels for different acceptors (ASNPwt, F290D, F290K) with varying % glucosylation.
<table>
<thead>
<tr>
<th>Enzymes</th>
<th>ASNPwt</th>
<th>F290K</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}/K_m$ (GF)</td>
<td>1.54</td>
<td>0.038</td>
</tr>
<tr>
<td>(nM·min⁻¹⁻¹)</td>
<td>0.074</td>
<td>3.63</td>
</tr>
<tr>
<td>$k_{cat}/K_m$ (Acceptor)</td>
<td>0.122</td>
<td>15.88</td>
</tr>
<tr>
<td>(nM·min⁻¹⁻¹)</td>
<td>287</td>
<td>6690</td>
</tr>
</tbody>
</table>
Figure 8.1
Figure 8.4
Figure 8.5

Figure 8.5
Figure 8.6
FESEPQTVTHS---TLQMPFK---AHDLIGEGTVSLNQDLTLPFQZYPVMW
FESEPQTVTHS---TLQMPFK---AHDLIGEGTVSLNQDLTLPFQZYPVMW
FSDGQTITQA---TLGFGPAF---AVDLGCMNPYREIGSSVRLQCPQRGMW
PESQAQIDEN---RRLRAGNRTGERFIDLYSQQTITSHHRTIPLFTPQVWM
PESDQTQEGN---KLRTAFGRPFVLNTWDFKRTYSECLVDIPYCLDWL
FESEPQVISAGN---LGLRHGFLGFLDFDLIQPQVSAPTADALTQPLQVPMW
FSDTQCHLEAR---KINMLIKRTEVSLUGFEGTVGTKELTRLELFYPMW
FEDITDAQUALR---HRLNHKLRTRKFTDFTGSSTPVWYTAQIKMLPDQYPMW
FESBQQTDEGR---HRLNHMLLRKTVDILQGQTVATKIAEHMLQGKYVMW
FESKHHQPPXS---CVDVILSSRSEQVDQLNQZEFYHQHIIFDADYWLN
FESQRQVIVSS---YTDGLASAQMVTDIASTTVKRTAVLVALGRLQVYMW
FESHEQKQISST---TLGLSSDWDVLGDSERVKPVAPQECZELAPYQVWM
FESPQHQLNRS---LALATLQASSEQLDLDNYDVDESSMLEAPYQVWM
KFELHLVHSLFLPKSLRLSIGLQTRAMNLNKEFDPDLSGDSLQFQIDLW
FESKEHHIRSLFH--QCCRALIDTCATDLTDFNLGSFRSFCSDGLDEYALDWLN
FEDNPQHLEQLQR---RSMQGFLDSLQVNGLYTGERFLFPNDQCVFAPQFY
FDCLKEFVEQLR---QLGFLESHLAMQATGRLVEQRQPQDFQAIAPQFY
FESNEKQYEWIN---NGQGQYKGLSQRLSGWQGAMCQKYGMAVMFLPMQY
FESNOKVAWIN---EDEOYDILDLHGRMVEQ’DPQFACQFY
FESNHITTTFPWTLQKLOGLHDVGELGES---PHLGGDLMELRFEAPRLDWLN
FESNEMLPLHLVRLSDLGHLVQOQDSGSA---FRKLDFVPFREDVYRALWLN
VTTKFRFPFPKKHAKRCSGDIYADGKLYQVFQDQFLGMP
LSTQCOVRQVLGVYQPLRHNLNDNS---LNIQDTEHAILDIACQVN
VSDAPQDFELPFEPAQFAQEDVDLW---APSAKPAGLSFYPIYGTV
FSDQOLVLELAVGVDWTLAIDAVDA---ASVDDVZSVDVPGYWMW
FSDQOLVLELAVGVDWTLAIDAVDA---ASVDDVZSVDVPGYWMW
FSDQOLVLELAVGVDWTLAIDAVDA---ASVDDVZSVDVPGYWMW
LVAR--
LEIA---
LEIA---
LEIA---
LETPQ---
LNPV---
LNPV---
LNPV---
LNPV---
LEPQ---
LARPQG---
LQVR---
LQVR---
LQVR---
LQVR---
LINAPQVP---
LQVP---
LYNN---
LSTQQ---
LMY---
LYEM---
LSBS---
LQ---
LQ---
LQ---
LQ---
LVAGG---
LTAGEPA---
LVQ---
LVRK---
LQSG---
LQSG---
LQ---
LQ---

Figure 8.8
<table>
<thead>
<tr>
<th>Acceptor</th>
<th>Acceptor Glucosylation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Glc transferred onto</td>
</tr>
<tr>
<td></td>
<td>acceptor derivatives</td>
</tr>
<tr>
<td>ASN Pwt</td>
<td>α-D-GlcNAc-OMe</td>
</tr>
<tr>
<td></td>
<td>α-D-GlcNAc-OMe</td>
</tr>
<tr>
<td>F290K</td>
<td>α-D-GlcNAc-OMe</td>
</tr>
<tr>
<td>ASN Pwt</td>
<td>β-D-GlcNAc-OMe</td>
</tr>
<tr>
<td></td>
<td>β-D-GlcNAc-OMe</td>
</tr>
<tr>
<td>F290K</td>
<td>β-D-GlcNAc-OMe</td>
</tr>
<tr>
<td>ASN Pwt</td>
<td>β-D-GlcNAc-OAlly</td>
</tr>
<tr>
<td></td>
<td>β-D-GlcNAc-OAlly</td>
</tr>
<tr>
<td>F290K</td>
<td>β-D-GlcNAc-OAlly</td>
</tr>
<tr>
<td>ASN Pwt</td>
<td>D-GlcNHTCA</td>
</tr>
<tr>
<td></td>
<td>D-GlcNHTCA</td>
</tr>
<tr>
<td>F290K</td>
<td>D-GlcNHTCA</td>
</tr>
</tbody>
</table>

**FIGURE 10**
FIGURE 11
A. LIBRARIES CONTAINING TWO VICINAL MUTATIONS (228-29, 289-90 and 330-31):

- Introduction of both mutations:
  - Template: wt with target vicinal mutations
  - PCR using degenerate primers (mixture of 276 (24G4) primers)
  - Digestion by CspI
  - Transformation into E. coli/TOP10

B. LIBRARIES CONTAINING TWO DISTANT MUTATIONS (228-290):

- Introduction of the first mutation:
  - Template: wt with target mutation (C289X or F290X)
  - PCR using degenerate primers (mixture of 19 primers)
  - Cleavage by CspI
  - Transformation into E. coli/TOP10 + plasmid extraction

- Introduction of the second mutation:
  - Template: C289X or F290X with target mutation F290X (or C289X)
  - PCR using degenerate primers (mixture of 19 primers)
  - Cleavage by CspI
  - Transformation into E. coli/TOP10

FIGURE 12