Exploration of the molecular determinants involved in alternansucrase specificity and stability

Manon Molina

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Institut National des Sciences Appliquées de Toulouse

---

Présentée et soutenue par

Manon MOLINA

Le 8 avril 2019

Exploration of the molecular determinants involved in alternansucrase specificity and stability

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Ingénieries microbienne et enzymatique**

Unité de recherche :
LISBP - Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés

Thèse dirigée par
Magali REMAUD-SIMEON et Sandrine MOREL

Jury

M. Marcelo Guerin, Rapporteur
M. Maher Abou Hachem, Rapporteur
Mme Marguerite Dols-Lafargue, Examinatrice
Mme Isabelle ANDRÉ, Examinatrice
M. Gianluca CIOCI, Examinateur
Mme Magali REMAUD-SIMÉON, Directrice de thèse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Institut National des Sciences Appliquées de Toulouse

---

Présentée et soutenue par

Manon MOLINA

Le 8 avril 2019

Exploration of the molecular determinants involved in alternansucrase specificity and stability

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Ingénieries microbienne et enzymatique**

Unité de recherche :

LISBP - Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés

Thèse dirigée par

Magali REMAUD-SIMEON et Sandrine MOREL

Jury

M. Marcelo Guerin, Rapporteur
M. Maher Abou Hachem, Rapporteur
Mme Marguerite Dols-Lafargue, Examinatrice
Mme Isabelle ANDRÉ, Examinatrice
M. Gianluca CIOCI, Examinateur
Mme Magali REMAUD-SIMÉON, Directrice de thèse

Last name: MOLINA **First name:** Manon

Title: Exploration of the molecular determinants involved in alternansucrase specificity and stability

Speciality: Ecological, Veterinary, Agronomic Sciences and Bioengineering. **Field:** Enzymatic and Microbial Engineering

Year: 2019 **Place:** Toulouse **Number of pages: 290**

The alternansucrase (ASR) from *Leuconostoc citreum* NRRL B-1355 is an α-transglucosylase belonging to the family 70 of glycoside hydrolases (GH70). In particular, the glucansucrases from this family use a cheap and abundant molecule, sucrose, to catalyze the formation of high molar mass (HMM) homopolymers of glucosyl units, called α-glucans, of interest for food, feed or health applications. ASR stands apart among these glucansucrases, being the only one to produce a particular α-glucan, the alternan, made of alternating α-1,6 and α-1,3 linkages in the main chain. Moreover, ASR produces a bimodal population of both HMM and lower molar mass (LMM) alternan. With a 45°C optimum temperature, ASR is also known as one of the most stable glucansucrases to date. However, in the absence of 3D structural data, most of the ASR determinants involved in linkage specificity, polymerization and higher stability remain unraveled. To get a deeper insight in ASR mechanism, we have solved the unliganded 3D structure of this enzyme at 2.8Å. Coupled to mutagenesis and molecular docking, our results shed the light on residues defining the +2 and +2' subsites (Trp675 and Asp772), in the prolongation of ASR subsites -1 and +1. The positioning of acceptor in either +2 or +2' subsite was submitted to control the linkage specificity in α-1,6 or in α-1,3 linkage formation respectively. Complexes of ASR with various sugar ligands were also obtained and highlighted a site on enzyme surface, in the catalytic domain, in which two amino acids (Tyr717 and Gln700) were shown to be involved in HMM alternan formation and in the alternan chain anchoring during polymerization. This site was characteristic of the ASR and could act as a bridge between the domain V and the catalytic site facilitating alternan processive elongation. Finally, a close inspection of the ASR structure combined to the comparison with the other GH70 3D structures available to date allowed the identification of amino acid insertion in loops of the domains A and C that could be involved in ASR stability. In particular, the ASR domain C was submitted to play a key role in stability, displaying higher level interactions than its counterparts. To evaluate this hypothesis, chimeras were constructed by domain C swapping between ASR and the branching sucrase GBD-CD2, a much less-stable enzyme. Preliminary results indicated the GBD-CD2 chimera to be more stable than the wild type, suggesting a role of the domain C in GH70 enzyme stability. In parallel, semi-rational and directed evolution of the ASR were initiated with the aim to obtain a variant with improved stability to temperature and organic solvents. Bringing all together, our results have deepened the knowledge of ASR structure, specificity, polymerization determinants and stability. Overall, it opens new paths of investigation for structure-function relationship studies of glucansucrases and for the conception of polymers with controlled structures and physicochemical properties.

Keywords: GH70, Glucansucrase, Alternansucrase, Crystal structure, Specificity, Stability

Doctoral school: SEVAB (Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries)

Laboratory: Laboratory of Biological Systems and Process Engineering (UMR CNRS 5504, UMR INRA 792), INSA de Toulouse

Nom: MOLINA **Prénom:** Manon

Titre: Exploration des déterminants moléculaires impliqués dans la spécificité et la stabilité de l'alternane-saccharase

Spécialité: Sciences Écologiques, Vétérinaires, Agronomiques et Bioingénieries. **Fillière:** Ingénierie microbienne et enzymatique

Année: 2019 **Lieu:** Toulouse **Nombre de pages: 290**

L'alternane-saccharase (ASR) de *Leuconostoc citreum* NRRL B-1355 est une α-transglucosylase de la famille 70 des glycoside hydrolases (GH70). En particulier, les glucane-saccharases de cette famille utilisent le saccharose, un substrat abondant et peu coûteux, pour catalyser la formation d'homopolymères d'unité glucose de haute masse molaire, appelés α-glucanes, trouvant des applications dans les domaines de l'agroalimentaire ou de la santé par exemple. L'ASR est une glucane-saccharase originale de par sa capacité à produire un α-glucane particulier, l'alternane, dont la structure est composée de liaisons osidiques α-1,3 et α-1,6 alternées dans la chaîne principale. De plus, l'ASR synthétise une population bimodale d'alternane de haute ou faible masse molaire. Avec une température optimale de 45°C, l'ASR est aussi connue comme étant l'une des glucane-saccharases les plus stables qui existent jusqu'à présent. Cependant, en l'absence de données structurales, la plupart des déterminants de la spécificité de liaison, de la polymérisation et de la plus haute stabilité de l'ASR restent incompris. Pour remédier à cela, nous avons résolu la structure de cette enzyme à 2,8Å. Combinant une étude par mutagénèse dirigée à du docking moléculaire, nos résultats montrent l'importance de résidus définissant les sous-sites +2 et +2' (Trp675 et Asp772), situés dans la prolongation des sous-sites -1 et +1. La spécificité de liaison serait contrôlée par le positionnement de l'accepteur dans l'un ou l'autre de ces sous-sites +2 ou +2' pour former une liaison α-1,6 ou α-1,3 respectivement. Des complexes de l'ASR ont également été obtenus avec différents ligands et ont concentré notre attention sur un site situé sur la surface de l'enzyme, dans le domaine catalytique, auquel appartiennent deux résidus (Tyr717 et Gln700) impliqués dans la formation de l'alternane de haute masse molaire et dans l'ancrage de la chaîne d'alternane pendant la polymérisation. Ce site caractéristique de l'ASR pourrait servir de pont entre le domaine V et le site catalytique facilitant l'élongation processive de l'alternane. Enfin, une analyse poussée de la structure de l'ASR associée à la comparaison avec les structures de GH70 disponibles à ce jour a permis de constater la présence d'insertions dans des boucles du domaine A et C, qui pourraient être impliquées dans la stabilité de l'ASR. En particulier, le domaine C pourrait jouer un rôle dans la stabilité, celui de l'ASR possédant plus d'interactions que les autres. Pour évaluer cette hypothèse, des chimères ont été construites entre l'ASR et l'enzyme de branchement GBD-CD2, une enzyme moins stable que l'ASR, par échange de leur domaine C respectifs. Les premiers résultats indiquent que la chimère de GBD-CD2 est plus stable que l'enzyme sauvage, suggérant en effet une implication du domaine C dans la stabilité des enzymes de la famille GH70. En parallèle, des approches semi-rationnelles et de l'évolution dirigée ont été initiées afin de faire évoluer l'ASR vers une plus grande stabilité à la température et aux solvants. Les résultats obtenus au cours de cette thèse approfondissent la connaissance des déterminants structuraux de l'ASR gouvernant sa spécificité de liaison, le mécanisme de polymérisation et sa stabilité. Globalement, ces travaux ouvrent des nouvelles pistes d'investigation concernant l'étude des relations structure-fonction des glucane-saccharases et la conception de polymères de structure et propriétés physico-chimiques contrôlées.

Mots-clés: GH70, Glucane-saccharase, Alternane-saccharase, Structure tridimensionnelle, Spécificité, Stabilité

École doctorale: SEVAB (Sciences Écologiques, Vétérinaires, Agronomiques et Bioingénieries)

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

Publications

Deciphering an undecided enzyme: investigations of the structural determinants involved in the linkage specificity of alternansucrase. Molina M., Moulis C., Monties N., Pizzut-Serin S., Guieysse D., Morel S., Cioci G., Remaud-Siméon M. *ACS Catal.,* **9** (2019). DOI: 10.1021/acscatal.8b04510.

A specific oligosaccharide binding site in domain A of alternansucrase is involved in alternan elongation. *In preparation*

Stabilization of the multi-domain branching sucrase GBD-CD2 by chimera construction using domain C swapping with alternansucrase. *In preparation*

Oral communication

Exploration of the molecular determinants involved in alternansucrase specificity & polymerization. Molina M., Moulis C., Pizzut-Serin S., Guieysse D., Morel S., Cioci G. & Remaud-Siméon M. 7th Symposium on the Alpha-Amylase Family, September 29- October 3, 2019, Smolenice Castle (Slovakia)

Posters

DSR-DP: a glucansucrase with catalytic efficiency boosted by organic solvents. Molina M., Claverie M., Vuillemin M., Cioci G., Guieysse D., Moulis C., Morel S. & Remaud-Siméon M. 12th Carbohydrate Bioengineering Meeting, April 23-26, 2017, Vienna (Austria)

DSR-DP: a glucansucrase with catalytic efficiency boosted by organic solvents. Molina M., Claverie M., Vuillemin M., Cioci G., Guieysse D., Moulis C., Morel S. & Remaud-Siméon M. EnzyBio 2017, May 9-12, 2017, Le Croisic (France)

Exploration of the molecular determinants involved in alternansucrase specificity & polymerization. Molina M., Moulis C., Pizzut-Serin S., Guieysse D., Morel S., Cioci G. & Remaud-Siméon M. 13th Carbohydrate Bioengineering Meeting, May 19-22, 2019, Toulouse (France)

Exploration of the molecular determinants involved in alternansucrase specificity & polymerization. Molina M., Moulis C., Pizzut-Serin S., Guieysse D., Morel S., Cioci G. & Remaud-Siméon M. 15th International Conference on Renewable Resources and Biorefineries, June 3-5, 2019, Toulouse (France)

Exploration of the molecular determinants involved in alternansucrase specificity & polymerization. Molina M., Moulis C., Pizzut-Serin S., Guieysse D., Morel S., Cioci G. & Remaud-Siméon M. 7th Symposium on the Alpha-Amylase Family, September 29- October 3, 2019, Smolenice Castle (Slovakia)

Remerciements

Merci tout d'abord aux professeurs Maher Abou Hachem et Marcelo Guérin, qui ont accepté d'être rapporteurs pour cette thèse. Un grand merci également au professeur Marguerite Dols-Lafargue qui a accepté d'être examinatrice pour cette thèse, ainsi qu'à la directrice de recherche Isabelle André pour avoir présidé ce jury. Les discussions ont été très enrichissantes et ont apporté un point de vue neuf et judicieux sur nos problématiques, le tout dans une très bonne ambiance et avec beaucoup de bienveillance. Je garde un excellent souvenir de cette journée de soutenance !

Je tiens ensuite à remercier chaleureusement Mag, ma directrice de thèse, pour son encadrement de qualité et tout le temps qu'elle a dégagé pour nos réunions, nos discussions et toutes les corrections. J'ai compris que rédiger était un art qui s'acquiert seulement après des années d'expérience ! Merci également Gianluca, pour ta décontraction en toutes circonstances et ta maîtrise de la purif, de la cristallo et de l'analyse en 3D. Je reste admirative de vos qualités respectives et de votre sens aiguisé de la recherche. Merci aussi aux trois autres encadrants sans qui ce travail n'aurait pas été possible : Claire pour tes encouragements, ton aide sur toutes les manips de biochimie... Et c'est grâce à toi que j'ai pu trouver cette thèse ! Sandrine pour tout l'accompagnement administratif et toutes les manips de RMN (il y en a eu des mutants à analyser !) et David également pour l'aide en RMN et en chimie. Vous avez tous apporté une plus-value à ma thèse et je vous remercie de votre confiance. Travailler avec vous a été très enrichissant et j'en garde un très bon souvenir.

Un grand merci également à l'équipe de Lionel Mourey à l'IPBS pour leur accueil, l'accès à la plateforme cristallo et toutes les collectes effectuées. Un merci particulier pour Sam, sa gentillesse et son avis expérimenté sur les manips de DSF. Enfin, merci à la plateforme ICEO pour l'accès aux plateformes de purification et d'analytique et notamment merci à Sandra, Nelly et Sophie pour leur aide expert.

Un immense merci aux collègues qui ont rendu la vie au labo plus belle et plus sympa chaque jour, chacun à leur façon, et qui m'ont aidé à un moment ou un autre : Maher, Mounir, Julien D., Emna, Sabine, Jiao, Manon C, Jérémy, Julian, Coco, Hayang, Gema, Azucena, Laure, Florent, Arthur, Mathieu, Pauline, Eli, Victor, Nathalie, Alex, Emma, Audrey, Christian, Marie, Vincianne, Zhongpeng, Julien R., Alexandra, Tarun, Benoit, David, Jelena, Akli, Gaby, Florence, Elisabeth, Sophie B., Sophie D., Claire D., Zhi, Dorian, Valentin, Vincent, et j'en oublie sûrement... Un merci particulier pour Thomas, stagiaire de quelques mois seulement qui a prouvé son efficacité et son intelligence.

Un petit mot également pour les équipes d'IC de Ramonville : Lolo, Ben, Panthère, Clo, Yannou, Martin, Jeanne, Antho, Clément, Claire ; c'était à chaque fois une joie de vous retrouver et une bouffée d'air frais !

Merci à ma famille pour leur soutien indéfectible et tous les bons moments passés ensembles : mes parents, mes sœurs Sophie et Coralie, sans oublier mes grands-parents de Rodez et ma grand-mère de Sainte-Colombe, la famille Jouve, mon parrain Patrice, ainsi que toute la famille Cépède. C'est grâce à vous que le moral tient et remonte dans les moments difficiles. Et merci de votre soutien et de vos pensées lors de ma soutenance !

Et enfin, merci à toi Maher d'avoir toujours été là, aux bons et aux mauvais moments, de m'avoir toujours encouragée et d'avoir cru en moi (plus que moi !).

Bonne lecture !

Table of Contents

Introduction

Polymers are used as basic materials in the manufacture of **innumerable** daily life **products**, as well as more sophisticated compounds in medicine, diagnostics and fine chemistry. Most of them are still today derived from petroleum-based chemistry. However, **environmental** and **sustainability** concerns push forward the research on renewable and recyclable bio-derived structures showing equivalent, new or improved physico-chemical properties. Among bio-sourced polymers, polysaccharides from plant, algae, fungi or microbial organisms are seen nowadays as promising options to replace **petroleum based products** with novel architectures. Many of them, and in particular the **plant** and **algal** polysaccharides such as starch, celluloses, pectins, guar and their derivatives are already at the basis of well-established bio-based industries with a wide range of applications in food products, pulp and papers, chemicals, pharmaceuticals, cosmetics and bioenergy industries (Prestegard et al., 2015). These polysaccharides still dominate the world market, but **microbial** polymer market, although more modest, is continuously growing. Indeed, the diversity found in the microbial world in regard to monomer composition, type of linkages, substituents and molar mass is impressive and offers a high potential for innovation (Freitas et al., 2011; Moscovici, 2015; Roca et al., 2015). The production of microbial polymers is also independent of environmental and climate fluctuations, and downstream processing are generally easy, which is another clear advantage. However, only a **few microbial polymers** are commercialized to date, including xanthan, gellan, hyaluronic acid, welan, clavan, fucopol, pullulan, and dextran.

In particular, **dextran** has been the first microbial biopolymer to be produced on an industrial scale and commercialized since 1948 as plasma substitute (Heinze et al., 2006; Leathers, 2005; Vettori et al., 2012). It is a homopolysaccharide of glucosyl units (α-glucan) produced from sucrose by lactic acid bacteria such as *Leuconostoc mesenteroides* NRRL B512-F. Overall, the panel of α-glucans produced by lactic acid bacteria from this cheap and abundant agro-resource is very wide and varying a lot in terms of **size**, **type** and arrangement of α-osidic **linkages** and degree of **branching** according to the bacterial strains and the enzymes they contain. All these factors define the physicochemical, biological, and mechanical properties of each α-glucan, and therefore its range of potential applications.

The bacterial enzymes at the origin of α-glucan production are called **glucansucrases** and are classified in the **family 70** of Glycoside Hydrolases (Lombard et al., 2014). They are efficient transglucosylases that do not require expensive nucleotide-activated sugars (NDP-sugar) as substrate or cofactors for the synthesis of very high molar mass polymers (from 10^3 to 10^9 g/mol) (Leemhuis et al., 2013; Monsan et al., 2010). Moreover, these enzymes are also able to perform a **transglucosylation reaction** from sucrose onto an exogenous acceptor, that can be a sugar or an

aglycone hydroxylated molecule (Koepsell et al., 1953; Monsan et al., 2010), leading to the production of interesting prebiotic oligosaccharides or gluco-conjugates (Leemhuis et al., 2013).

In the past decades, considerable progress has been made in the characterization of the structure-function relationships of GH70 α-transglucosylases, thanks to the identification, cloning and characterization of **atypical glucansucrases** found in natural diversity, as well as the resolution of the first 3D structures, since 2010 (4 to date in the CAZy database). Characterization at molecular level has long been difficult due to the large size of these enzymes (150 to 300 kDa), complicating heterologous productions in *E. coli*, protein purification and crystallisation assays. Notably, some amino acids involved in glycosidic linkage specificity of certain enzymes have been identified, as well as some determinants governing the size of the produced polymers.

However, we are still far from having identified all the residues that would allow for a tight control of α-glucan size and structure, with the aim of being able to produce tailor-made α-glucans. Another major bottleneck of glucansucrase widespread use is their lack of **robustness** and **stability**. Indeed, all of these enzymes known to date are originated from **mesophilic** lactic acid bacteria, and to our knowledge, **no thermostable** variants have been found in Nature or succesfully engineered. Yet, enzyme stability is crucial for the optimization of enzyme-based processes. Thus, trying to **understand** and **improve** glucansucrase **stability** determinants is of the utmost importance for future processes development.

To participate to the effort of a better understanding of GH70 enzyme structure-function relationships, we decided to choose –after a first screening of a collection of GH70 enzymes available in the lab[*]- the **alternansucrase** (ASR) from *Leuconostoc citreum* NRRL B-1355 as model of study to identify the determinants involved in its product specificity (osidic linkages and polymer size) as well as those involved in stability. Indeed, ASR is a long known glucansucrase producing, from sucrose, a peculiar α-glucan named **alternan** that contains **alternated α-1,6 and α-1,3 linkages** in the main chain, a **unique** feature among all the α-glucans described to date. The mode of ASR linkage **alternation** has not been elucidated yet, constituting an intriguing mechanistic question at molecular level. Moreover, this enzyme is one of the **most stable** known glucansucrases, even if relatively low (half-life time of 6 hours at 40°C). Identifying the structural features that could explain it –in comparison with less stable glucansucrases- could open the route for an improvement of this stability by rational engineering.

[*] See Annex I for screening results, p. A1

In parallel, a strategy of directed evolution was also initiated, by screening ASR libraries of mutants onto high amount of organic solvents such as DMSO or at higher temperatures.

In this context, the Chapter I of this manuscript is a literature review dividing in two parts; the first one presents an overview of the glucansucrases and the second part is an extensive review of the ASR and alternan discovery, characterization and uses.

The three other chapters are presented in the form of scientific articles. The first two aim to decipher ASR mechanism and investigate the determinants responsible for alternated linkages in the polymer main chain (Chapter II) and high molar mass alternan formation (Chapter III). The last chapter focuses on the ASR improvement of stability by both random and rational strategies, including an investigation of the potential role of the domain C in this property (Chapter IV).

Chapter I

Glucansucrases

&

Alternansucrase

a literature review

Part 1- Glucansucrases: an overview

I. Glucansucrases: origin, role and classification

I.1. Producing organisms and physiological role

Glucansucrases from the family 70 of glycoside hydrolases are α-transglucosylases produced by **lactic acid bacteria (LAB)**, a group of Gram positive carbohydrate-fermenting bacteria. To date, they have been found in several LAB genera including *Lactobacillus*, *Leuconostoc*, *Streptococcus* (Sidebotham, 1974), and more recently, *Weissella* (Kang et al., 2009) or *Oenococcus* (Dimopoulou et al., 2014); and only in mesophilic strains. The polysaccharides synthesized by glucansucrases from sucrose were initially identified in sucrose-containing foods and beverages, where they caused **thickening**. These rheological changes were first attributed to microbial action by Louis Pasteur in 1861. Van Tieghem identified the causative agent in 1878 and named it *Leuconostoc mesenteroides*. In 1940, the first glucansucrase was characterized by Hehre and collaborators from a sucrose-broth of *Ln. mesenteroides* and shown to synthesize a linear α-1,6 linked **homopolymer of glucosyl units** named dextran, concomitantly with fructose release (Hehre, 1941; Hehre and Sugg, 1942).

Later, the diversity of homopolysaccharides of glucosyl units, generally called **α-glucans**, produced by LAB from *Acetobacter, Betabacterium, Leuconostoc, Streptobacterium and Streptococcus* was highlighted in the work of Jeanes and collaborators (Jeanes et al., 1954). Ninety-six different strains were grown on sucrose. The synthesized polymers were isolated and characterized. The study evidenced that the **sizes**, **viscosity** and type of **osidic linkages** of the α-glucans could vary with the producing strain and that one single strain could synthesize different α-glucans, indicating that several glucansucrases were likely produced. This was later confirmed by enzyme biochemical characterization (Côté and Robyt, 1982a, 1983; Sato et al., 1984; Smith et al., 1998), gene cloning (Aoki et al., 1986) and genome sequencing (Biswas and Biswas, 2012; Kim et al., 2008; Laguerre et al., 2012; Passerini et al., 2014, 2015). For example, five glucansucrases (DSR-A, DSR-B, DSR-E, DSR-M and DSR-DP) were identified in *Ln. citreum* NRRL B-1299 (Passerini et al., 2015), a strain producing α-glucans with around 70% of α-1,6 linkages and 30% of α-1,2 linkages (Kobayashi and Matsuda, 1977).

Glucansucrases are usually **secreted** in the culture broth. Depending on the LAB-producing strain, their expression can be **constitutive** (for most glucansucrases from *Streptococcus* sp.) or **induced by sucrose** (for most glucansucrases from *Leuconostoc* sp.) (Monsan et al., 2001). Notably, mutant

strains of *Leuconostoc* were modified to produce glucansucrases constitutively (Kim and Robyt, 1994, 1995; Mizutani et al., 1994). Their main physiological function seems to be linked with their ability to catalyze **extracellular high molecular mass** (HMM) polymer synthesis from sucrose. These polymers participate in **biofilm** formation and can protect the bacteria against environmental stresses (*e. g.* dessication, biocides, antibiotics, phagocyte attack). They can also mediate the adhesion to surfaces (Flemming and Wingender, 2010). In particular, the α-glucans formed by the *Streptococcus* genus are very well-studied due their contribution to the dental plaque formation, the tooth surface colonization and the development of dental caries (Loesche, 1993).

I.2. Classification

I.2.1. Glucansucrases

Glucansucrases are classified in the family 70 of Glycoside Hydrolases (**GH70**) and in the **GH-H** clan according to the CAZy (Carbohydrate-Active enZymes) classification, that lists and categorizes enzymes able to "assemble, modify or breakdown oligo- and polysaccharides" (Lombard et al., 2014). The GH-H clan, formerly known as the α-amylase superfamily, gathers enzymes from families 13, 70 and 77 of glycoside hydrolases that all possess a catalytic $(\beta/\alpha)_8$ barrel. Notably, the $(\beta/\alpha)_8$ barrel version of glucansucrases is **circularly permutated** (MacGregor et al., 1996). Glucansucrases are active on sucrose whereas GH13 and GH77 family enzymes are mainly starch-modifying and 4-α-glucanotransferases, respectively (Leemhuis et al., 2013). Glucansucrases are **multi-domain large** enzymes. They usually contain more than **1,600 amino acids** on average, compared to only 600 amino acids for the enzymes from GH13 and GH77 families (Monchois et al., 1999b). To date, 565 enzyme sequences are classified in the GH70 family. More than 88% of them are putative enzymes. They mostly come from *Streptococcus* sp. (45.8%), *Lactobacillus* sp. (29.2%), *Leuconostoc* sp. (20.4%) and *Weissela* sp. (5.3%). Other genus identified as GH70-producers represent less than 3% of the listed enzymes and are thus marginal (0.2% for *Azotobacter, Burkholderia, Fructobacillus, Geobacillus* and *Oenococcus*; 0.4% for *Bacillus*, 0.5% for *Exiguobacterium* and 0.7% for *Pediococcus*). 66 enzymes out of these 565 are characterized; among them the glucansucrase enzymes are clearly predominant: 59 glucansucrases (GS) are listed, for only six glucanotransferases and one branching sucrase (CAZy database, November 2018) (see pages 9-10 for a description of the different activities found in this family).

The glucansucrases can be distinguished by the structure of the polymers they produce from sucrose. Indeed, the α-glucans can vary a lot in terms of osidic linkages as well as of degree, length and arrangement of branching. **Dextransucrases** (DSR; E.C.: 2.4.1.5) synthetize mainly α-1,6 linked dextrans displaying various amounts of α-1,2 or α-1,3 linked branching points; **reuteransucrases**

(2.4.1.-) synthetize reuteran comprising both α-1,6 and α-1,4 glucosidic linkages and **mutansucrases** (2.4.1.125) catalyze the formation of mutan, an α-glucan containing mainly α-1,3 linked glucosyl units (Figure 1) (Remaud-Siméon et al., 2000; van Hijum et al., 2006). Finally, a fourth category of glucansucrases is reported, namely the **alternansucrases**, the enzyme at the heart of our research work, to which we will dedicate a specific part (see Part 2 - the story of alternansucrase, page 46).

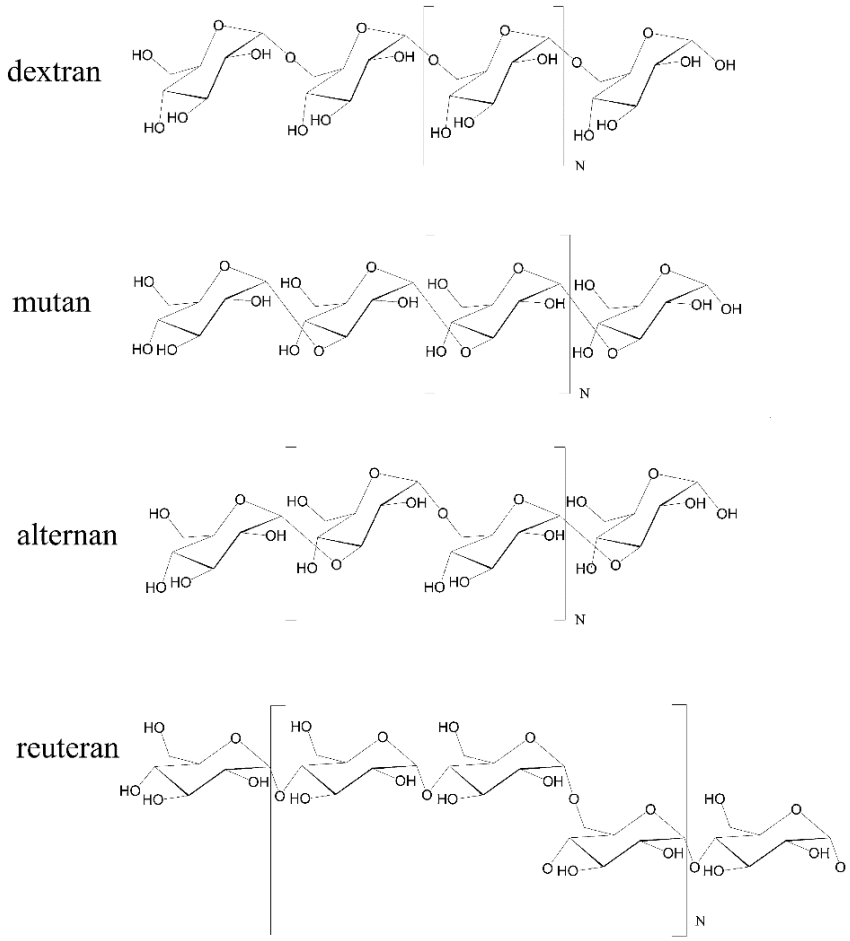


Figure 1: "The basic structures of the α-glucans synthesized by glucansucrases. The α-glucans are classified according to the dominant linkage type in the main chain." From (Leemhuis et al., 2013).

I.2.2. Branching sucrases

In **2002**, the first gene coding for the **bifunctional** glucansucrase from *Ln. citreum* NRRL B-1299 was isolated and sequenced. The recombinant enzyme, DSR-E, was produced in *E. coli*. This enzyme is the largest glucansucrase described with 2,835 residues and a predicted molar mass of 313,000 g/mol and the first one distinguished by the presence of two catalytic domains (Bozonnet et al., 2002). The first catalytic domain (CD1) is dedicated to the formation of dextran with 86% α-1,6 linkages, 11% α-1,3 linkages and 3% α-1,4 linkages whereas the second one (CD2) is specific for the α-1,2 branching synthesis (Fabre et al., 2005). From sequence analyses, both domains were placed in the GH70 family. The truncation of the first 1141 residues led to the construction of a new enzyme called GBD-CD2, the first engineered branching sucrase described. Indeed, it cannot act as a polymerase but is specific for the introduction of α-1,2 linked branches into linear exogenous dextran molecules, using sucrose as glucosyl donor (Fabre et al., 2005).

Since **2002,** other natural branching sucrases have been discovered essentially through genome mining. Like GBD-CD2, they use sucrose as a substrate but cannot perform the HMM **α-glucan** synthesis. An acceptor *e.g.* dextran has to be added in the medium (Moulis et al., 2016). To date, two types of branching sucrases have been identified, both originated from *Leuconostoc* or *Lactobacillus* sp.: **α-1,2** branching sucrases catalyzing the formation of α-1,2 glucosidic linkage (GBD-CD2 (Brison et al., 2012), BRS-A (Passerini et al., 2015), BRS-D (Vuillemin et al., 2016)) and **α-1,3** branching sucrases specific for α-1,3 osidic linkage formation (BRS-B, BRS-C (Vuillemin et al., 2016)). Of note, **another bifunctional** enzyme was recently reported in *Lactobacillus kunkeei* DSM 12361. Both the full-length enzyme, named GtfZ (Met1-Lys2621), and the truncated version GtfZ-CD1 (Asn50-Ile1209) could not be successfully overproduced in *E. coli*. A truncation similar to that performed in DSR-E enabled the generation of an engineered α-1,3 branching sucrase GtfZ-CD2 (Asp1212-Asp2264) (Meng et al., 2018).

I.2.3. Glucanotransferases

In **2004**, the GTFB enzyme from *Lactobacillus reuteri* 121 was found to be inactive on sucrose, yet displaying a sequence very similar to those of glucansucrases, justifying its classification in the GH70 family (Kralj et al., 2004a). Further characterization led to the conclusion that this enzyme - as other GtfB-like enzymes discovered since then- use **maltodextrins or starch** as substrate to catalyze the hydrolysis of α-1,4 glycosidic linkages from the non-reducing end of a donor and synthetize α-1,6 linkages (**4,6-α-glucanotransferases**) (Kralj et al., 2011) or α-1,3 linkages (**4,3-α-glucanotransferases;** *Lactobacillus fermentum* NCC 2970) (Gangoiti et al., 2017b). These enzymes possess a circularly permuted $(\beta/\alpha)_8$ barrel (Kralj et al., 2004b), like glucansucrases.

Recently, putative genes of GH70 family enzymes have been identified in uncommon organisms (*Exiguobacterium, Azotobacter, Bacillus, Paenibacillus, Geobacillus*) and the corresponding enzymes were found to be **glucanotransferases** as GtfB-like enzymes, but with a different domain organization (not the circular permutation and absence of one domain) likely to **GH13** family enzymes (see III.1. From GH13 to GH70 family, page 29). These enzymes also use maltodextrin or starch as substrate to catalyze the formation of α-1,6 linkages to whether isomalto/malto-oligosaccharides (for GtfC from *Exiguobacterium sibiricum* 255-15) (Gangoiti et al., 2015) or reuteran-like polymer (for GtfDs from *Azotobacter chroococcum* NCIMB 8003 and *Paenibacillus beijingensis* DSM 24997) (Gangoiti et al., 2016, 2017a). To that date, these α-glucanotransferases are the only GH70 family enzymes to be found in non-lactic acid bacteria (*e.g. Azotobacter, Paenibacillus)*.

II. Mechanism and Structure

II.1. Glucansucrase reaction mechanism

Unlike Leloir type-glycosyltransferases, glucansucrases do not require expensive Nucleotide-activated sugars (NDP-sugar) as substrate (Monsan et al., 2010). They also do not need cofactors and simply use **sucrose** as a glucosyl donor, an abundant and low-cost resource (Monchois et al., 1999b). Glucansucrases adopt a two-step **α-retaining mechanism** (Koshland, 1953) (Figure 2) involving the formation of a covalent β-glucosyl-enzyme intermediate (Mooser et al., 1991; Mooser and Iwaoka, 1989), which can further react with different types of acceptors. Three catalytic residues are essential for catalysis: an aspartate, a glutamate and a second aspartate residue playing the role of a **nucleophile**, an **acid/base** catalyst and a **transition state stabilizer** (TSS), respectively (Leemhuis et al., 2013).

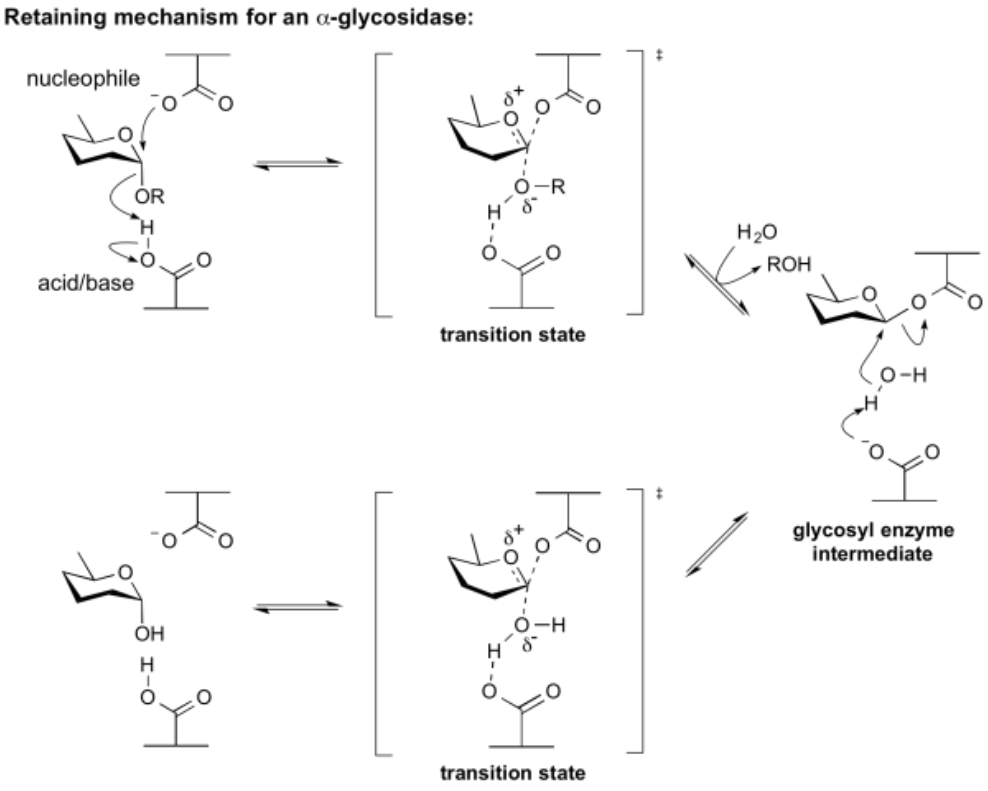


Figure 2: Scheme of the Koshland mechanism. From (Withers and Williams, 2007).

II.1.1. The formation of the glucosyl-enzyme intermediate

The first step is known as the glycosylation step and corresponds to the cleavage of the α-1,2-β osidic linkage of sucrose, that occurs between subsites -1 and +1 of the enzyme (Figure 3). To do so, a nucleophilic attack is exerted by the nucleophile on the anomeric carbon of the glucosyl unit thanks to the assistance of the acid/base catalyst that gives its proton to the fructosyl moiety. This allows fructose release and the formation of an oxocarbenium transition state leading to the covalent β-glucosyl-enzyme intermediate.

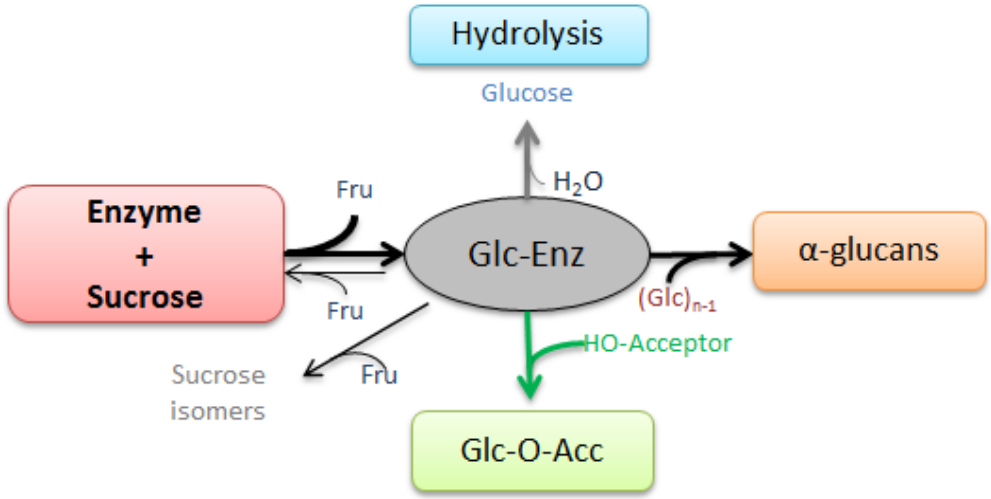


Figure 3: Scheme of the main three types of reaction catalyzed from sucrose by glucansucrases.

II.1.2. The transfer on acceptor molecules

The de-glucosylation step corresponds to the transfer of the glucosyl unit, depending on enzyme affinity for the different molecules present in the reaction medium that can act as acceptor. There is indeed a competition (Luzio and Mayer, 1983) between (Figure 3):

- the non-reducing end of a growing **glucan chain** (**polymerization reaction**) (Moulis et al., 2006)
- the **water** molecule (**hydrolysis reaction**), a minor reaction relative to each enzyme (see Table 1 page 15 for examples). The more the proportion of glucosyl units transferred to water (hydrolysis percentage) is high, the less the enzyme is an efficient transglucosylase.
- the **fructose** released from sucrose cleavage (**sucrose isomers formation**). The molecules that can be formed are leucrose (α-D-Glc*p*-(1→5)-β-D-Fru*p*) and isomaltulose (α-D-Glc*p*-(1→6)-β-D-Fru*f).* Glucosyl transfers onto fructose are predominant at the end of the reaction, when fructose is in a large excess (Moulis et al., 2006).
- the **glucose** released from hydrolysis

- the **sucrose** (β-D-Fru*f*-2↔1-α-D-Glc*p*) molecule itself. In 2008, isomelezitose (α-D-Glc*p*-(1→6)- β-D-Fru*f*-2↔1-α-D-Glc*p*) was isolated for the first time produced by alternansucrase. It is not understood yet if it arose from sucrose or isomaltulose glucosylation (Côté et al., 2008). Later, isomelezitose was found to be produced by many glucansucrases (eleven strains tested) in different yield (Côté and Skory, 2017).
- any hydroxylated acceptor that can be added in the reaction medium (**acceptor reaction**). One example is the pioneer work of Koepsell *et al.* that found certain sugars (isomaltose, maltose, glucose, α-methyl glucoside) to act as glucosyl acceptor when added in the reaction medium whereas others (fructose, leucrose, melibiose and galactose) were found to be less good acceptors for *Ln. mesenteroides* NRRL B-512 dextransucrase (Koepsell et al., 1953). In contrast, much more recently, melibiose was found to be as good acceptor as isomaltose for ASR from *Ln. citreum* NRRL B-1355 (Côté et al., 2003). This reflects the **diversity** of glucansucrase **catalytic site shapes** that determines acceptor affinity.

II.1.3. Initiator molecules for glucan formation

All these small molecules (sucrose, glucose, sucrose isomers) can serve as **initiator** of polymerization and can be used in turn as acceptor to yield oligosaccharides and α-glucans. As examples, for dextransucrase DSR-S from *Ln. mesenteroides* NRRL B-512F and ASR from *Ln. citreum* NRRL B-1355, the first products of reaction were found to be fructose, glucose and small oligosaccharides that were mainly elongated from sucrose molecules. Thus, both **glucose** coming from sucrose hydrolysis and **sucrose** itself act as acceptor for α-glucan synthesis initiation. Glucose is preferred when present in sufficient amount and then, the transfer of glucosyl residues to produce oligosaccharides of higher DP and HMM dextran or alternan is favored, resulting in polymer formation (Moulis et al., 2006). However, leucrose and isomaltose were detected earlier for ASR, suggesting that this enzyme recognize glucose and fructose as acceptors more rapidly than DSR-S (Moulis et al., 2006). Similarly, both the dextran produced by GTF-S3 from *S. sobrinus* and the reuteran produced by GTFA from *L. reuteri* 121 were found to have a sucrose molecule at the terminal-end (Cheetham et al., 1991; Dobruchowska et al., 2013). In contrast, for mutansucrase GTF-I from *S. sobrinus*, sucrose is first hydrolyzed, leading to glucose and fructose released. Only the **glucose** is then used as acceptor to form nigerose, and glucosyl transfer result in nigero-oligosaccharides formation of increasing DP (Komatsu et al., 2011).

II.1.4. Mode of glucan elongation

Following the initiation phase, the **mode of α-glucan elongation** can be processive (single chain mechanism) or non-processive (multi chain mechanism) by addition of the glucosyl units at the non-reducing end of acceptor molecules (see Table 1 page 15 for examples) (Moulis et al., 2006). Both DSR-S and ASR were described to be semi-processive transglucosylases because of the presence of both (i) a glucan polymer with a maximum size (processive mode) and (ii) oligosaccharides varying in DP (non-processive mode) (Moulis et al., 2006). DSR-S dextran was first detected and reached its maximum size after only 45 minutes (equivalent to 23% sucrose consumption). In contrast, GTF-I mutansucrase or DSR-M dextransucrase were described to be non-processive (Claverie et al., 2017; Komatsu et al., 2011).

II.1.5. Glucan populations

Depending on the enzyme, one, two, or more **populations of α-glucan** can be formed that vary in size from one enzyme to another (see Table 1 page 15 for examples). For instance, DSR-M only forms one single population of low molar mass (LMM) dextran whereas DSR-DP forms five different populations out of them one is of high molar mass (HMM). At high sucrose concentration, the processivity is lost (Meng et al., 2015a) and both low molar mass (LMM) glucan production and acceptor reaction onto sucrose are favored over HMM glucan production (Moulis et al., 2006).

II.1.6. Glucan diversity

When Jeanes *et al.* analyzed the glucans produced by 96 strains of LAB, they classified them in three different classes: A (0-2% of α-1,3 linkages), B (3-6% of α-1,3 linkages), C (>6% of α-1,3 linkages) or "Heterogeneous" glucans when different fractions were isolated. The observed nature of the glucan was also reported and ranked from gum, flocculent to fine powders that can be long (elasticity), short, crumbly, stringy, smooth, pasty or fluid (Jeanes et al., 1954). This reflects again the huge **variety** of α-glucan **size** and **structure**, even when comparing two dextransucrases or two reuteransucrases (see Table 1 page 15) and despite high similarities in the catalytic domain with the conserved residues in the motifs (see II.2.1. Conserved residues of the N-terminal catalytic domain, page 16 for motif sequences). This suggests the specificity not to be governed by these conserved residues and a much more **complex mechanism** involving the **interplay** of many residues that will be further discussed in the structure-function part (see III.3 Structure-function relationships, page 34).

Table 1: Illustration of the diversity of the α-glucan produced by a set of glucansucrases. References: 1= (Brison et al., 2012); 2= (Passerini et al., 2015); 3= (Claverie et al., 2017) ; 4= (Moulis et al., 2006) ; 5= (Joucla et al., 2006); 6= (Meng et al., 2016c); 7= (Kralj et al., 2004b); 8= (Dobruchowska et al., 2013); 9= (S. Kralj et al., 2005); 10= (Meng et al., 2014, p. 940); 11= (Vuillemin et al., 2018); 12= (Claverie et al., 2019b)

Enzyme name	GBD-CD2	DSR-DP	DSR-MΔ2	DSR-S	ASR C-	GTFA-ΔN	GTFO-ΔN	GTF180	DSR-OK
Organism	*Ln. citreum* NRRL B-1299	*Ln. citreum* NRRL B-1299	*Ln. citreum* NRRL B-1299	*Ln. mesenteroides* NRRL B-512F	*Ln. citreum* NRRL B-12355	*L. reuteri* 121	*L. reuteri* ATCC 55730	*L. reuteri* 180	*O. kitaharae* DSM 17330
Type	Branching sucrase	Glucansucrase	Glucansucrase	Glucansucrase	Glucansucrase	Glucansucrase	Glucansucrase	Glucansucrase	Glucansucrase
Specificity (polymer)	α-1,2	>90% α-1,6 dextransucrase[2]	100% α-1,6 dextransucrase[2]	100% α-1,6 dextransucrase	58% α-1,6 / 20% α-1,3[5] alternansucrase	42% α-1,6 / 58% α-1,4 reuteransucrase[6]	21% α-1,6 / 79% α-1,4 reuteransucrase[6]	67% α-1,6 / 33% α-1,3 [10]	dextransucrase
Residues following the TSS (Motif IV)	^{2323}KGV^{2325}	^{571}AES^{573}	^{791}SEV^{793}	^{663}SEV^{665}	^{768}YDA^{770}	^{1134}NNS^{1136}	^{1134}NNS^{1136}	^{1137}SNA^{1139}	^{574}SEV^{576}
Hydrolysis (%)	87[1]	n.d.	3[3]	2[4]	5[4]	9[6]	43[6]	24[10]	2[12]
Number of populations from sucrose (except DP1 and DP2)	0	5[2]	1[2]	2[4]	2[4]	2[7]	2[9]	1	1
Structural characteristics	/	LMM1= 580 g/mol[2] LMM2= 2000 g/mol[2] LMM3= 4240 g/mol[2] LMM4= 9300 g/mol[2] HMM= >2,000,000 g/mol[2]	LMM= 27,000 g/mol[2]	LMM= DP2-25[4] (4,000 g/mol) HMM= > 10,000,000 g/mol[4]	LMM= 1,300 g/mol[5] HMM= 1,700,000 g/mol[5]	LMM HMM= 48,000,000 g/mol[7]	LMM HMM= 42,000,000 g/mol[9]	HMM= 22,600,000 g/mol[10]	HMM >1,000,000,000 g/mol[11]
Processivity	/	n.d.	non-processive[3]	semi-processive[4]	semi-processive[4]	semi-processive[8]	n.d.	n.d.	semi-processive[12]

II.2. Glucansucrase primary structure

In **1987**, the first glucansucrase sequences from *Streptococcus mutans* GS-5 (GTF-B*)* (Shiroza et al., 1987), *Streptococcus mutans* LM7 (GTF-C*)* (Pucci et al., 1987) or *Streptococcus sobrinus* MFe28 (GTF-I) (Ferretti et al., 1987) were released.

Analysis of the first 14 available sequences enabled to submit an organizational scheme including four different parts: a **signal peptide**, a N-terminal **variable region,** a **catalytic domain** and a C-terminal **glucan binding domain** (Figure 4) (Monchois et al., 1999b).

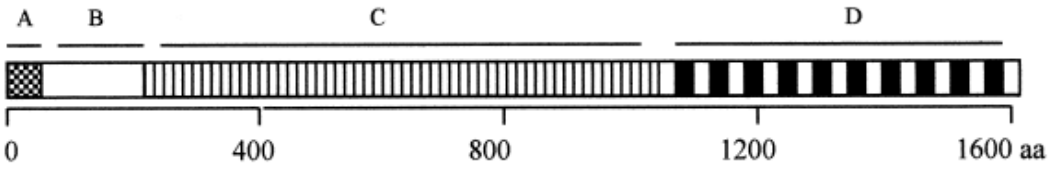


Figure 4: Schematic structure of glucansucrases for which encoding genes have been cloned. A, signal peptide; B, variable region; C, N-terminal catalytic domain; D, C-terminal glucan binding domain. From (Monchois et al., 1999b).

II.2.1. Conserved residues of the N-terminal catalytic domain

The catalytic nucleophile was first identified thanks to the isolation of **glucosyl-enzyme** complex of *S. sobrinus* GTF-S and GTF-I (Mooser et al., 1991; Mooser and Iwaoka, 1989).

Sequence analysis and comparison with related sequences of enzymes from the GH13 family enabled the identification of the putative catalytic **nucleophile**, acid/base catalyst and **transition state stabilizer** and also revealed that GH13 and GH70 family were related and shared common sequence **motifs** (motifs I to VII) in the catalytic domain, but ordered differently along the sequence due to the **circular permutation** event that may have occurred during evolution (MacGregor et al., 1996; Mooser et al., 1991). The corresponding sequence alignment of several glucansucrases of known 3D structures is presented in Figure 5. A total of **seven strictly conserved** residues were identified in GH13 family enzymes (Uitdehaag et al., 2002). Six of these seven residues are also conserved in GH70 family enzymes: the catalytic triad (Asp, Glu and Asp in motifs II, III and IV respectively), the arginine upstream of the first aspartate (motif II), the histidine upstream of the second aspartate (motif IV) and the aspartate in motif I (Vujičić-Žagar et al., 2010). The seventh conserved residue is an histidine but it is replaced by a glutamine in GH70 family enzymes motif I (Figure 5) (Vujičić-Žagar et al., 2010). There are **three other motifs** (V-VII) less well conserved among GH13 family enzymes that can be also found in GH70 family enzymes: motifs V, VI and VII (Figure 5) (Janeček, 1997, 2002; Janeček et al., 2014; Jespersen et al., 1993).

Of interest, the amino acids in the vicinity of acid/base catalysts and in particular those following the transition sate stabilizer (Motif IV, dashed line box in Figure 5) are often conserved in enzyme showing similar linkage specificity: the triplet YDA is only found in alternansucrases (*Ln. citreum* B-1355, LBAE-C11, KM20, ABK-1, B-1501 and B-1498 strains; see Part 2- II.1. Origin and discovery, pages 46-47), the triplet SEV is found in mostly dextransucrases (DSR-S, DSR-M, GTF-S) and mutansucrases (GTF-SI, GTF-I) whereas the triplet NNS is relative to reuteransucrases (GTFA, GTFO) (Table 1, Figure 5).

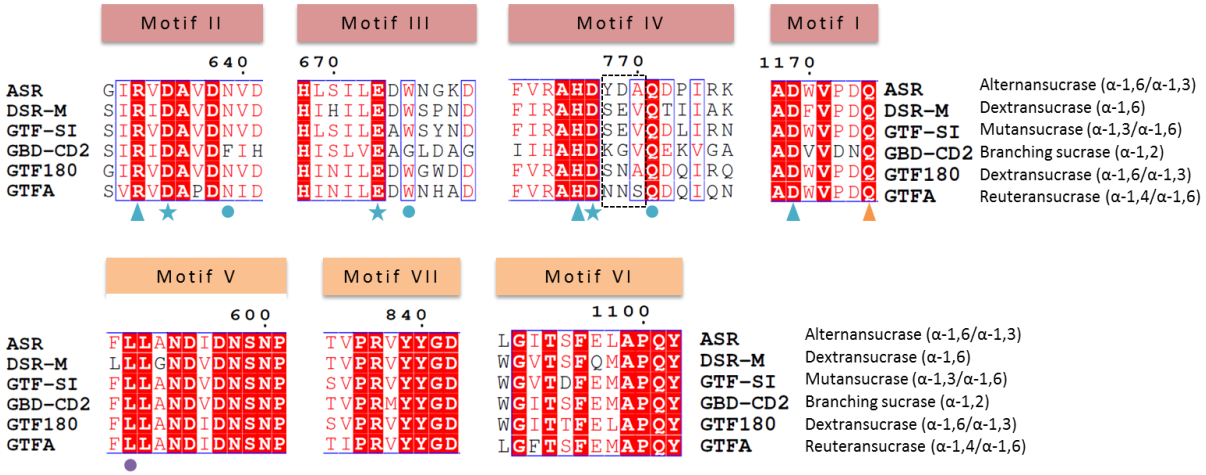


Figure 5: Sequence of highly conserved motifs I-VII in the sequences of sucrose-acting GH70 available structures. Stars: catalytic residues. Blue arrows and stars: residues conserved between the GH70 and GH13 family enzymes. Orange arrow: residue not conserved between the two families. Purple circle: residue found to interact with maltose in 3D structure (see II.3.2. Domain A, page 21). The alignment was created using ENDscript 2 (Robert and Gouet, 2014).

II.2.2. Repeated sequences in the glucan binding domain

Now, focusing on the glucan binding domain (**GBD**) defined as the C-terminal part of glucansucrases sequence (Figure 4), its name comes from the results of different biochemical studies that revealed it was involved in glucan binding (Funane et al., 1998; Komatsu et al., 2007; Moulis et al., 2006; Shah et al., 2004; Singh et al., 1993; Suwannarangsee et al., 2007; Wong et al., 1990; Wren et al., 1991).

Sequence analysis of the GBD of seven glucansucrases from *Streptococcus* sp. allowed the identification of the fundamental **YG-repeat** by Giffard *et al.* that defined this sequence by the presence of one or more aromatic residues (usually tyrosine) followed by glycine 3 or 4 residues downstream. The consensus sequence was:

NDG**YYF**xxx**G**xxhx(G/N)xhhh

(h= hydrophobic residue, x= poorly conserved residue) (Giffard and Jacques, 1994)

Consecutives repeats are separated by a basic residue (K or R). Their number varies in different glucansucrases. The YG-repeats are similar to **cell-wall** (CW) binding repeats and were classified into four different subgroups: A (Giffard and Jacques, 1994), B, C (Gilmore et al., 1990) or D-repeat (Giffard et al., 1993). A-repeat consensus sequence was described as:

WYYFDanGkaVTGaQtInGqtlYFdqdGkQVKG

(Capital letters: conserved residues among 96 A-repeat from 20 proteins) (Shah et al., 2004)

Such YG-repeats were also identified in the N-terminal part of certain glucansucrases (the variable region, Figure 4), in particular in the sequences of glucansucrases from *Lactobacillus* sp. (Kralj et al., 2004a).

II.3. Glucansucrase 3D structure

Less than ten years ago, a further step was taken with the resolution of the structure of the GTF180 glucansucrase from *Lactobacillus reuteri* (PDB IDs: 3HZ3, 3KLK, 3KLL) (Vujičić-Žagar et al., 2010). It provided new insights in the comprehension of the enzyme mechanism and domain organization. Glucansucrase structure was found to adopt a **U-shape** schematically visualized in Figure 6. This structure defined **five distinct domains**: domains A, B and C, structurally close to those of GH13 family enzymes, and domains IV and V, unique to glucansucrases (Figure 6). As a result of the circular permutation, each glucansucrase domain is built of a **non-contiguous chain** except for domain C (Vujičić-Žagar et al., 2010).

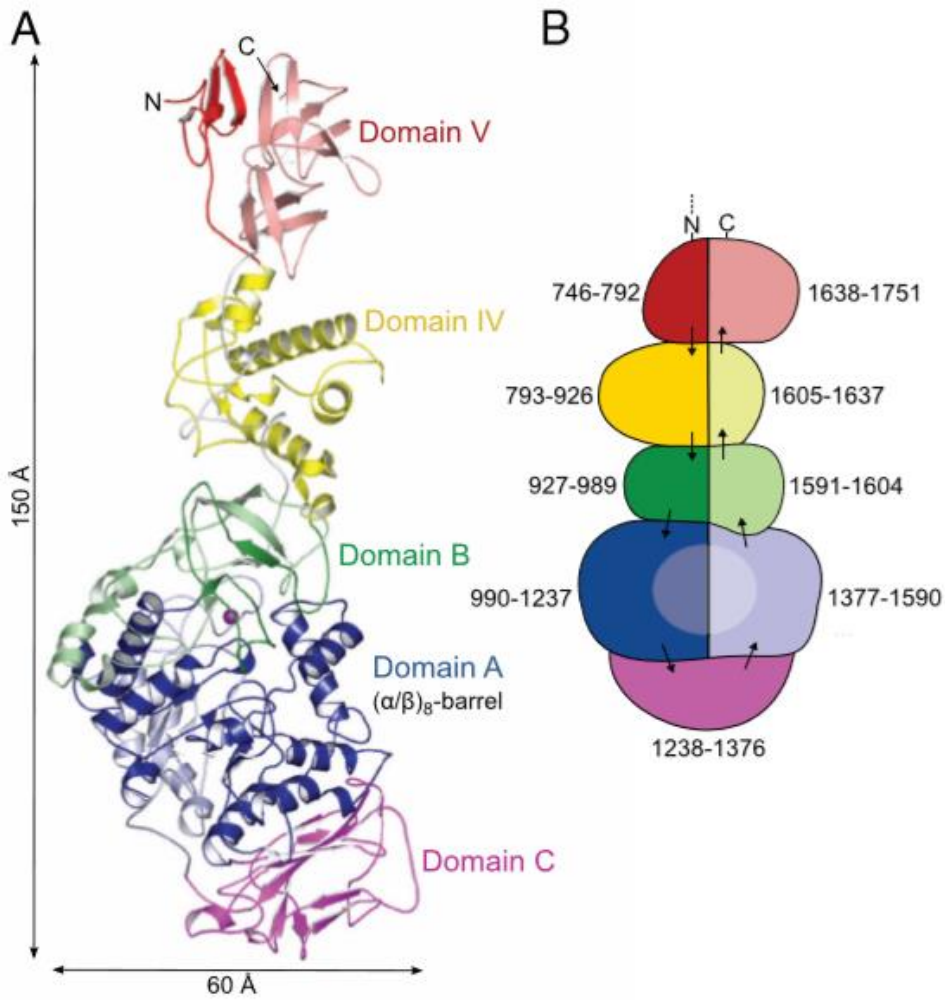


Figure 6: Overall structure of *L. reuteri* 180 GTF180-ΔN. (A) Crystal structure (B) Schematic representation of the "U-shaped" course of the polypeptide chain. From (Vujičić-Žagar et al., 2010).

II.3.1 Available structures

There is now several 3D structures besides the glucansucrase GTF180-ΔN (α-1,6/α-1,3) (Vujičić-Žagar et al., 2010) mentioned above: the mutansucrase GTF-SI (catalytic core only) (α-1,3/α-1,6) (Ito et al., 2011), the branching sucrase ΔN_{123}-GBD-CD2 (α-1,2) (Brison et al., 2012, 2016), the reuteransucrase GTFA-ΔN (α-1,4/α-1,6) (Pijning et al., 2012) and the dextransucrase DSR-MΔ2 (α-1,6) (Claverie et al., 2017). Importantly, all the solved structures were shown to adopt the same U-shaped fold with five domains and a circularly permutated $(\beta/\alpha)_8$ barrel (Figure 7). Different conformations of the domain V were obtained, that will be discussed in the Domain V and Domain IV parts, pages 25-27.

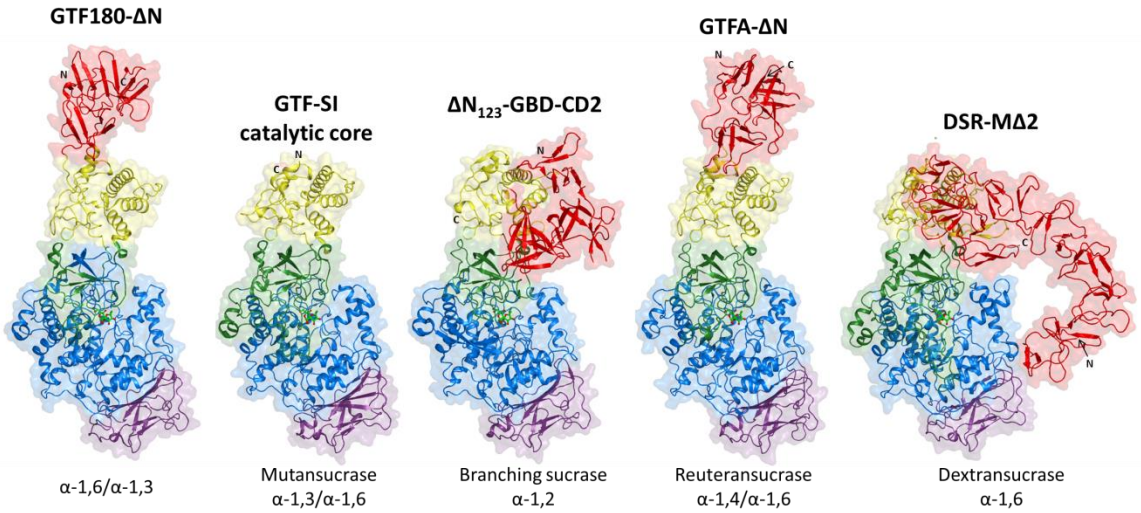


Figure 7: Domain definition of the five sucrose-acting GH70 structures solved to date. PDB IDs: 3KLK (GTF180-ΔN), 3AIB (GTF-SI-catalytic-core), 3TTQ (ΔN_{123}-GBD-CD2), 4AMC (GTFA-ΔN), 5NGY (DSR-MΔ2). Red: domain V; yellow: domain IV; green: domain B; blue: domain A; purple: domain C. Definition of the domains: (Vujičić-Žagar et al., 2010) for GTF180, (Ito et al., 2011) for GTF-SI, (Brison et al., 2012) for GBD-CD2, (Pijning et al., 2012) for GTFA, (Claverie et al., 2017) for DSR-M. Sucrose manually docked from 3HZ3.

II.3.2. Domain A

The domain A comprises the structural elements forming the **catalytic $(\beta/\alpha)_8$ barrel** and the **catalytic triad** (Figure 6, Figure 8). It also includes two α-helices localized between β_4 and α_5 elements of the barrel (dashed line box in Figure 8) that are found in GH70 family but not in GH13 family. These two α-helices are linked by a loop named loop A1 (1151-1160 in Figure 8) and were later called "**H1/H2 subdomain**" (Brison et al., 2012).

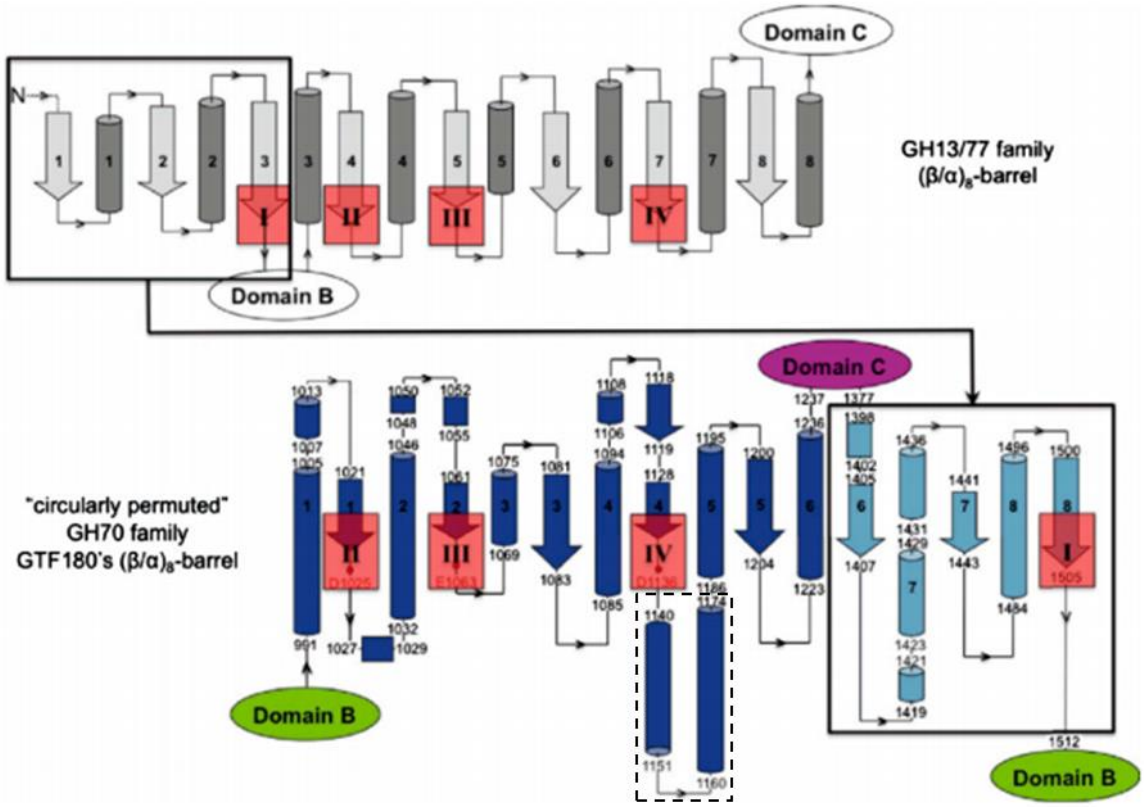


Figure 8: Schematic representation of the elements of the catalytic $(\beta/\alpha)_8$-barrel in GH13/77 (upper panel) versus GH70 family enzymes (lower panel, GTF180 numbering). Red boxes: conserved motifs I-IV. From **(Vujičić-Žagar et al., 2010)**

The GTF180-ΔN complexes with sucrose (PDB ID: 3HZ3) and maltose (PDB ID: 3KLL) (an acceptor very well recognized by most of the GH70 glucansucrases) enabled the identification of residues that are part of **-1, +1** and **+2** donor/acceptor **substrate binding sites** referring as to the nomenclature defined by Davies *et al.* (Davies et al., 1997). For glucansucrases, subsite -1 and +1 correspond to the sites where glucosyl and fructosyl moieties of sucrose are accommodated, respectively (Figure 9). Subsite +1 is also involved in the accommodation of acceptor molecules. For a given acceptor, different hydroxyl groups can be placed in favorable position to intercept the glucosyl enzyme. This is directly linked to the enzyme linkage specificity.

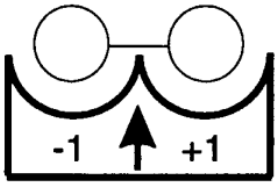


Figure 9: Schematic drawing of the sugar-binding subsites from (Davies et al., 1997). Arrow: point of cleavage.

For both complexes, the crystals were soaked and cryoprotected with sucrose (30 minutes with 25 mM) or maltose (15 minutes with 25 mM then two hours at 250 mM) (Vujičić-Žagar et al., 2010).

In GTF180-ΔN:sucrose complex (PDB ID: 3HZ3), ten residues, strictly conserved in motifs II to IV (Figure 5), are found in interaction with sucrose: the catalytic nucleophile (**Asp**1025, motif II, GTF180 numbering), the catalytic acid/base (**Glu**1063, motif III), the transition state stabilizer (**Asp**1136, motif IV), **Arg**1023 and **Asn**1029 (motif II), **Trp**1065 (motif III), **His**1135 and **Gln**1140 (motif IV) and **Gln**1509 (motif I) (Figure 10A). **Asp**1504 (motif I) interact with **Tyr**1465, a strictly conserved residue outside the motifs, that stacks the glucosyl moiety in subsite -1. Two other very conserved residues outside the motifs, **Asn**1411 and **Asp**1458 (loop A2 as defined recently (Bai et al., 2017)), were found to interact with sucrose (Figure 10A, Figure 11) (Vujičić-Žagar et al., 2010).

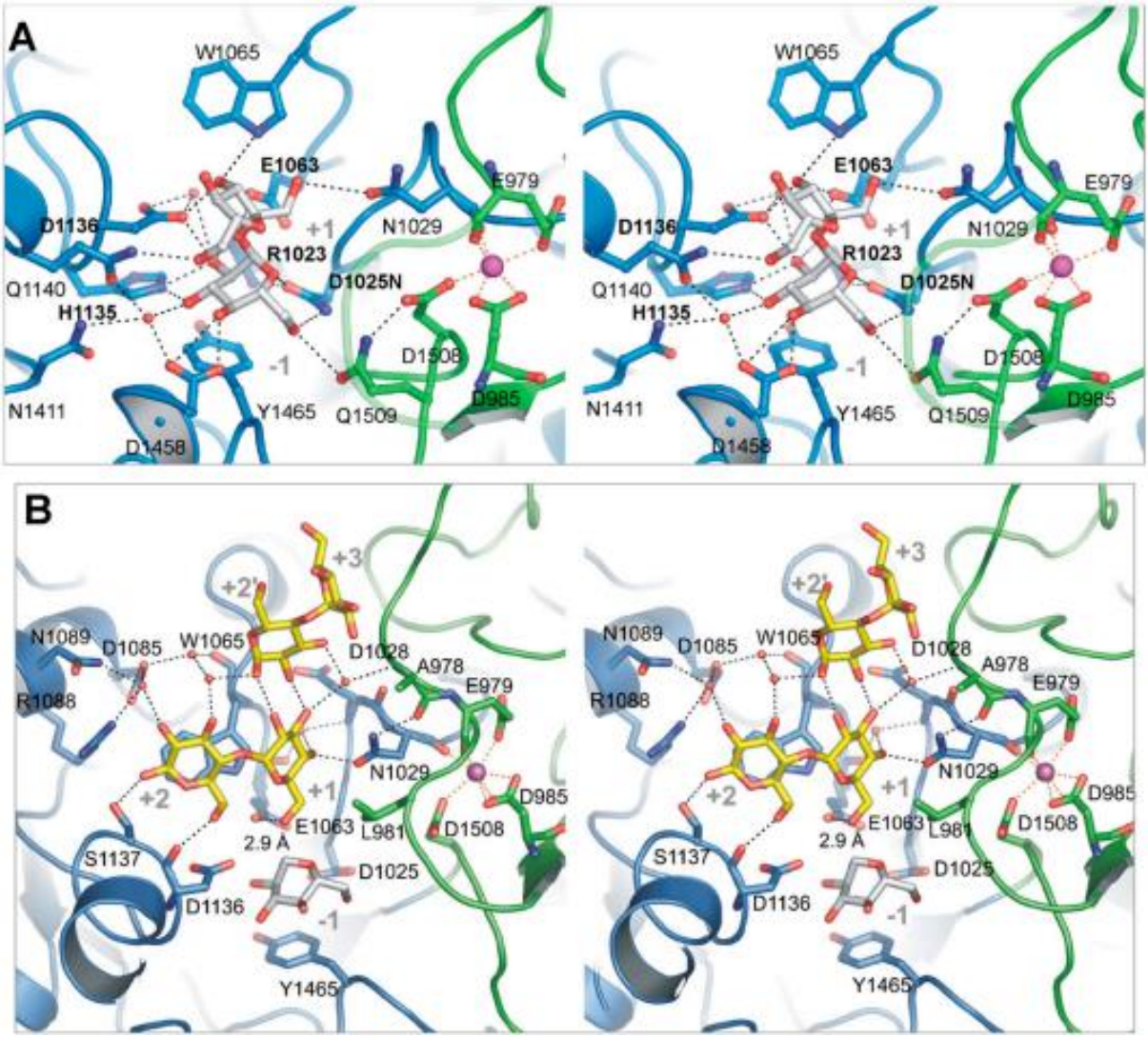


Figure 10: (A) Sucrose bound at subsites -1 and +1 in GTF180 Asp1025 (catalyst nucleophile) mutant. Bold: strictly conserved residues in GH70 and GH13 family enzymes (Asp1504 not shown). (B) Maltose bound at subsites +1 and +2 or subsite +2 and +3 with sucrose superimposition. From **(Vujičić-Žagar et al., 2010)**.

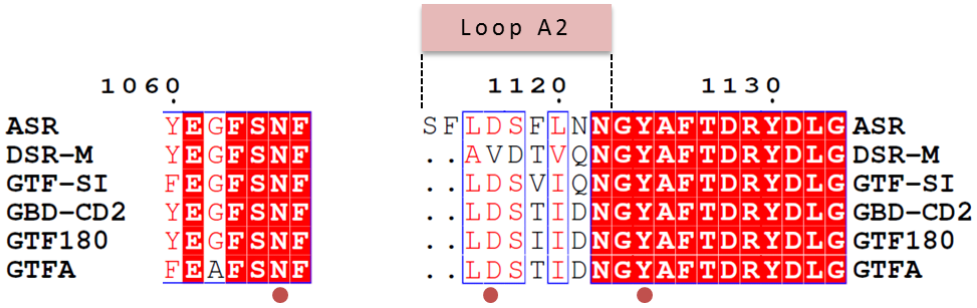


Figure 11: Sequence alignment of conserved stretches outside the motifs I-VII using sequences of sucrose-acting GH70 available structures. Pink circle: Residues found to interact with sucrose. The alignment was created using ENDscript 2 (Robert and Gouet, 2014).

Another sucrose molecule was found in the domain B but it was suggested to be of no significance for glucansucrases as the binding residues were found to be not conserved and a mutation of one stacking residue (W1531) had no effect on GTF180 product size (Vujičić-Žagar et al., 2010).

Four molecules of maltose were found in the complex GTF180-ΔN:maltose (PDB ID: 3KLL) at binding sites M1, M2, M3 and M4. One of them is positioned in subsites +1 and +2 (M1) and another one in two other subsites that were named +2' and +3 (M2) (Figure 10B). The molecule bound in the subsites +1 and +2 exposed the O6 of the non-reducing glucosyl unit (in +1 subsite) in a position favorable for glucosylation and panose formation (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc). Maltose binding was also stabilized by the same residues as those involved in sucrose binding and two leucines: Leu981 (motif V, Figure 5) and Leu938 (loop B2 from domain B) (Figure 10B) (Vujičić-Žagar et al., 2010).

The other binding sites of maltose molecules (M2, M3 and M4) were found to be not conserved in glucansucrases, thus it was suggested that these sites are of no importance for glucansucrases in general. M1, M2 and M4 are in the domain A whereas M3 is in the domain B (Vujičić-Žagar et al., 2010).

DSR-MΔ2 was also solved recently with a sucrose molecule in its catalytic site (PDB ID: 5O8L). The same residues were found in interaction with sucrose (Claverie et al., 2017).

GTF-SI catalytic domain (residues 244-1163) structure was also solved with one single maltose molecule in subsites +1 and +2 (PDB ID: 3AIB) (Ito et al., 2011) that superimpose with the one of GTF180-ΔN:maltose complex in subsites +1 and +2. Another complex was obtained with acarbose (PDB ID: 3AIC), a pseudotetrasaccharide found to be an inhibitor of streptococcal glucansucrases (Newbrun et al., 1983).

Recently, other ligands than sucrose and maltose were obtained in the domain A of the α-1,2 branching sucrase ΔN_{123}-GBD-CD2: soaking with D-glucose (PDB ID: 4TVD) allowed the identification of four binding sites in domain A (A-1, A2, A3, A4). One of them was positioned in subsite -1 (A-1) (Brison et al., 2016).

II.3.3. Domain B

The domain B is located next to the catalytic domain and is made of a highly twisted antiparallel five or six β-sheets (Figure 6) (Ito et al., 2011). This domain is essential to enzyme function and important for enzyme **specificity** as demonstrated by mutagenesis studies that targeted loop B1 or B2 (see III.3 Structure-function relationship, page 34).

Soaking with D-glucose allowed the observation of one glucose molecule at the interface of domains B, IV and V (site B-IV-V) in the enzyme ΔN_{123}-GBD-CD2 (PDB ID: 4TVD) (Brison et al., 2016). At another location in the domain B, sucrose (site S2) and maltose (site M3) molecules were also observed in GTF180-ΔN but this site was described of no importance for glucansucrases (Vujičić-Žagar et al., 2010).

II.3.4. Domain C

The domain C is composed by eight-stranded β-sheets forming a Greek key motif resembling that found in the GH family 13. This domain is **conserved** but its **function** has **not** been **elucidated** yet (Ito et al., 2011). It is the only domain made of **one contiguous segment** at the basis of the "U" shape (Figure 6). Interestingly, the domain C is the only domain for which no complexes have been obtained to date. Thus, its role is likely not related to glucan binding (more information is presented in the Chapter IV, page 156).

II.3.5. Domain V

The domain V is built of N-terminal and C-terminal segments of the protein and thus includes the so-called GBD, shown to be involved in **glucan binding** (see II.2.2. Repeated sequences in the glucan binding domain, page 17). Notably, no structure comprising entire domain V have been solved yet and it is only in 2016 that the interaction between oligosaccharide molecules and the domain V was structurally evidenced (Brison et al., 2016). Indeed, the soakings using D-glucose 100 mM (PDB ID: 4TVD), isomaltotriose 50 mM (PDB ID: 4TTU) or gluco-oligosaccharides 40 mM (PDB ID: 4TVC) with branching sucrase ΔN_{123}-GBD-CD2 resulted in the observation of isomaltosyl and isomaltotriosyl residues in the truncated domain V. These complexes allowed the mapping of **sugar binding pockets** and the identification of key residues interacting with the oligosaccharides. In particular, two sugar binding pockets (V-K and V-L) were structurally described. They consist of three consecutive elements: generally six β-sheets arranged in three β-hairpins (Figure 12A) and contain between 60 and 81 residues which correspond to **two A-repeats** (see II.2.2. Repeated sequences in the glucan binding domain, page 17). Sequence analysis of the glucan binding domain of DSR-E, the enzyme

from which GBD-CD2 is originated, enabled the identification of 12 putative sugar binding pockets in total. It also highlighted the conservation of four residues found to bind the sugar in pockets V-K and V-L: an aromatic residue representing a **stacking platform** (Tyr1834 in V-K and Tyr1914 in V-L), a second stacking platform (Trp1839 in V-K), a glutamine (Gln1879 in V-L and Gln1951 in V-L) and a lysine (Lys1881 in V-K and Lys1953 in V-L) (Brison et al., 2016).

Similarly, DSR-MΔ2 was soaked with isomaltohexaose (IM6) 50 mM (during 10 minutes) and a complex DSR-MΔ2:IM4 was obtained (two glucosyl units of the IM6 were not visible in the density map) (PDB ID: 5NGY). In this complex, isomaltotetraose (IM4) was also bound in the domain V, in a pocket named V-A similar to those found for GBD-CD2. In this pocket V-A, the isomaltotetraose ligand is in interaction notably with Tyr180 (stacking platform), Tyr187, Gln217, Lys219, Leu236 (main chain) and Tyr238 (Figure 12A). Several other putative binding pockets (V-A, V-B and V-C) were also identified by sequence and structural alignments. The biochemical characterization of the DSR-MΔ2 mutants in which the **stacking residues** (Tyr180 and Tyr264 for pockets V-A and V-B respectively) were replaced by alanine showed for the first time the **functionality** of the **sugar binding pockets** V-A and V-B whereas the pocket V-C was proposed to be non-functional (Claverie et al., 2017) (see III.3.3. Structure-guided mutants, page 41).

Additionally, the DSR-MΔ2 3D structure and SAXS analysis revealed a particular "**horse-shoe**" shape of the enzyme, observed for the first time (Figure 12B) (Claverie et al., 2017).

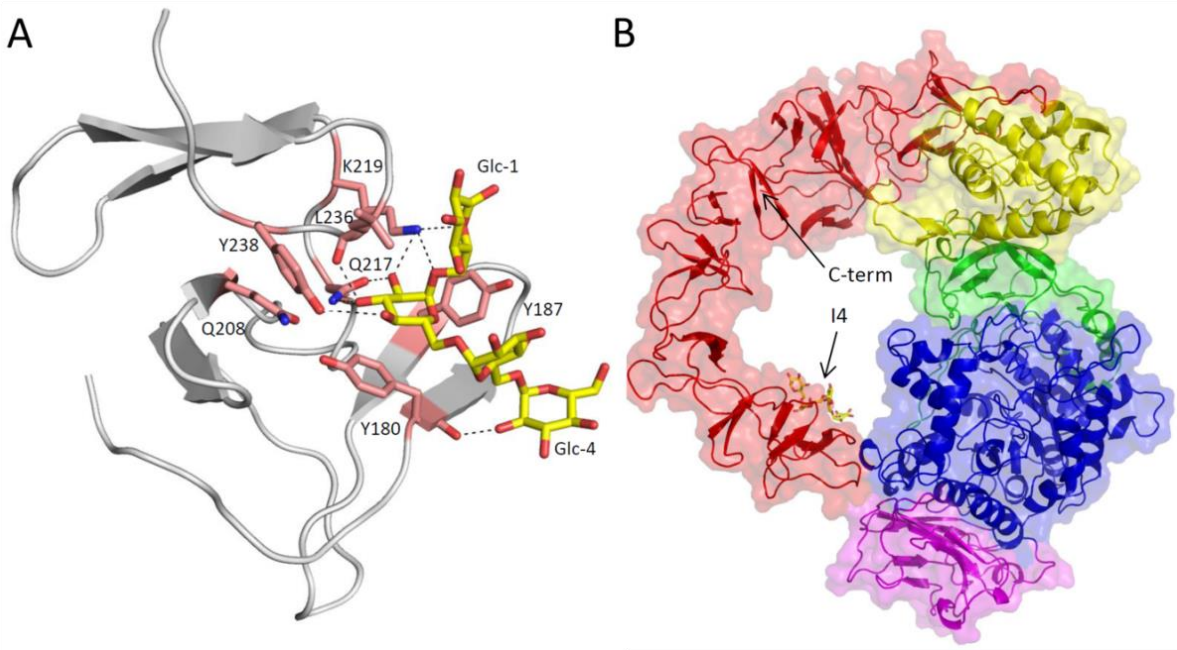


Figure 12: (A) Sugar binding pocket V-A of DSR-M in complex with isomaltotetraose (yellow sticks). (B) Particular "horse-shoe" shape of DSR-M. From (Claverie et al., 2017).

The domain V was found to be **not** always **necessary** for enzymatic **activity** in certain glucansucrases such as GTFA, DSR-M and GTF-I from *S. downei* MFe 28 (Claverie et al., 2017; Kralj et al., 2004b; Monchois et al., 1999a) whereas its deletion resulted in almost completely inactive enzyme in GBD-CD2 branching sucrase, DSR-S and DSR-OK (Claverie et al., 2019b; Fabre et al., 2005; Moulis et al., 2006). Claverie *et al.* showed that DSR-MΔ2 adopts a **distributive** mode (non-processive) to synthesize from 100 mM sucrose a polymer of 23,000 g/mol. In the absence of the domain V, the size of the polymer synthesized by DSR-MΔ5 is reduced to 16,000 g/mol, showing that the domain contributes to polymer elongation. It was suggested that an interplay between the domain V and the catalytic domain would allow to bring longer **chains** in the **proximity** of the **active site** to promote long chain extension (Claverie et al., 2017). This phenomenon is in accordance with the study of Moulis *et al.*, in which it was shown that gradual removal of the YG repeats of domain V led to the synthesis of polymer of lower sizes and switches the semi-processive mode of elongation to a more distributive one (Moulis et al., 2006).

II.3.6. Domain IV

Domain IV connects domain B and V (Vujičić-Žagar et al., 2010). It is suggested that this domain could play the role of a **hinge**, swinging domain V toward and away from the catalytic core to bring the domains A and V closer together and facilitate glucan extension (Ito et al., 2011). Small Angle X-Ray scattering of GTF180 as well as 3D structures (3KLK in triclinic form and 4AYG in orthorhombic form) supports this hypothesis and highlight the **dynamic** behavior of the enzyme in solution. Indeed, different conformations of the protein were obtained in which the domain V was either far, or in contrast close to domain A (boomerang-like shape) (Figure 13) (Pijning et al., 2014).

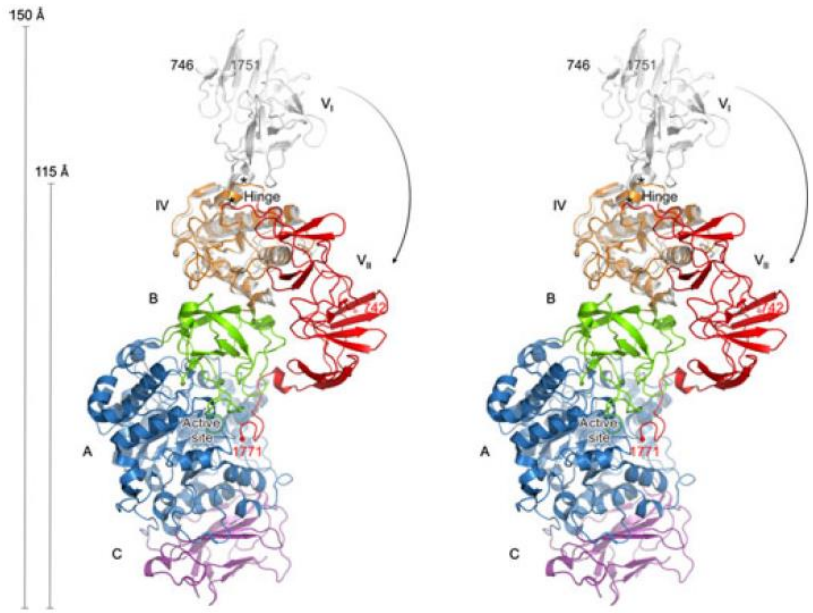


Figure 13: "Superposition of GTF180-ΔN I (grey, only domains IV and V are shown) and GTF180-ΔN II crystal structures." From (Pijning et al., 2014).

In the ΔN_{123}-GBD-CD2:glucose complex (PDB ID: 4TVD), four of the nine binding sites observed involved residues from the domain IV: one glucose molecule was identified at the interface between domains IV and V, another one at the interface between domains B, IV and V and two others on the surface of the domain IV (Brison et al., 2016).

II.3.7. Glucanotransferase structure

Recently, the first 3D structure of a 4,6-α-glucanotransferase of the family GH70 was also solved (Bai et al., 2017) that brought new elements concerning evolution. Thus, there are only **six enzyme structures** in GH70 family whereas the number grows to 117 enzyme structures for GH13 family (CAZy database, November 2018). All PDB IDs of GH70 family enzymes and corresponding complexes are listed in annex (Annex II, p. A3).

III. Evolution and structure-function studies

III.1. From GH13 to GH70 family

III.1.1. Glucanotransferases: a recently discovered evolutionary intermediate

As previously mentioned, the relationship between **GH13** and **GH70** families has been established since the nineties with the observation that they both adopt a $(\beta/\alpha)_8$ barrel (MacGregor et al., 1996). Moreover, a phylogenetic tree of α-amylase family members (now GH-H clan) highlighted the evolutionary relationship between *S. downei* glucansucrase, amylases and CGTases found in the same cluster (Janeček, 1997).

The three domains A, B and C of glucansucrases were found to superimpose to the A, B and C ones of GH13 family enzymes (Vujičić-Žagar et al., 2010). Vujičić-Žagar *et al.* in 2010 proposed a model of "permutation per duplication" that could explain the evolution process that could have taken place. The more recent discovery of the **glucanotransferases** that use **starch** instead of sucrose as substrate and more precisely, the identification and characterization of glucanotransferases in **non-LAB bacteria** (GtfC, GtfD; see I.2.3. Glucanotransferases, page 10) nuanced this model. Indeed, GtfC-like and GtfD-like enzymes lack the domain V but not the domain IV and do not possess the circular permutation (motif order: I, II, III, IV) suggested them to be an **evolutionary intermediate** between GH13 and GtfB-like glucansucrases, that use also maltodextrins or starch as a substrate but have the same domain organization than glucansucrases (circular permutation, motif order: II, III, IV, I and domains A, B, C, IV, V) (Figure 14) (Gangoiti et al., 2015, 2016, 2017a). Thus, the circular permutation would have occurred after the insertion of domain IV (Figure 15) (Meng et al., 2016b).

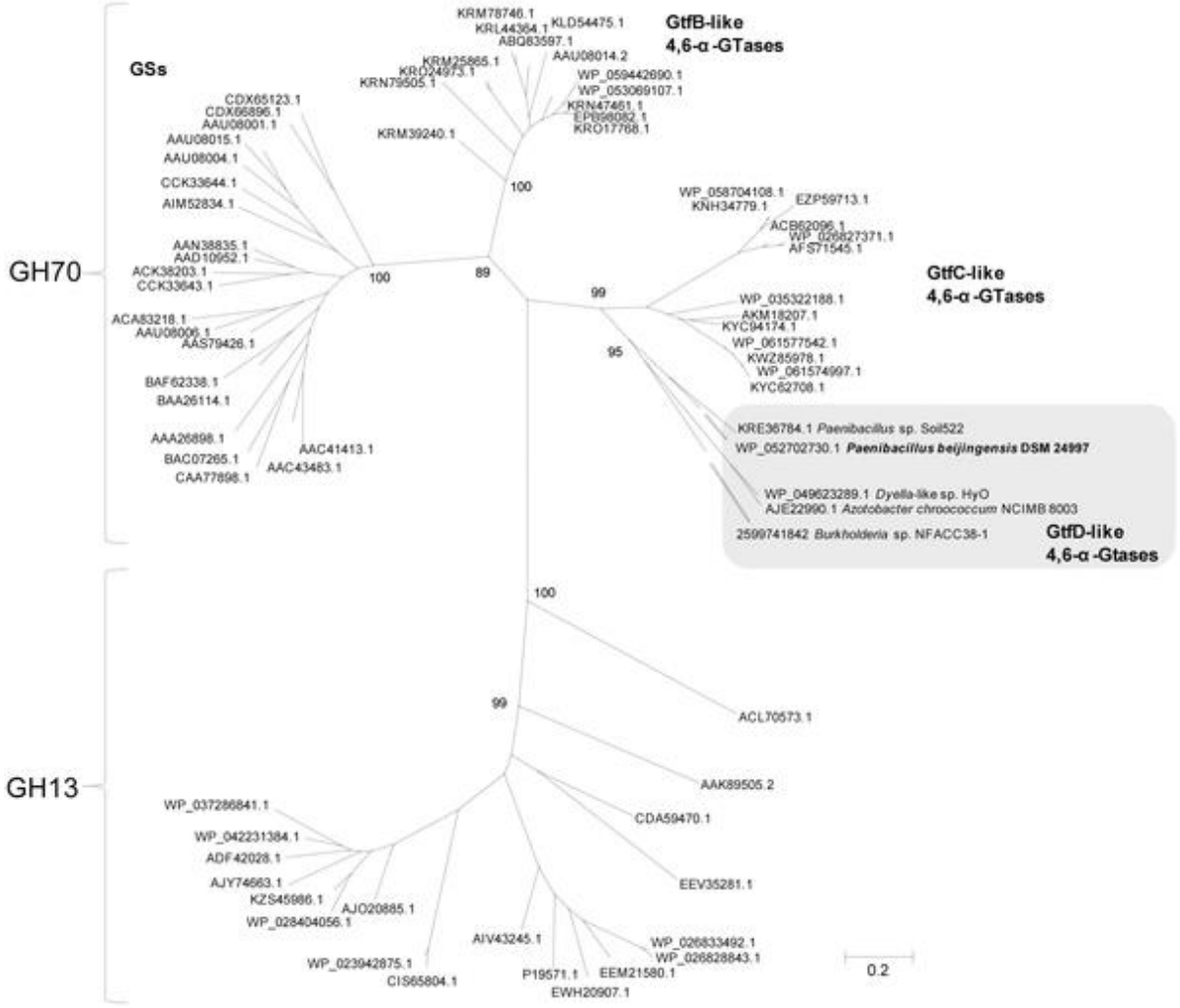


Figure 14: "Unrooted phylogenetic tree of representative family GH13 and GH70 protein sequences identified by BLASTp searches using the *A. chroococcum* GtfD 4,6-α-Gtase protein as query". From (Gangoiti et al., 2017a)

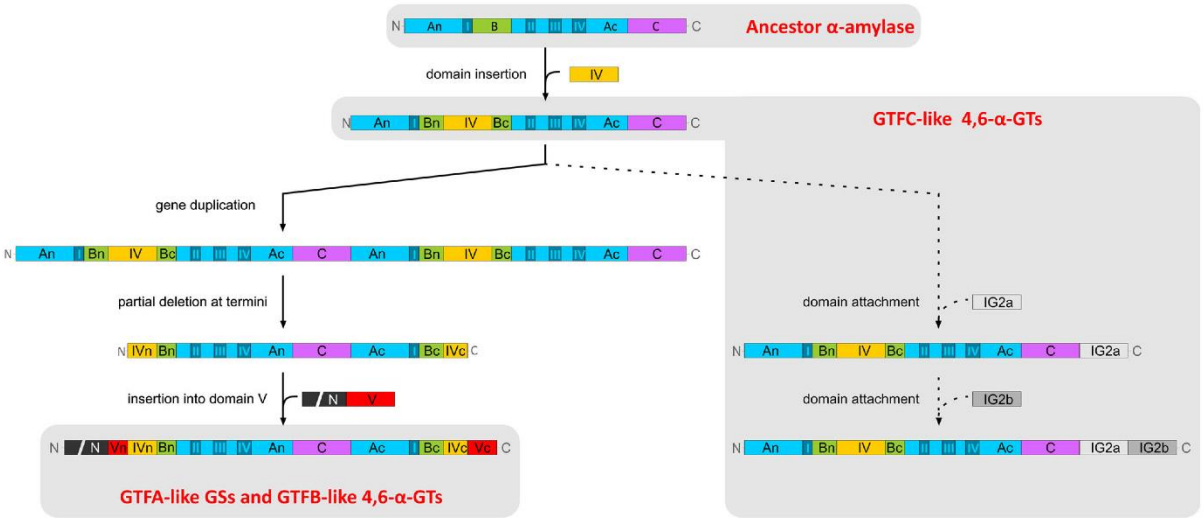


Figure 15: Possible evolutionary pathway according to the "permutation per duplication model". From (Meng et al., 2016b)

Structure resolution and analysis of the 4,6-glucanotransferases (GtfB) from *L. reuteri* 121 (PDB IDs: 5JBD, 5JBE, 5JBF) revealed that the active-site architecture of GtfB is intermediate between GH13 α-amylases and GH70 glucansucrases. Notably, loops elongation or shortening could have helped the transition from maltooligosaccharide substrate to sucrose substrate (Figure 16). The presence of **transposase** around the genes of 4,6-α-glucanotransferase and glucansucrase organized in tandem in *L. reuteri* 121 suggests this gene pair to be the result of **gene duplication** event. It was also suggested that the changes in our **dietary** (from starch to sucrose intake) was important factors in the evolution from α-amylases to glucanotransferases and later to glucansucrases in probiotic strains such as Lactobacillus reuteri; allowing the formation of extracellular polymers from different substrates (Bai et al., 2017).

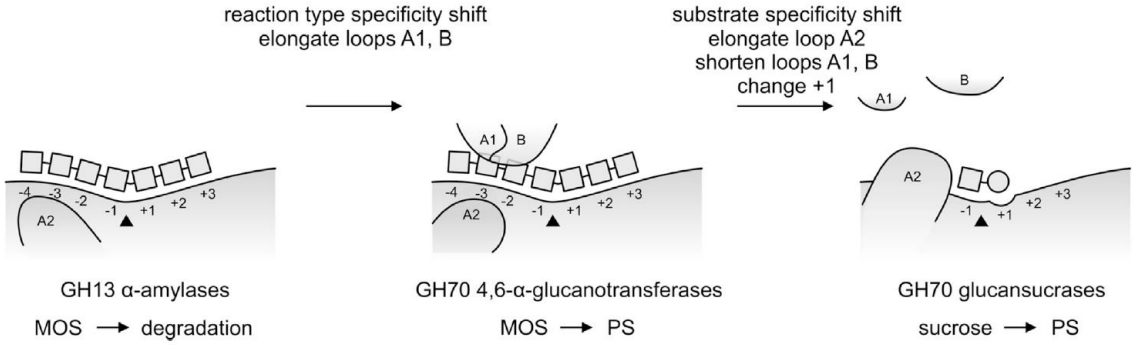


Figure 16: Illustration of the evolution of GH70 glucansucrases from GH13 α-amylase. From (Bai et al., 2017)

III.1.2. Branching sucrases: an evolution process unraveled to that date

Branching sucrases are the third subgroup in GH70 family (see I.2.2. Branching sucrases, page 9). Phylogenetic analysis highlighted that they form a **cluster** in a distinct branch, apart from glucansucrases and glucanotransferases, and were divided in two part: one corresponding to α-1,2 specificity and the other to α-1,3 specificity (Vuillemin et al., 2016).

Recently, an **atypical dextransucrase** named Gsy (*Leuconostoc mesenteroides* BD3749) was characterized and described as an evolutionary intermediate between branching sucrases and glucansucrases based on sequence analysis. Indeed, Gsy has the "apparent" **sequence** of a **branching** sucrase but is able to perform the polymerization reaction. It was suggested that the residues Met693, Arg695 (motif IV, Gsy numbering), Ile549 (motif V) and Ser771 (motif VII) play an important role in specificity because the residues were different from glucansucrases and branching sucrases (Figure 17) (Yan et al., 2018). Moreover, both the glucansucrase conserved asparagine in motif II and the tryptophan in motif III are not present in Gsy and the motif following the TSS in Gsy (^{699}RGS^{701}) is

not present in glucansucrases neither in branching sucrases (Figure 17). However, very few is known about branching sucrase and glucansucrase evolution pathway to that date.

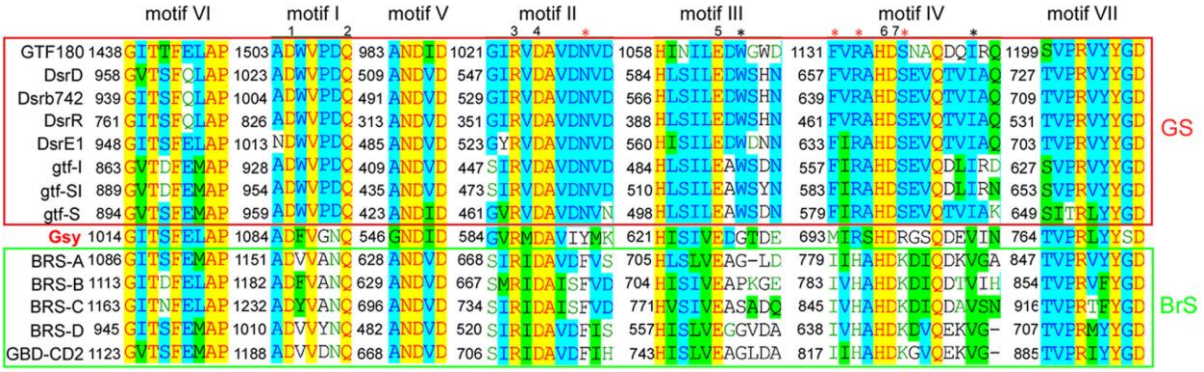


Figure 17: "Sequence-based analysis of motifs I to VII of the catalytic core of different GH70 enzymes." From (Yan et al., 2018)

III.2. Gene acquisition by GH70 family enzyme producers

In *Ln. citreum*, the importance of **phage mediated horizontal transfer** was suggested as the dextransucrase DSR-DP was likely originated from such type of transfer (Passerini et al., 2015). Additionally, phylogenetic analysis of 44 glucansucrases sequences, including 20 *Streptococci* glucansucrases, led to the submission that the transfer of glucansucrase gene from lactic acid bacteria (*e.g. Lactobacillus* and *Leuconostoc*) to *Streptococci* occurred when *Streptococci* species encountered such LAB from fermented food. The transfer would have been mediated by **horizontal transfer using transposons**. Our consumption of refined sugars would have been a secondary selection pressure favoring the acquisition of multiple *gtf* genes by gene duplication (as GTF-S, GTF-SI and GTF-I are usually found in tandem in *Streptococci* strains) (Hoshino et al., 2012). Phylogenetic analysis of a larger dataset of 66 glucansucrases (including gene sequenced from chimpanzee and macaque *Streptococcus* strains) revealed a monophyletic relationship (*e.g.* derivation from one common ancestor) among the streptococcal enzymes, regardless of the glucan formed, contrary to enzymes from *Lactobacillus* and *Leuconostoc* species; for which a polyphyletic relationship was described (Figure 18). The common ancestor of streptococcal glucansucrases was dated to 124-140 million years ago (Mya) that could fit with the divergence time of marsupials and was likely a dextran-producer. The apparition of mutan-producer glucansucrases (only found in *S. mutans* strains to that date) was dated around 85-90 Mya and would be related to the selective advantage resulting in the production of adhesive glucan (mutan) (Argimón et al., 2013). Indeed, modern humans dental plaque is formed of 70% water-insoluble glucan (WIG, mutan) and less than 2% of water-soluble glucan (WSG, dextran) (Hotz et al., 1972). Diversification of streptococcal glucansucrases (GTF-I, GTF-

S, GTF-SI) is thus not governed by the change in human diet. These results are not is accordance with the one of Hoshino mentioned above (Argimón et al., 2013).

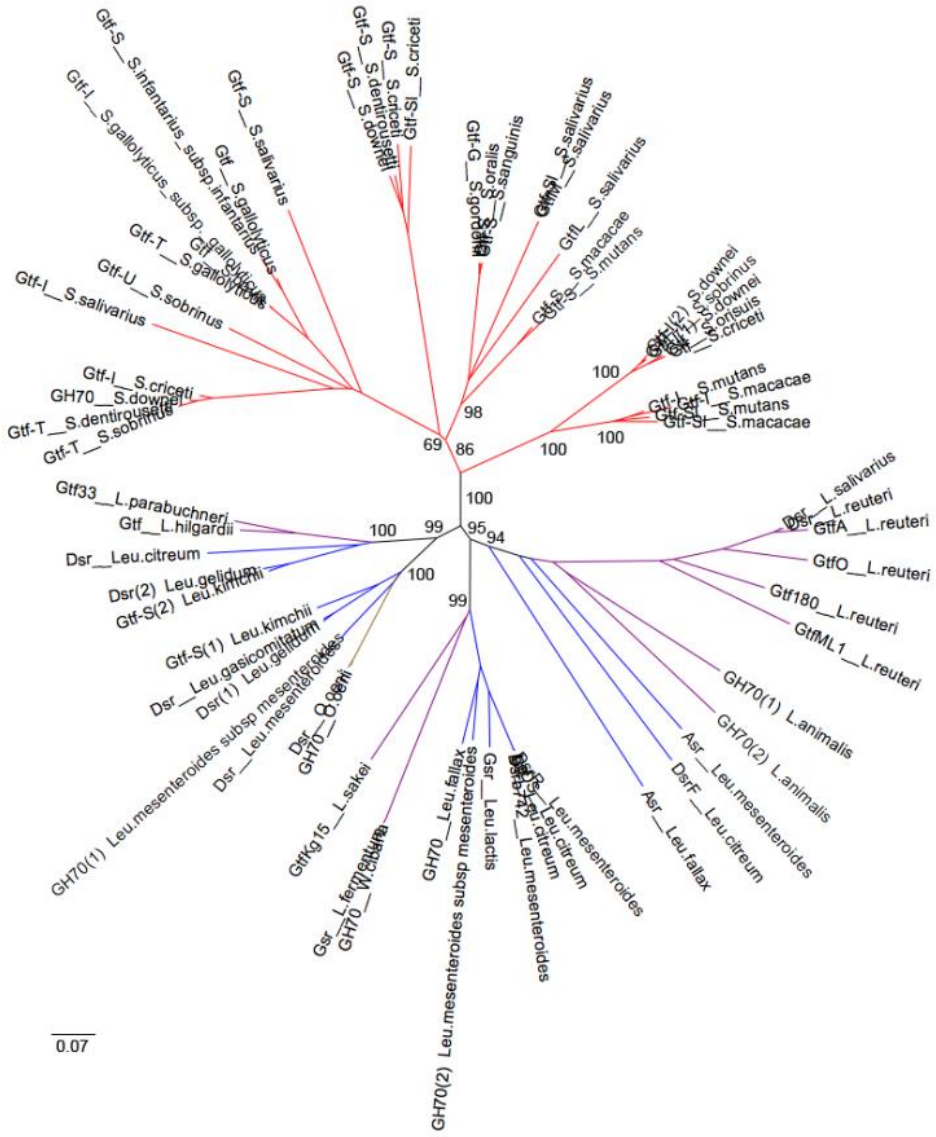


Figure 18: Phylogenetic tree of the dataset of 66 glucansucrases. From (Argimón et al., 2013)

III.3. Structure-function relationship

Sequence and structural comparisons of glucansucrases coupled with mutagenesis studies and biochemical characterization have allowed deciphering the role of various amino acids in catalysis, linkage specificity, or ability to perform polymerization. In a first time, they allowed the identification of the major residues implied in **catalysis**: residues from the catalytic triad (DED) and the confirmation of the importance of the strictly conserved residues in the conserved motifs. Then, studies highlighted the importance of certain residues for **activity, linkage specificity and/or polymer size or proportion**. Since 2011, less than a decade, five sucrose-active GH70 family enzyme 3D structures have been solved, enable the rationally design of enzyme mutants, not only based on sequence alignment and analysis, but also on the three-dimensional positioning of the residues.

A rather large number of studies have been released on glucansucrases or the branching sucrase GBD-CD2: we count a total of **37 papers** that used **16 different enzymes** as template for mutagenesis. However, it is quite messy to compare the result of these studies because enzyme name or numbering has sometimes changed, the numbers can be very different depending on the enzyme (from 451 to 2210 for the nucleophile catalyst number), and GenBank accession numbers are rarely indicated. Hopefully, the major part of the sequences can be retrieved using the **CAZy database** (http://www.cazy.org/GH70.html). To circumvent these problems, we decided to use the recently **ASR structure** that we solved (see Chapter II, page 76) **as reference** (the largest structure to that date), to find and align the mutant enzyme sequences in order to report the corresponding position in ASR. This tool will facilitate the identification of already targeted position in the literature and the **comparison of mutation effect** from one enzyme to another. Accumulating such knowledge could be a key in **structure-function relationship understanding,** and thus facilitate the rational design of glucansucrases for the production of **tailor-made polysaccharides** with many applications in food, feed or health fields.

All the positions targeted in the 37 studies that found an equivalent in ASR (identified by sequence alignment using Clustal Omega or a structural alignment when available) are localized on ASR structure in the Figure 19 and their location is indicated in the Table 2.

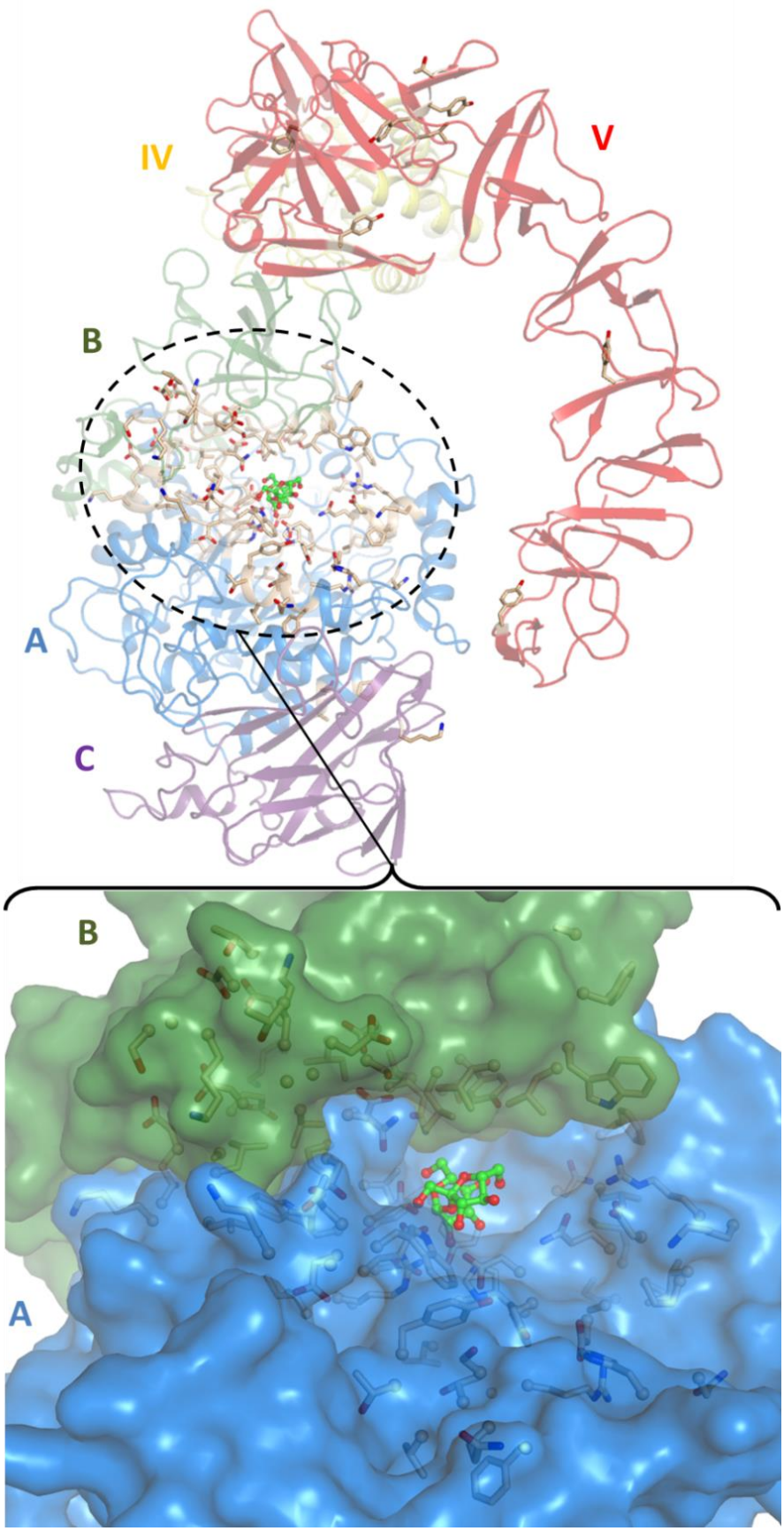


Figure 19: Mapping of all targeted mutations.
The equivalents of the residues in stick were targeted in at least one study. Red: Domain V, Yellow: domain IV, green: domain B, Blue, domain A, Purple: domain C. Sucrose manually docked from 3HZ3.

Table 2: Localization and number of mutants studied in all GH70 family enzymes

Domain	Motif	ASR equivalent	Nb. of mutants
A	V	595-606	15
	II	632-647	32
	III	673-678	29
	IV	764-859	105
	VI	1098-1101	1
	I	1169-1176	11
	none	598-872 1031-1176	56
B	none	529-597 1177-1270	32
C	none	873-1030	1
IV	none	400-528 1271-1303	1
V	none	1-399 1304-1425	23
TOTAL			306

III.3.1. Mutagenesis studies on very conserved residues

The Table 3 gathers some of the first studies performed on the catalytic triad of glucansucrases. The results clearly indicate that the mutations performed onto the nucleophile catalyst, the acid/base catalyst or the transition state stabilizer (TSS) are deleterious and result in a mutant without any detectable activity or less than 0.3% residual catalytic efficiency.

Similarly, mutants of the residues found in interaction with sucrose in the GTF180-ΔN:sucrose complex have been targeted, before or after structure solving (see Annex III, Table 1 page A6). The results showed that residues Arg633 (Motif II, ASR numbering), His766 (motif III), Gln1174 (motif I) and Tyr1124 (glucose stacking -1) are essential for activity. Both Asn639 (motif II) and Trp675 (motif III) mutants have also a negative impact on activity and increase the hydrolysis. To note, Trp675 and Gln771 (motif IV) mutants decrease the proportion of high molar mass polymer synthesis.

Table 3: First mutagenesis results on the catalytic triad. References: 1= (Kato et al., 1992); 2= (Monchois et al., 1997); 3= (Devulapalle et al., 1997) ; 4= (Kralj et al., 2004b) ; 5=(Swistowska et al., 2007)

Ref.	Enzyme			Localization				ASR equivalent	Effets	
	Enzyme name and GenBank accession nb	Organism	Catal. Residues	Domain	Motif	Precise localization	Mutant	Residue	Active?	Residual activity OR residual catalytic efficiency (%)
1	GTF-I AFM81411.1	*S. mutans* GS-5	D451 E489 D562	A	II	WILD TYPE		D635	YES	100
						Nucleophile catalyst	D451E		NO	
						Nucleophile catalyst	D451N		NO	
						Nucleophile catalyst	D451T		NO	
2	DSR-S AAD10952.1	*Ln. Citreum* NRRL B-512F	D551 E589 D662			WILD TYPE			YES	100 (1040 $U.g^{-1}$)
						Nucleophile catalyst	D551N		NO	
3	GTF-I	S. downei	D415 E453 D526			WILD TYPE			YES	100 (k_{cat}/K_m= 346 $min^{-1}.mM^{-1}$)
						Nucleophile catalyst	D415N		YES	0,06
4	GTFA AAU08015	*L. reuteri* 121	D1024 E1061 D1133			WILD TYPE			YES	100
						Nucleophile catalyst	D1024N		YES	0,1
3	GTF-I	*S. downei*	D415 E453 D526		III	WILD TYPE		E673	YES	100 (k_{cat}/K_m= 346 $min^{-1}.mM^{-1}$)
						Acid/Base catalyst	E453Q		YES	0,03
4	GTFA AAU08015	*L. reuteri* 121	D1024 E1061 D1133			WILD TYPE			YES	100
						Acid/Base catalyst	E1061Q		YES	0,1
5	GtfR BAA95201.1	*S. oralis* ATCC10557	D516 E554 D627 Model: D190 E228 D301		IV	WILD TYPE		D767	YES	100
						TSS	D627E		YES	0,3
3	GTF-I	*S. downei*	D415 E453 D526			WILD TYPE			YES	100 k_{cat}/K_m= 346 $min^{-1}.mM^{-1}$
						TSS	D526N		YES	0,3
4	GTFA AAU08015	*L. reuteri* 121	D1024 E1061 D1133			WILD TYPE			YES	100
						TSS	D1133N		YES	0,3

III.3.2. Mutations impacting the specificity

Considering specificity, one example could be the results obtained by mutagenesis onto the same position that was highlighted in five different studies performed onto five different glucansucrases, both dextransucrases and mutansucrases (Table 4). The corresponding position in ASR is Asp772. This residue is in the domain A, in the motif IV, exactly five residues downstream of the TSS.

Sequence analysis on enzymes from *Streptococci* revealed that an aspartate is present at this position for GTF-SI-like and GTF-I-like enzymes (mutan/insoluble glucan/IG producers) whereas a threonine is present in GTF-S-like enzymes (dextran/soluble glucan/SG producers). The swapping between aspartate and threonine -either in GTF-I or GTF-S- resulted in an impressive increase in soluble glucan (from 0 to 24%) or insoluble glucan (from 14 to 85%) respectively (Shimamura et al., 1994). Similar effect was observed with the mutation of the threonine of DSR-S to an aspartate (Remaud-Siméon et al., 2000) and of the aspartate of another GTF-I to a threonine: the mutant switched from insoluble glucan production to both soluble and insoluble glucan production (Monchois et al., 2000b). It was also shown that this position was important for oligosaccharide size (Monchois et al., 2000b). Yet, DsrI appears to be an exception, displaying a threonine at this position but forming water insoluble glucan (Côté and Skory, 2014).

.

Table 4: Mutagenesis results on position equivalent to D772 in ASR. References: 1= (Kato et al., 1992); 2= (Remaud-Siméon et al., 2000); 3=(Monchois et al., 2000b); 4= (Côté and Skory, 2014); 5= (Shimamura et al., 1994)

Ref	Enzyme			Mutants	ASR equivalent	Effets					
	Enzyme name and GenBank accession nb	Organism	Catal. Residues			Active?	Residual activity (%)	HMM proportion (%)	Hydrolysis (%)	NMR (%) or Residual glucan-synthetic activity (%)	
										α-1,3 or IG	α-1,6 or SG
1	GTF-I AFM81411.1	*S. mutans* GS-5	D451 E489 D562	WT	D772	YES	100				
				D567T		YES	75,6				
5	GTF-S AFM81328.1	*S. mutans* GS-5	D465 E503 D584	WT						14	86
				T589D						85	15
				T589E						98	2
	GTF-I AFM81411.1		D451 E489 D562	WT						100	0
				D567T						76	24
2	DSR-S AAD10952.1	*Ln. citreum* NRRL B-512F	D551 E589 D662	WT		YES				5	95
				T667R		YES				13	87
3	GTF-I AAC63063.1	*S. downei* MFe28	D453 E491 D564	WT		YES	100 12.07 $U.mg^{-1}$	61	17		
				D569E		YES	81	72	17		
				D569T		YES	72	78	14		
				D569N		YES	48	71	17		
				D569H		YES	43	72	15		

				D569V		YES	42	76	14		
				D569S		YES	36	74	15		
				D569I		YES	31	73	15		
				D569A		YES	29	74	15		
				D569R		YES	16	72	18		
				D569Y		YES	14	73	16		
				D569L		YES	12	73	17		
4	Dsr-I	*Ln. mesenteroides* NRRL B-1118	D533 E571 D649	WT						44	56
				T654E						62	38
				T654K						68	32
				T654Q						71	29
				T654C						66	34
				T654R						78	22
				T654N						70	30
				T654I						74	26
				T654S						62	38
				T654H						71	29
				T654D						50	50
				T654G						52	48
				T654Y						43	57

III.3.3. Structure-guided mutants

The mutations performed in GTF180, GBD-CD2, GTFA and DSR-M (after 3D structure resolution) were mapped on each corresponding structure (Figure 20). GTF180 has been the most used template for mutation studies (86 mutants constructed and characterized), before that of GTFA, with 27 mutants.

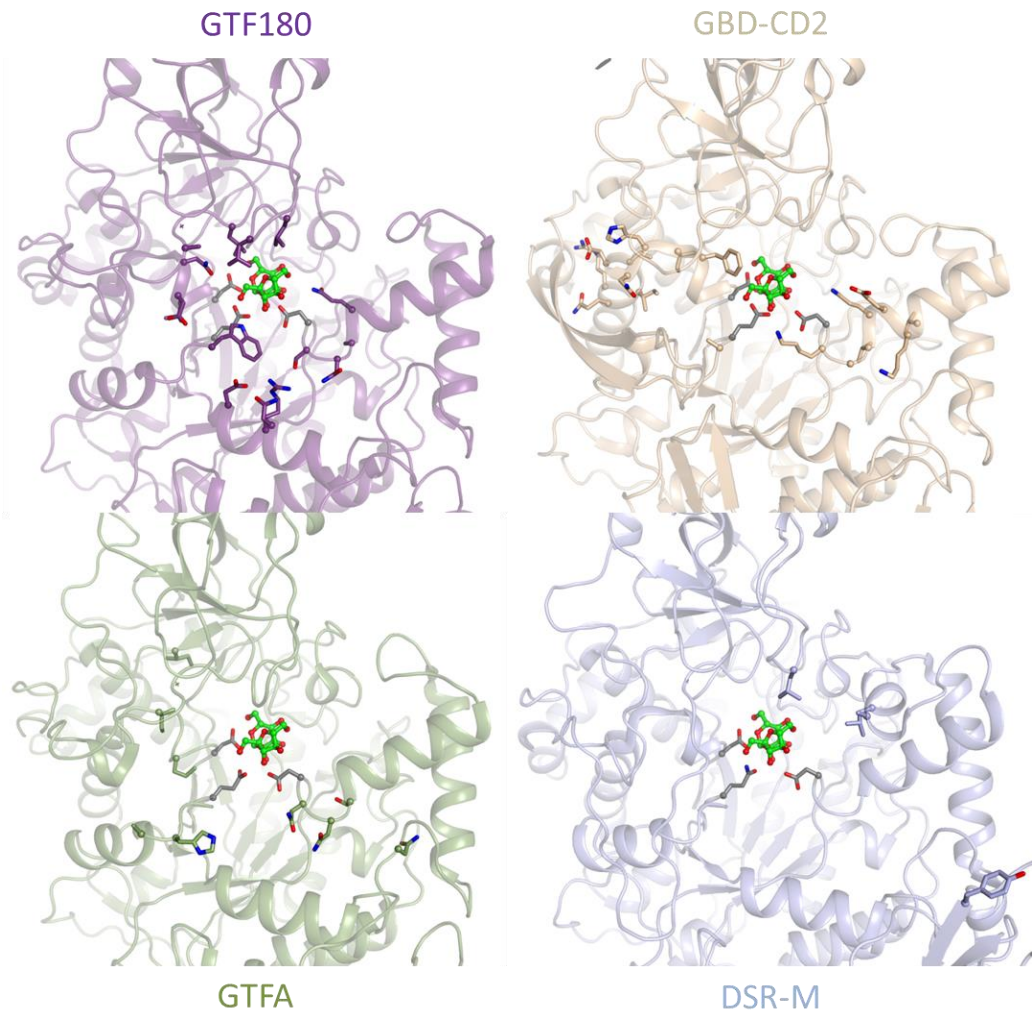


Figure 20: Mapping of mutants from sucrose-active GH70 structures

For GBD-CD2, all the mutants constructed resulted in a decrease of activity with no impact on the linkage specificity (α-1,2 branching) (see Annex III, Table 2 page A10). To date, it is the only branching enzyme that has been mutated for structure-function studies. Mutation of residues Ala2249, Gly2249 (or both) to tryptophan was not sufficient to switch the enzyme activity from a branching sucrase to a polymerase (Brison et al., 2012). To note, multiple mutations in several sequence stretches of Motif II and IV (eight and seven residues respectively) to the corresponding sequences of DSR-S resulted in an enzyme that synthesized from sucrose a product stainable by Schiff coloration likely corresponding to a polymer (Fabre et al., 2006).

For DSR-M, as mentioned in the paragraph describing the role of Domain V (see II.3.5. Domain V, page 25), mutants of the stacking residue in the sugar binding pockets V-A and V-B (Tyr158A and Tyr241A) highlighted the functionality of the pocket and their importance for polymer elongation (see Annex III, Table 3 page A11).

From this study, the QxK motif was suggested to be essential for pocket functionality (Claverie et al., 2017). Interestingly, Gln1321 and Lys1323 were also mutated in repeat of GBD1A (a truncated version of the GBD of *S. downei* GtfI that contains four A repeats, as identified at that time). The binding of GBD1A was evaluated using biotinylated dextran and Gln1321Ala, Gln1321Asn and Lys1323Arg were shown to not affect the binding ability whereas the Lys1323Ala mutation resulted in a loss of 75% binding (Shah et al., 2004). This suggests the importance of a basic residue at this position (Arg or Lys). Moreover, mutation of the residues Trp1227 and Trp1292 into alanine, in the third and fourth repeat of GBD1A, (likely to correspond to the second stacking of sugar binding pocket) resulted in a complete loss of biotinylated dextran affinity whereas no effect was observed with mutation Trp1163Ala in the second and repeat (Shah et al., 2004). Yet, sequence alignment suggested these three tryptophan to play the same role of second stacking interaction in sugar binding pockets (Brison et al., 2016).

Mutations performed in GTF180 targeted 14 different positions both in domain A (motif II, III, IV, helix α_6 of the catalytic barrel) and in domain B (loop B1, loop B2). The results showed the key role of some residues from both domains concerning activity (*e.g.* Leu938, Asp1028, Asn1029, Trp1065), specificity (*e.g.* Leu940, Asp1028) or polymerization (*e.g.* Trp1065, Leu940) (see Annex III, Table 4 page A12). Mutant of GTFA were also constructed in the domain A and also shed the light on residues important for enzyme catalysis (*e.g.* the catalytic triad Asp1024, Glu1065, Asp1133 or Asn1134) or specificity (Asn1134) (see Annex III, Table 5 page A17).

As these four enzymes do not have the same specificity, the comparison is not well-to-do. We would need more structures coupled to more biochemical characterization to better understand the determinants of specificity, polymerization or branching for each enzyme.

III.4. Applications

To finish, α-glucans produced by enzymes from GH70 family represent a quite important class of polysaccharides in terms of applications. Notably, the **dextran** from *Ln. mesenteroides* NRRL B-512F has been the **first** microbial biopolymer produced on an industrial scale and commercialized since **1948** to be used as **plasma substitute** (Heinze et al., 2006; Leathers, 2005; Vettori et al., 2012).

Nowadays, the market mainly concerns linear dextrans or iron and sulfate derivatives of average molar mass of 20, 40, 60 and 70 kg/mol obtained by partial acid hydrolysis and solvent fractionation. It was estimated to be of 170 Million US$ in 2017 and is dominated by **biomedical** applications (blood expanders, iron carriers or anticoagulant agents (iron or sulfate dextrans)) and **analytical** applications (versatile chromatographic supports known as Sephadex® gels, commercialised by GE Healthcare). This market size is forecast to grow at 3.2% CAGR (Compound Annual Growth Rate) to reach 250 Million US$ in 2025. The industrial players are Pharmacosmos, PK Chemicals, Meito Sangyo, Polydex Pharm, Jinyang Biological Pharmaceutical ("Global Dextran Market Insights, Forecast to 2025," 2018). Dextrans are also used in **food** for their thickening properties, **cosmetics** or **flocculants** for mineral extractions. Furthermore, novel applications are recently emerging for α-glucans. They concern the use of **mutan** type of polysaccharides for **biomaterial** applications. These α-1,3 glucans produced by *Streptococci* species from sucrose have also raised attention in the last five years due to their insolubility and mechanical properties which can be further enhanced by **chemical modifications** such as carboxymethylation (DuPont patent (Dennes et al., 2015; Paullin et al., 2014)) thus offering the possibility to develop versatile platforms of new α-glucan-based biomaterial (NUVOLVE™).

Another important application concerns the use of glucansucrases to produce **prebiotics**. The prebiotics definition was recently updated to « a substrate that is selectively utilized by host microorganisms conferring a health benefit » by experts from the International Scientific Association for Probiotics and Prebiotics (ISAPP) (Gibson et al., 2017). The α-1,2 branched oligosaccharides obtained by the native DSR-E glucansucrase or those obtained by combining the action of a dextransucrase and an α-1,2 branching sucrase were shown to possess prebiotic properties (Djouzi et al., 1995; Sanz et al., 2005b; Sarbini et al., 2011, 2013; Valette et al., 1993). The isomaltooligosaccharides and α-1,2 branched oligosaccharides obtained by maltose acceptor reaction with DSR-E glucansucrase (BioEcolans®, Solabia) are commercialized on the scale of around 80 t/year by SOLABIA for cosmetic applications and stimulation of beneficial skin flora (Sarbini et al., 2013).

Furthermore, glucansucrase can also be used to produce **glucoderivatives** and catalyse the **glucosylation** of diverse acceptors such as flavonoids to improve their solubility and protect them against oxidation (Malbert et al., 2018; Morel et al., 2017) or for the synthesis of glyco-co-polymers (André et al., 2018).

These main current and potential applications are summarized in the Table 5. Of note, the applications or potential applications of alternan and oligoalternans will be reviewed in detail in the following part of this literature review (see III. Structural characterization and potential applications of alternan, page 60).

In summary, the current investigations and number of patents related to dextrans and other α-glucans or glucooligosaccarides and glucoderivatives (18,111 patents since 2015, google patent consultation February 26, 2018), their chemical functionalization and their applications are impressively high, providing obvious evidences of α-glucan popularity and potentialities for innovation. In addition, the discovery of novel glucansucrases in the natural diversity as well as the engineering of these enzymes give today new opportunities to access to new biopolymers from a renewable resource and develop sustainable process targeting environmentally friendly and sustainable applications.

Table 5: α-Glucan applications. References: 1= (Soetaert et al., 1995); 2= (Valette et al., 1993); 3= (Djouzi et al., 1995); 4= (Wang et al., 2010); 5= (Leemhuis et al., 2013); 6= (Lynch et al., 2018); 7= (Monsan et al., 1995); 8= (Finkenstadt et al., 2011); 9= (Semyonov et al., 2014); 10= (Wangpaiboon et al., 2018)

Domain	Glucan	Applications
Health	Dextran	Blood plasma substitute (40-70 kDa dextran (clinical))[1]
		Microcarrier in tissue or cell culture[1]
	All (dextran, reuteran, alternan, isomalto/ maltopolysaccharides)	Fibre, potentially prebiotic* [2,3]
	Reuteran (*L. reuteri* TMW1.656)	Inhibition of enteroxigenic *E. coli* -induced hemagglutination of porcine erythrocytes[4]
	DEAE-dextran	Cholesterol lowering agent[1]
Food	Dextran	Gelling, viscosifying, emulsifying agent[5]
		Improving quality of sourdough breads, including gluten-free, (increase in dough viscosity, crumb softness, specific volume and decrease in firmness and staling rate. Prebiotic health benefits can also be added) and avoid the use of commercial texturizing additives (hydrocolloids)[6]
Feed	GOS	Improve the daily weight gain of pigs, broilers and calves[7]
Cosmetic	GOS	Promotion of the growth of beneficial skin lactic bacteria and prevention of the growth of detrimental bacteria[7]
Other	Dextran	Size exclusion chromatography (cross-linked dextran, Sephadex) [1]
		Aqueous two-phase separation systems (dextran/polyethyleneglycol)[1]
		Protection of steel against corrosion for *Ln. mesenteroides* NRRL B-1498 dextran[8]
	Dextran, Alternan	Nanoparticules forming[9,10]

Part 2- The story of alternansucrase

I. The alternansucrase from *Leuconostoc citreum* NRRL B-1355

I.1. Origin and discovery

The strains characterized as alternansucrase-producing strains were first discovered by Jeanes *et al*. (Jeanes et al., 1954). The screening of 96 polymer-producing LAB from sucrose carbon source highlighted three strains - *Leuconostoc mesenteroides* NRRL B-1355, NRRL B-1498, and NRRL B-1501 – that produced **two types** of α-glucans. These glucans were separated by fractional precipitation with ethanol. The **L** (for Less soluble) and **S** (for Soluble) **fractions** precipitated with 36-38% and 38-41% (v/v) of ethanol respectively (Figure 21) (Wilham et al., 1955). Using periodate oxidation, the most soluble α-glucans (recovered in the S-fraction) were shown to contain a higher percentage of **α-1,3 linkages** than the less soluble ones (Table 6). These first findings indicated as mentioned in the Part 1 that several GH70 encoding genes were likely present in these strains.

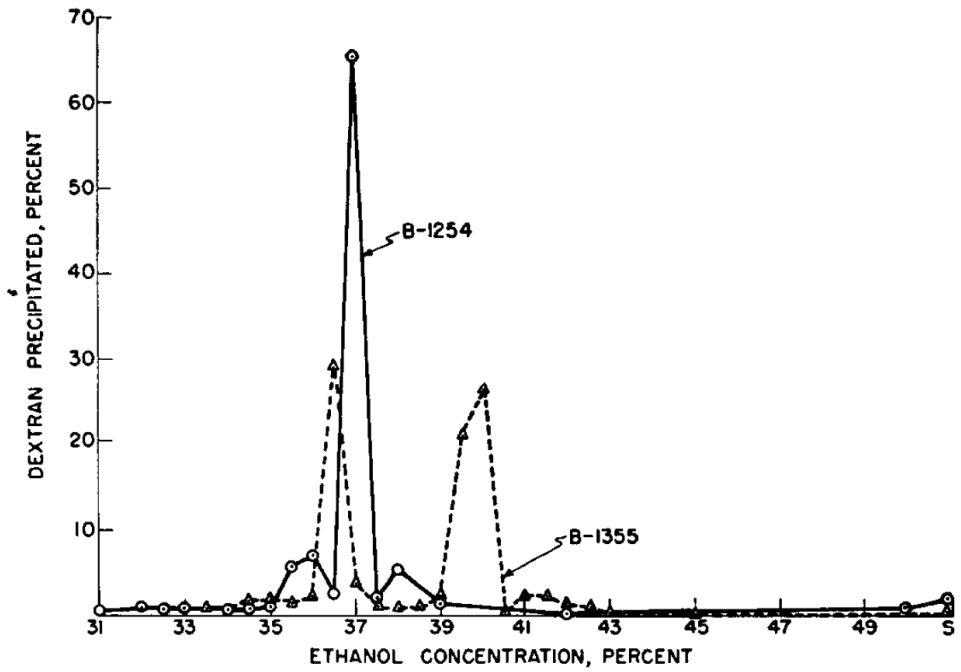


Figure 21: Fractional precipitation curves. From (Wilham et al., 1955)

Table 6: Polymer characterization of the –L and –S fractions of *Ln. mesenteroides* NRRL B-1355, 1498 and 1501. Adapted from (Jeanes et al., 1954)

Specie	Strain NRRL B-	Fraction	Type of links (%), periodate oxidation			1,3-like links (%), infrared absorption data	Appearance of dilute aqueous solution
			1,6-	1,4-like	1,3-like		
Leuconostoc mesenteroides	1355	L	88	9	3	n.d.	Very turbid
		S	57	8	35	35	Marked bluish opalescence
	1498	L	94	6	0	n.d.	Very turbid
		S	62	11	27		Marked bluish opalescence
	1501	L	93	7	0		n.d.
		S	65	15	20		Marked bluish opalescence

Then, several studies focused on the structural characterization of the **α-glucan fraction S** of *Ln. mesenteroides* NRRL B-1355. Acetolysis (Goldstein and Whelan, 1962), methylation (Seymour et al., 1977) and NMR (Seymour et al., 1976) experiments revealed a unique structure made of α-1,3 and α-1,6 osidic linkages alternation in the main chain. The fraction S was thus named **alternan** (more details on the polymer structure are given in the part III. Structural characterization and potential application of alternan, page 60) and the enzymatic activity responsible for its production, **alternansucrase** (ASR) (Côté and Robyt, 1982a). The enzyme responsible for the activity involved in the second fraction (–L fraction) was found to be a dextransucrase producing the same product than the DSR-S from *Ln. citreum* NRRL B-512F (Smith et al., 1994) or the DSRB from *Ln. citreum* NRRL B-1299: a soluble dextran with more than 95% α-1,6 linkages (Monchois et al., 1998). The enzyme was named DSRC (CAB76565.1), its gene was sequenced (Argüello-Morales et al., 2000a). Later, the presence of a third dextransucrase of higher molecular mass (240 kDa) producing dextran with α-1,2 branches was highlighted in *Ln. citreum* NRRL B-1355 strain (Smith et al., 1994, 1998). To note, the genome of *Ln. citreum* NRRL B-1355 has not been sequenced yet.

Jeanes *et al.* identified three different ASR-producer strains (*Ln mesenteroides* NRRL B-1498, B-1501 and B-1355, recently reclassified as *Ln. citreum* (Bounaix et al., 2010)). Using ASR sequence and genome mining, **four other** strains producing an enzyme highly similar to the ASR (more than 97% identity) were identified in *Ln. citreum* (Table 7). Among them, the *Ln. citreum* LBAE-C11 strain was isolated from ripe firm traditional wheat sourdough from the "Centre Technique de la Conservation des Produits Agricoles" (CTCPA, Auch, France) (Amari et al., 2015; Bounaix et al., 2010; Robert et al., 2009). In contrast, *Ln. citreum* KM20 is originated from Kimchi, a popular fermented Korean food made with vegetables, traditionally using Korean cabbage and/or East Asian giant white radish (Kim

et al., 2008) and *Ln. citreum* ABK-1 was isolated from a Thai dessert named Khow-tom-mud (Thai banana in sticky rice) (Wangpaiboon et al., 2018).

Table 7: BLASTp hits ranked by descending order of identity percentage using *Ln. citreum* NRRL B-1355 ASR as query.

Specie	Strain	Enzyme	Identity with B-1355 ASR (%)	Nb. of identical res. with B-1355 ASR	Nb. of res.	GenBank accession
Leuconostoc citreum/ mesenteroides	NRRL B-1355	ASR	100	2057	2057	CAB65910.2
	LBAE-C11	ASR	98	2013	2057	CCF26526.1
	KM20	ASR	98	2012	2057	ACA83110.1
	ABK-1	ASR	98	2012	2057	AIM52834.2
	EFEL 2700	ASR	98	2012	2057	WP_100666134.1
	NRRL B-1501	ASR	97	2003	2057	AGZ03657.1
	NRRL B-1498	ASR	97	2003	2057	AGZ03656.1
	NRRL B-1299	DSR-M	61	1313	2065	CDX66895.1

I.2. Native enzyme characterization

To produce the native ASR, Côté and collaborators first cultured *Ln. citreum* NRRL B-1355 in 10 L fermentor at 25°C on **sucrose-containing medium** in order to induce the enzyme production. Ethanol purification was performed to separate the S and L fractions and attempt enzyme purification, but only 12% of enzymatic activity was recovered. So, the culture supernatant was dialyzed and concentrated to yield an enzymatic preparation of **0.84 U.mL^{-1}**, which was submitted to **dextranase** digestion and **size exclusion** chromatography. This allowed the separation of the ASR from the DSR (Figure 22). Another purification method based on chromatography using O-(Phenoxyacetyl) cellulose (**PA-cellulose**) also enabled to separate the two activities (Figure 23) (Côté and Robyt, 1982a).

To note, **polysaccharides at high concentration** interfered with protein concentration determination (Côté and Robyt, 1982a), explaining why protein concentration and specific activity are not indicated.

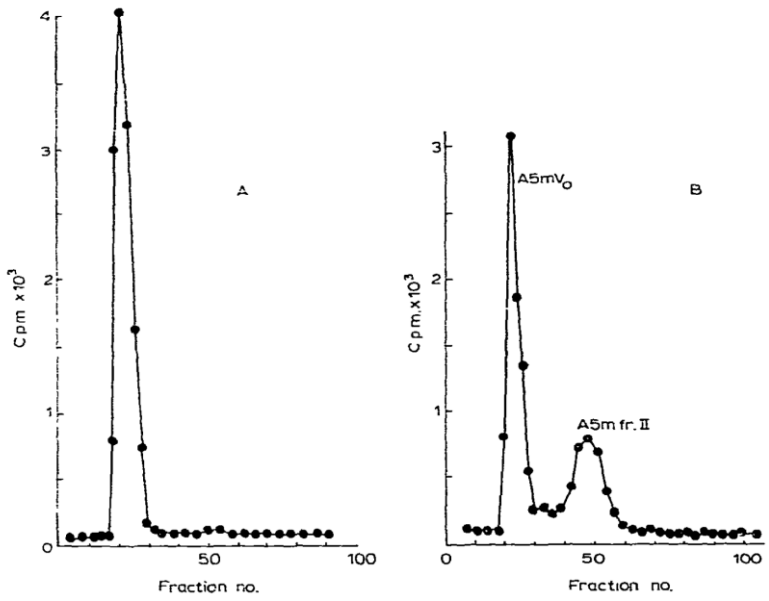


Figure 22: Elution on Bio-Gel A5m of (A) dialyzed and concentrated culture supernatant and (B) dextranase-treated, dialyzed and concentrated culture supernatant. From (Côté and Robyt, 1982a)

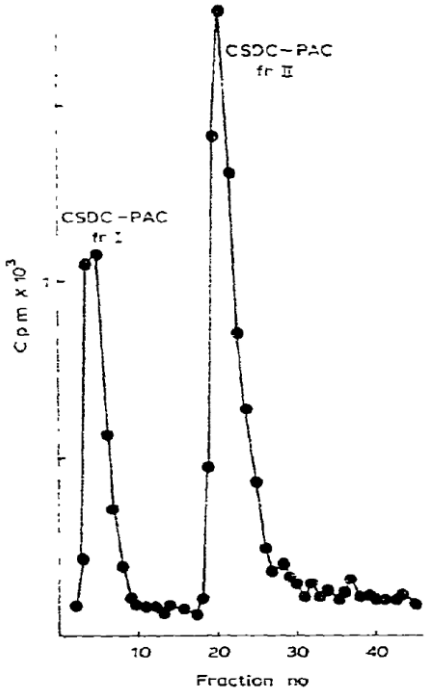


Figure 23: Elution on O-(phenoxyacetyl)cellulose. Vertical axis: proportional to glucansucrase activity (14C-glucose from 14Csucrose). From fraction 15, eluent switched from 20 mM pH 5.3 pyridine-acetate buffer to 0.5% triton X-100 in buffer. From (Côté and Robyt, 1982a)

A fed-batch production using an **alkaline sucrose solution** to regulate the pH enabled the enzyme production level to be enhanced up to **3.5 U.mL^{-1}** (López-Munguía et al., 1990; Monsan and Lopez, 1981). Aqueous two-phase partition between polyethylene-glycol and the dextran/alternan present in the culture broth enabled the recovery of both ASR and DSR enzymes in the **dextran/alternan** phase, demonstrating that alternan and/or dextran could **bind** to **ASR**. This first purification step was followed by a second one consisting of a **thermal treatment** for 20 minutes at 45°C. This treatment inactivated DSR while the ASR remained untouched. The profile of oligosaccharides produced from maltose with this enzymatic preparation (Figure 24b) showed an increase of oligoalternan production with a concomitant decrease of oligodextran synthesis and was similar to the profile obtained with the preparation purified by O-(Phenoxyacetyl) cellulose (PA-cellulose) (Figure 24c).

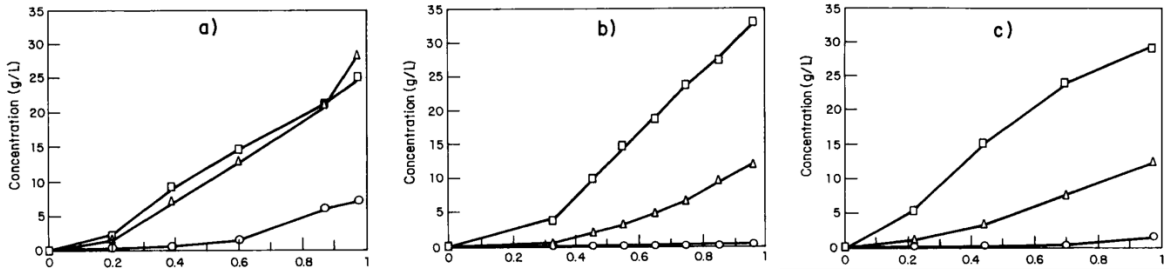


Figure 24: "Evolution of the main acceptor reaction products obtained in the presence of maltose (50 g/L) and sucrose (100 g/L) in batch reactions with 0.5 U/mL of enzyme: (a) supernatant enzymes of *L. mesenteroides* NRRL B-1355 cultures; (b) heat-treated alternansucrase after PEG extraction; (c) purified alternansucrase by PA-cellulose gel filtration as described previously. Circle: oligodextran of degree of polymerization (DP) 5; triangle: mixture of oligodextran DP4 and oligoalternan DP5; square: oligoalternan DP4." X-axis: conversion. From (López-Munguía et al., 1990)

Reaction products from maltose were also analyzed using HPLC, the purified enzymatic preparation of ASR and the *Ln. mesenteroides* NRRL B-512F dextransucrase. Oligoalternans were distinguished by comparison with dextransucrase linear products (Figure 25) (López-Munguía et al., 1993).

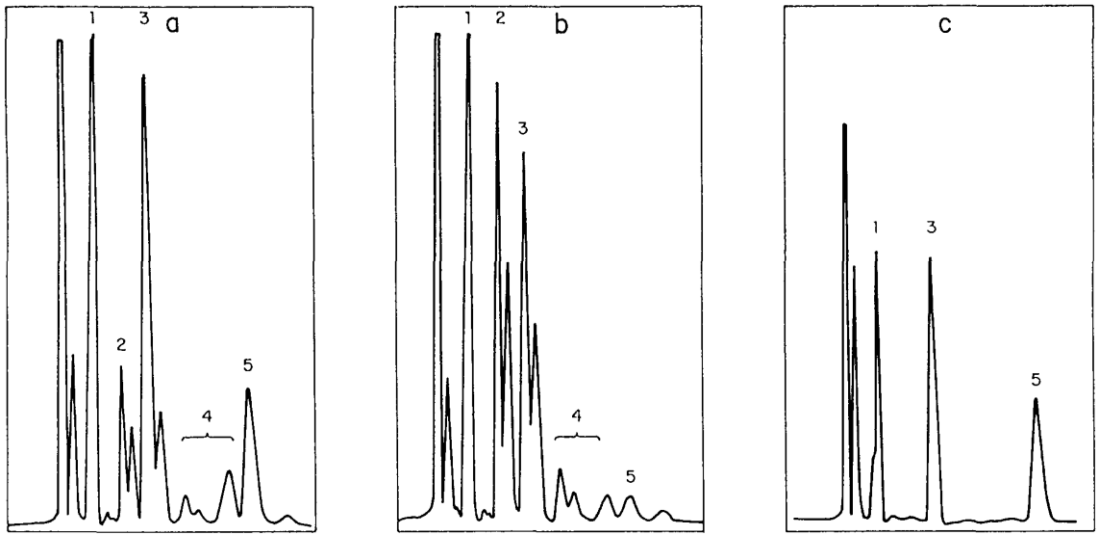


Figure 25: "Chromatograms of acceptor products obtained with different enzyme preparations: (a) Enzymes from *Ln. mesenteroides* NRRL B-1355 supernatant; (b) Purified alternansucrase; (c) *Ln. mesenteroides* NRRL B512F dextransucrase. (1= panose, 2= oligoalternan DP4,3; 3= oligodextran DP4 and/or oligoalternan DP; 4= oligoalternans , 5= oligodextran DP5)." From (López-Munguía et al., 1993)

Biochemical parameters of ASR

ASR is one of the most stable glucansucrase with an optimum temperature of **45°C**, eight degrees higher than that of DSR-S (López-Munguía et al., 1993). The optimum pH of ASR was found to be **5.5** using 20 mM phosphate buffer at room temperature (Côté and Robyt, 1982a; López-Munguía et al., 1993) at both 30°C and 40°C. Enzyme specific activity was not determined due to the presence of contaminant glucan after purification.

Different inhibitors of the native ASR were tested. 3-deoxy-3-fluoro-α-D-glucopyranosyl fluoride was found to fully inhibit ASR activity whereas Triton X-100, Tween 80, L-Serine, L-Threonine and Zwittergent 3-12 did not impact its activity (Côté and Robyt, 1982a).

I.3. Mutated strains

Several mutants of the *Ln. citreum* NRRL B-1355 strain were constructed to circumvent the problems of (i) sucrose-inducible *asr* gene expression resulting in the production of glucansucrases together with **contaminating dextrans** and **alternans** and (ii) **DSR** production **contamination**. For example, the constitutive mutant B-1355C was generated using ethyl methane sulfonate mutagenesis and produced **0.34 U.mL^{-1}** of glucansucrase activity without sucrose induction (Kim and Robyt, 1994). This mutant was subjected to another mutagenic treatment with N-methyl-N'-nitro-N-nitrosoguanidine and was further improved, leading to the B-1355C/CF10G3 mutant displaying **2.0 U.mL^{-1}** of glucansucrase activity (Kitaoka and Robyt, 1998).

The mutant strain B-21138 was obtained using UV mutagenesis and found to be genetically stable displaying a two-fold reduced glucansucrase activity with 88% representation of ASR versus 69% for the native strain (Leathers et al., 1995). This mutant strain was used to obtain the highly stable **B-21297** mutant strain which overproduces the ASR (four times more than B-1355) without production higher level of DSR (Figure 26) (Leathers et al., 1997).

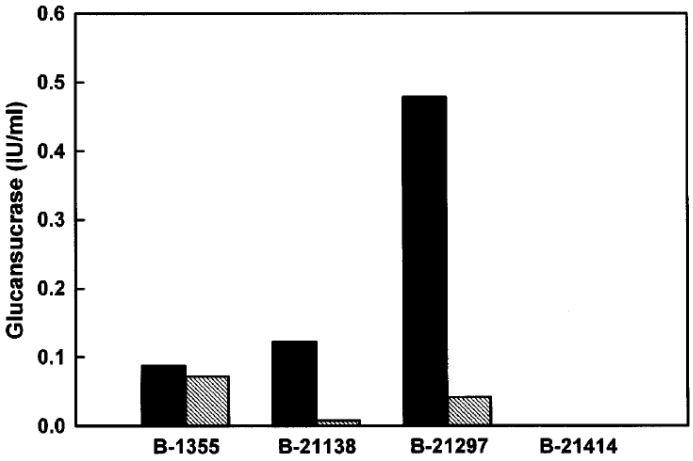


Figure 26: "Glucansucrase production by 2-day cultures of *Ln. mesenteroides* on liquid medium containing sucrose. Solid bars: alternansucrase activity. Hatched bars: dextransucrase activity". From (Leathers et al., 1997)

Finally, UV mutagenesis of the B-1355 strain led to the R1554 mutant (NRRL B-23191), which secreted ASR in the culture supernatant in soluble form and led to a preparation exhibiting **1.5 U.mg^{-1}** of proteins. Another mutant R1588 (NRRL B-23192) also obtained by UV mutagenesis produced a cell wall-associated form of ASR with a specific activity of **1.2 U.mg^{-1}** of dry matter (Smith et al., 1998).

II. Recombinant enzyme characterization and preliminary results on structure-function relationship studies

II.1. ASR primary structure

To facilitate ASR production, purification and initiate structure-function studies, the *asr* gene from *Ln. citreum* NRRL B-1355 was first **cloned** in a pGEM-T constitutive vector and **expressed in *E. coli*** JM109. No alternansucrase activity was detected in the culture supernatant, whereas it reached 160 $U.L^{-1}$ of culture in the sonicated cell extracts (Argüello-Morales et al., 2000b). The sequence of ASR revealed that the protein contains **2057** amino acids and has a predicted molecular mass of **228.971** kDa. It was the largest glucansucrase ever described at that time. Sequence analysis highlighted the presence of a 39 amino acid **signal peptide** (Nielsen program analysis (Nielsen and Krogh, 1998)). The residues **Asp635**, **Glu673** and **Asp767** were proposed to play the role of the catalytic triad (nucleophile, acid/base and TSS respectively) and the four conserved motifs related to the GH13 family were identified in the order II, III, IV and I, showing that the ASR also displays a **circularly permuted** $(\beta/\alpha)_8$ catalytic barrel like all the GH70 glucansucrases. When released, the ASR sequence distinguished from other glucansucrase sequences by several unique features including (i) a variable region (N-terminal) **100 amino acids longer** those found in the other glucansucrases , (ii) the presence of ASR-unique sequence stretches in motif III (^{676}**NGK**^{678}) and motif IV (^{768}**YDA**^{770} instead of the common SEV), (iii) a very **large GBD** containing around 700 amino acids which was 200 amino acids longer than the GBD of other glucansucrases (C-terminal) (Figure 27) (Argüello-Morales et al., 2000b). Additionally, seven repeats containing the motif "**APY**" were found in the C-terminal part of the protein (Janeček et al., 2000; Joucla et al., 2006) (Figure 27). These APY repeats were also identified in the inulosucrase EIS from *Ln. citreum* CW28 (Olivares-Illana et al., 2003), in the α-1,3 branching sucrase BRS-B from *Ln. citreum* NRRL B-742 (Vuillemin et al., 2016), in the dextransucrase DSR-M and in the α-1,2 branching sucrase BRS-A from *Ln. citreum* NRRL B-1299 (Passerini et al., 2015). To note, the APY repeat deletion in the EIS, BRS-B or DSR-M enzyme did not impact significantly the specificity and efficiency of the enzymes (Claverie et al., 2017; Olivares-Illana et al., 2003; Vuillemin et al., 2016). In inulosucrase EIS, the thermal stability was twice reduced at 40°C with a half-life decrease from 70 minutes for the full-length enzyme to 20 minutes for the APY-truncated mutant (Olivares-Illana et al., 2003) indicating that APY motifs stabilized the enzyme. No specific function could be attributed to these repeats for alternansucrase, glucansucrase or branching sucrase as their deletion did not severely impacted their properties.

Cell-Wall (CW) binding repeats, that were previously shown to be involved in glucan binding in other glucansucrases (see II.2.2. Repeated sequences in the glucan binding domain, page 17), were also identified in both the variable region (10 repeats) and in the C-terminal region (seven repeats) (Figure 27) (Joucla et al., 2006).

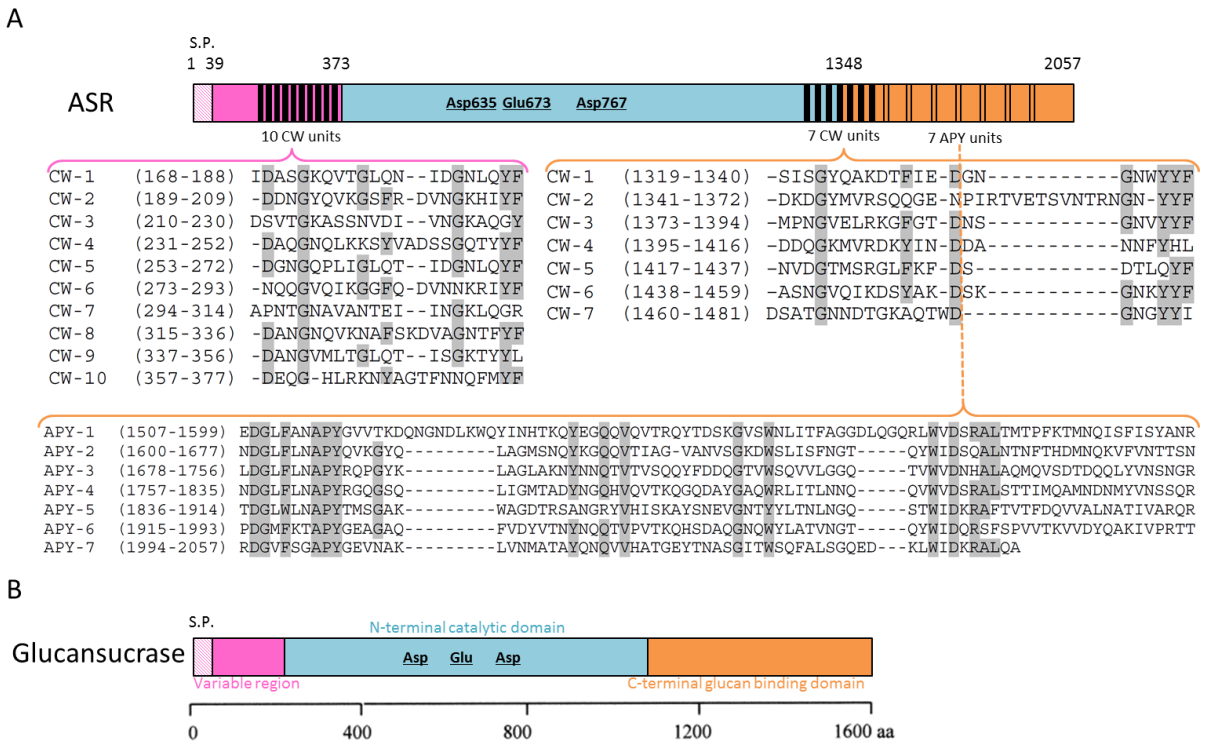


Figure 27: (A) ASR sequence and identified putative cell-wall (CW) and APY repeats. Domains delimitation from (Argüello-Morales et al., 2000b). Adapted from (Joucla et al., 2006).
(B) Schematic representation of glucansucrases sequence for which encoding gene has been cloned before 1998. Adapted from (Monchois et al., 1999b)

ASR recombinant production from a pBad vector expressed in *E. coli* TOP10 reached a 4-fold increase with **661 U.L^{-1}**. Only 26% of the total activity was recovered in the soluble fraction (Joucla et al., 2006). The same ASR was also produced using pE-SUMO plasmid in *Lactococcus lactis* LM0230 as a host. The expression in *E. coli* using this plasmid only yielded **insoluble** protein (Côté et al., 2017). Production level in *L. lactis* was not indicated.

II.2. Truncated version of previously characterized ASR

To overcome the problem of protein solubility when produced with *E. coli*, a truncated enzyme **ASR-C-APY-del** (Met1-Gly1425) deleted of the seven APY repeats and of two and a half CW repeats was constructed. The ASR C-APY-del was produced from the *pBad* plasmid expressed in *E. coli* TOP10 and a level of **774 U.L^{-1}** of culture was obtained with 73% of activity recovered in the **soluble** fraction. In addition, SDS-PAGE showed that the enzyme was less degraded and kept the same linkage specificity as the full-length enzyme (Figure 28) (Joucla et al., 2006).

ASR C-APY-del was shown to produce, from sucrose, a bi-modal population of glucan, comprising a **High Molar Mass** (HMM) fraction of 1,700,000 g.mol^{-1} and a **Low Molar Mass** (LMM) fraction of 1,300 g.mol^{-1} as estimated by Sixe Exclusion Chromatography calibrated with dextran standards (Figure 28) (Joucla et al., 2006).

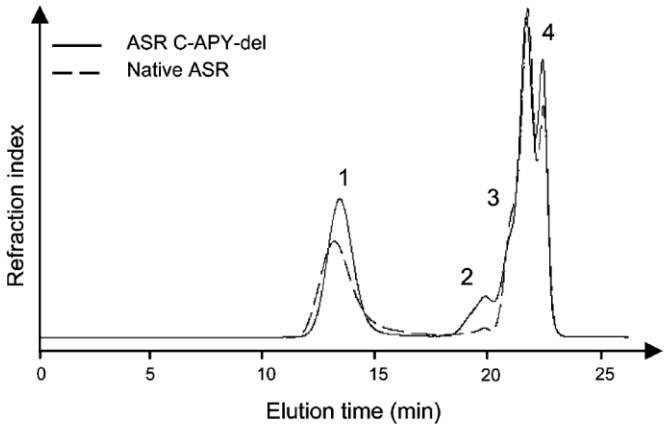


Figure 28: "High performance size exclusion chromatography analysis of products synthesized by the native alternansucrase from *L. mesenteroides* NRRL B-23192 and ASR C-APY-del. 1: Polymer of 1700 kDa, 2: oligosaccharides of 1.3 kDa, 3: disaccharides, 4: monosaccharides. From (Joucla et al., 2006)

The recombinant ASR C-APY-del was purified on a Ni-NTA column using a N-terminal His-tag. The pure enzyme displayed a specific activity of **160 U.mg^{-1}** of protein (the concentration being determined using the Bradford method) corresponding to a turnover of 404 s^{-1}, which makes ASR C-APY-del a very efficient enzyme compared to other glucansucrases (Joucla et al., 2006). In comparison, a K_m and a k_{cat} value of 32.2 ± 3.2 mM and 290 ± 12 s^{-1} were determined for the recombinant ASR from *Ln. citreum* ABK-1 (Wangpaiboon et al., 2018). The optimum temperature of the recombinant ASR C-APY-del was found to be 45°C (Joucla, 2003), the same as for the native full-length enzyme (López-Munguía et al., 1993) and the recombinant ASR from *Ln. citreum* ABK-1 (Wangpaiboon et al., 2018).

ASR was already known as one of the most stable enzyme in the GH70 family. Notably, ASR truncation (ASR-C-APY-del) impacted positively the enzyme stability with an increase of its half-life time at 30°C, 40°C and 50°C (Table 8) determined with the recombinant enzymes (Joucla, 2003), contrary to what was observed with the inulosucrase EIS (Olivares-Illana et al., 2003).

Table 8: Half-life time of full-length ASR versus ASR C-APY-del at 30°C, 40°C and 50°C (Joucla, 2003)

	30°C	**40°C**	**50°C**
Full-length ASR	75 h	6 h	2 h
ASR C-APY-del	87 h	20 h	10 h

Bringing together all these data, the following ID card of *Ln. citreum* B-1355 ASR production and characterization can be provided (Table 9).

Table 9: Alternansucrase ID card. References: 1= (Joucla et al., 2006); 2= (Joucla, 2003); 3=(Argüello-Morales et al., 2000b); 4= (López-Munguía et al., 1990); 5= (López-Munguía et al., 1993) ; 6= (Côté and Robyt, 1982a).

ID CARD					
Enzyme name	Alternansucrase				
Enzyme nickname	ASR				
Class	GH70 - Glucansucrase				
Origin	*Ln. citreum* NRRL B-1355				
Enzyme Production					
Construction	ASR C-APY-del	ASR			
Production	Recombinant	Recombinant		Native	
Host	*E. coli* TOP10	*E. coli* TOP10	*E. coli* JM109	*Ln. citreum* NRRL B-1355	*Ln. citreum* NRRL B-1355
Plasmid	pBad	pBad	pGEM-T	/	/
Tags	His (C-terminal)	His (C-terminal)	none	/	/
Production level (U.L^{-1} culture)	774[1]	661[1]	160[3]	703.5[4]	840 (total GT activity)[6]
S.A. on sucrose (U.mg^{-1})	160[1]	*n.d.*	*n.d.*	*n.d.*	*n.d.*
Kinetic parameters	*n.d.*	*n.d.*	*n.d.*	*n.d.*	*n.d.*
Melting temperature (°C)	*n.d.*	*n.d.*	*n.d.*	*n.d.*	*n.d.*
Optimum temperature (°C)	45[2]	*n.d.*	*n.d.*	45[5]	*n.d.*
Half-life times	87h (30°C)[2] 20h (40°C)[2] 10h (50°C)[2]	75h (30°C)[2] 6h (40°C)[2] 2h (50°C)[2]	*n.d.*	*n.d.*	*n.d.*
Optimum pH	*n.d.*	*n.d.*	*n.d.*	5.5[5]	5.5[6]
Size (without tags)	1,425[1]	2,057[3]	2,057[3]	*n.d.*	*n.d.*
Molar mass (kDa)	161.5	229.0[3]	229.0[3]	*n.d.*	*n.d.*
Product					
HMM polymer mass (g.mol^{-1})	1,700,000[1]	1,700,000[1]	*n.d.*	*n.d.*	*n.d.*
HMM polymer linkages	42% α-1,3[1] 58% α-1,6[1]	*n.d.*	*n.d.*	*n.d.*	*n.d.*

II.3. Enzyme mechanism and mutants

Comparison with other GS sequences guided a few mutagenesis studies to get insight into the structural features involved in specificity.

Three mutants were constructed using the *asr-C-APY-del* as template. In the first one, the ASR-unique motif **YDA** downstream the TSS Asp767 in the conserved motif IV was replaced by the triplet SEV mainly found in dextransucrases. The mutation resulted in a mutant showing 1.4% residual activity and synthesizing both oligoalternans (glucooligosaccharides with α-1,6 and α-1,3 osidic linkages) and oligodextrans (with only α-1,6 linkages) from maltose acceptor reaction. Notably, the YDA-SEV mutant was unable to elongate oligoalternans above DP4, contrary to the wild type enzyme, what resulted in its accumulation. In contrast, the oligodextrans were further elongated via α-1,6 linkage synthesis (Moulis et al., 2006). The two other ASR C-APY-del mutants also targeted unique features of ASR. The specific insert loop ^{1109}**NYGGM**^{1113} was deleted and the triplet ^{676}**NGK**^{678} in motif III was changed with the SGN triplet, often encountered in DSR sequences. Both mutants were severely affected and showed 6% and 9% of residual activity compared to the wild type, respectively (Joucla, 2003).

The full-length ASR was also recently used as template to introduce mutations at position **544** (in the loop B1 of domain B) and replace the **leucine** residue by a glutamate, a serine or a proline. The mutants were expressed in *Lactococcus lactis* and compared to the wild type for their ability to synthesize isomelezitose. The best yield was reached with the mutant Leu544Pro (23% yield from sucrose versus 2.5% for the wild type) (Côté et al., 2017).

In conclusion, data issued form mutagenesis studies are scarce and are still lacking to elucidate the mechanism behind the linkage specificity of ASR. One hypothesis was advanced, in which the glucose moiety in the +1 subsite would not be well stabilized, whereas the glucosyl moiety in +2 subsite (identified at that time by analogy with α-amylase subsite) would interact with **Tyr768** (of the motif YDA) and **orient** the position of glucosylation occurring in **+1** subsite and thus the linkage specificity (Moulis et al., 2006). However, in the absence of three dimensional structures this hypothesis could not be assessed and remained **speculative**.

III. Structural characterization and potential applications of alternan

III.1. Alternan (HMM) polymer

III.1.1. HMM alternan stucture

Since its discovery in 1954, different analytical methods were used to gain insight in the alternan structure produced by the strain B-1355:

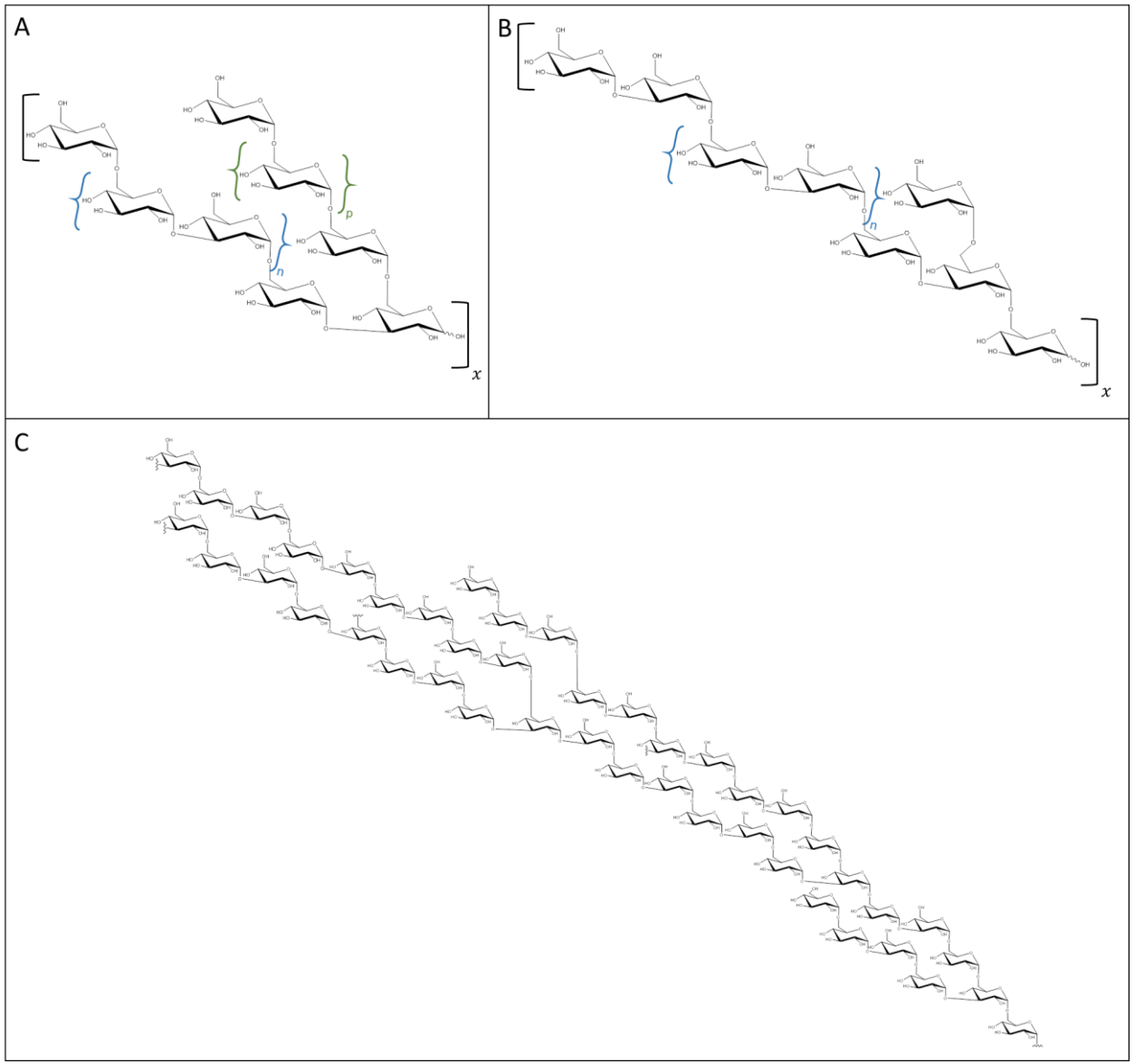
- **periodate oxidation**; that revealed 35% of α-1,3 linked units, 8% α-1,4 linked units and 57% α-1,6 linked (Jeanes et al., 1954)
- **optical rotatory shifts** of the cuprammonium solution; that highlighted non α-1,6 linked unit in the main chain (Scott et al., 1957)
- **infrared** absorption data type II; that gave a percentage of 35% of α-1,3 linked units (Jeanes et al., 1954)
- partial **acetolysis** (technique that preserve non α-1,6 bonds); that resulted in 21% glucose, 2% isomaltose, 20% nigerose, 2% isomaltosylglucose and notably, no nigerotriose. Thus α-1,3 linkages are not arranged consecutively in the sequence (Goldstein and Whelan, 1962). With acetolysis, the products obtained are glucose (47%) and nigerose (31%) (Torii and Sakakibara, 1974). This experiment was repeated and again, only glucose and nigerose (25%) were obtained (Joucla et al., 2006) confirming the isolation of α-1,3 linked glucosyl units.
- **Smith reaction**; that gave glycerol (from non-reducing end groups and α-1,6 linked units), 1-O-glucopyranosylglycerol (from a 3-substituted glucose with 1,6 linkages with the next glucose) and showed that the majority of α-1,3 bonds are isolated from each other (no nigerose obtained) (Goldstein and Whelan, 1962). Only glucose (54.6%) and glycerol (45.4%) were obtained, which is in accordance with the absence of consecutive α-1,3 bonds in the polymer (Misaki et al., 1980).
- **Methylation**; similar results were obtained showing the presence of four type of linked glucose (Table 10).
- **NMR**; spectra were compared to other strains to assign peaks (Seymour et al., 1976, 1979b, 1979a) and the signals were integrated to evaluate linkage percentages: the integration of ^{1}H spectra gave 42% of α-1,3 linkages and 58% of α-1,6 linkages (Joucla et al., 2006). Similar values were obtained recently with a ratio of α-1,3:α-16 linkages of 1:1,48 (Dertli et al., 2018) which correspond to 40% and 60% of α-1,3 linkages and of α-1,6 linkages, respectively. These values are exactly the same as those reported for the alternan produced by the ASR from *Ln. citreum* ABK-1 (Wangpaiboon et al., 2018).

- **Enzymatic digestion** using different dextranase as the isomaltodextranase from *Arthrobacter globiformis* (Sawai et al., 1978), the exodextranases from *S. mitis 439* and *A. globiformis* T6 (Hare et al., 1978) or the exo-G_2-dextranase of *A. globiformis* T6. Surprisingly, the latter resulted in the release of 85% isomaltose and 15% α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc (Misaki et al., 1980) indicating that this enzyme probably cleaved both α-1,3 and α-1,6 linkages. The endo-dextranase from Sigma and ICN Pharmaceuticals (Côté and Robyt, 1982a) and the dextranase from *Penicillium* sp. or commercial dextranases from *Penicillium* sp. (Sigma) and *Chaetomium erracticum* (Bio-Cat Inc.) were also used and alternan was found to be not fully resistant to the digestion by these dextranases. These results highlighted the existence of localized region of consecutive α-1,6 linkages (Leathers et al., 2002, 2009).

Table 10: Methylation analysis of HMM alternan

ASR Strain	Glucopyranose methylation (%)				
	terminal	3-substitued	6-substitued	3,6-substitued	Reference
B-1355	6.9	35	46.9	11.2	Seymour, 1977 (six)
B-1355	11	34	44	11	Hare et al., 1978
B-1355	9.8	35.3	45.1	9.8	Misaki et al., 1980
B-1355	10	35	45	10	Leathers et al., 1995
B-1355	13.3	25	53.6	8.1	Moulis et al., 2006
B-1355	7.7	31.3	49.3	11.7	Leathers, 2009
B-1355	13.1	35.4	40.2	11.3	Dertli et al., 2018
B-1501	14.6	31.5	43.5	10.5	Dertli et al., 2018
ABK-1	16	26	47	11	Wangpaiboon et al., 2018

All these studies concluded to the presence of alternating α-1,3 and α-1,6 linkages in the main chain, as well as some consecutive α-1,6 linkages and around 10% of branches. The proposed models are illustrated in the Figure 29.


A
B
C

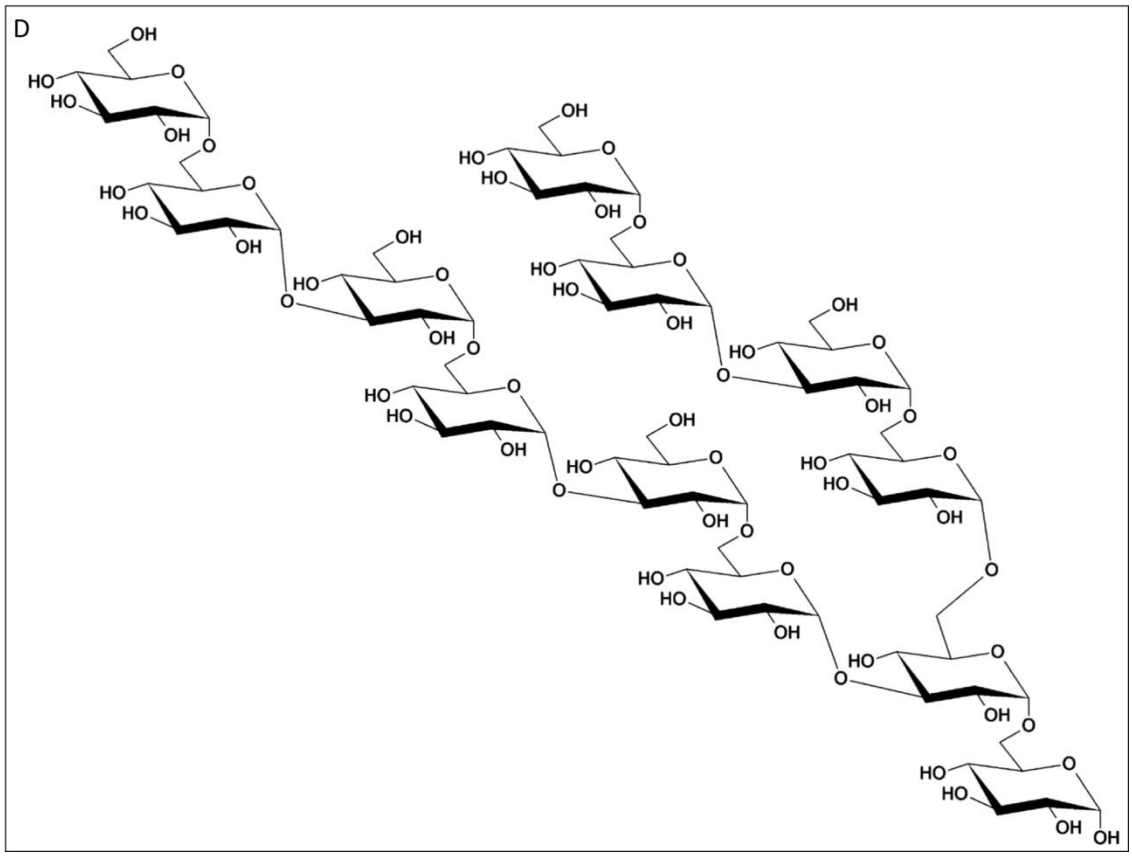


Figure 29: Proposed models of alternan structure. (A) Adapted from (Seymour et al., 1979b); n=5.1 and p=0.7. (B) Adapted from (Seymour et al., 1977) n=3. (C) Adapted from (Misaki et al., 1980). (D) Proposed structure of alternan building block from (Côté, 2002). Adaptations constructed using Chemdraw.

III.1.2. HMM alternan: size and application

The HMM fraction of alternan from *Ln. citreum* NRRL B-1355 was determined to be of around 50 millions g.mol^{-1} using off-line MALS, SEC/MALS or HDC/MALS (Hydrodynamic Chromatography). Slightly higher values were reported for the HMM polymer produced from the mutant strain B-21297 (59,800,000 ± 1,600,000 g.mol^{-1}) (Isenberg et al., 2010). Alternan finds applications in several fields. As, the polymer is **more soluble** in water and less viscous than dextrans, it was shown to be a good substitute of **Arabic gum** for food applications after a sonication treatment used to lower its molar mass (Côté, 1992) or endodextranase treatment (Leathers et al., 2002, 2009). Additionally, the hydrolysis of alternan with isomaltodextranase also lowered its molecular weight to the equivalent of DP 3500 and this "limit alternan" was rheologically comparable to **maltodextrins** of DP 10 (Côté, 1992). Recently, the alternan produced by *Ln. citreum* ABK-1 ASR was shown to form **nanoparticles** or **films** of interest for potential applications in nanotechnology (Wangpaiboon et al., 2018).

III.2. Alternan (LMM) oligomer

III.2.1. LMM alternan and oligoalternan production

As previously mentioned, ASR also synthesizes a LMM fraction (of around 1,300 g.mol^{-1}) from sucrose and can also catalyzes **transglucosylation** reaction from sucrose to many different types of sugar **acceptors**. The first acceptors to be assessed were maltose, nigerose, methyl α-D-glucoside, isomaltose, D-glucose, methyl β-D-glucoside (Côté and Robyt, 1982b). Then, maltodextrin, maltitol and isomaltooligosaccharides were tested (López-Munguía et al., 1993). Côté *et al.* (2003) extended the list and ranked 73 acceptors on their ability to reduce alternan synthesis (Table 11). In this classification maltose, maltitol, nigerose, gentiobiose and panose are placed in the top five most efficient acceptors leading to oligosaccharides at the cost of alternan formation.

Table 11: Acceptor rank. From (Côté et al., 2003)

Monomers	*Alternan (Rel. %)*	*Di- & Oligomers*	*Alternan (Rel. %)*
dulcitol	101	α,α-trehalose	102
mannitol	100	lactulose	91
myo-inositol	99	galactosyl-arabinose	90
D-arabinose	99	melezitose	90
D-lyxose	99	*cyclo*Glc$_4$ (cycloalternan)	90
xylitol	97	stachyose	89
D-sorbose	97	kojitriose	86
L-rhamnose	96	β-dodecyl maltoside	83
β-methyl-D-xyloside	96	leucrose	81
β-methyl-D-galactoside	94	sophorose	80
L-fucose	94	lactose	79
D-ribose	93	planteose	76
N-acetyl-D-mannosamine	93	cellobiose	72
L-lyxose	92	palatinose	69
D-altrose	92	raffinose	67
D-mannose	90	gentianose	66
N-acetyl-D-glucosamine	90	maltotriose	65
2-deoxy-D-galactose	89	laminaribiose	58
β-methyl-D-mannoside	89	turanose	57
N-acetyl-D-galactosamine	88	6-*O*-α-D-Glc-trehalose	56
L-sorbose	88	6,6'-di-*O*-α-D-Glc-trehalose	54
D-xylose	86	isomaltotriose	50
D-psicose	85	melibiose	45
D-galactose	83	isomaltose	44
sorbitol	82	theanderose	42
2-deoxy-D-ribose	81	6"-*O*-α-D-Glc panose	36
D-fructose	78	3'-*O*-α-D-Glc isomaltose	31
D-talose	77	kojibiose	27
L-arabinose	75	panose	25
D-allose	75	gentiobiose	25
α-methyl-D-galactoside	73	nigerose	23
α-methyl-D-mannoside	70	maltitol	18
D-tagatose	65	maltose	11
L-glucose	63		
D-quinovose	61		
β-methyl-D-glucoside	60		
2-deoxy-D-glucose	59		
D-glucose	53		
β-octyl D-glucoside	52		
α-methyl-D-glucoside	20		

III.2.2. LMM alternan and oligoalternan structure

Quantitative ^{1}H-NMR analysis of LMM alternan of average DP33 produced by *Ln. citreum* NRRL B-1355 ASR gave 63% of α-1,6 linkages and 37% of α-1,3 linkages. Additionally, methylation experiment of the same fraction revealed the presence of 8.8% terminal glucopyranose residue, 21.6% 3-substituted glucopyranose, 62.3% 6-substituted glucopyranose, 6.9% 3,6-substituted glucopyranose (Grimaud et al., 2018). The percentage of di-substituted glucose signaling the presence of branching is thus slightly lower in LMM alternan than in HMM alternan (the average for B-1355 alternan being 10.7 ± 0.6 %). Methylation of the LMM alternan fraction only purified from leucrose and fructose revealed the presence of 25.8% terminal glucopyranose residue, 20.5% 3-substituted glucopyranose, 48.8% 6-substituted glucopyranose and 4.9% 3,6-substituted glucopyranose (Moulis et al., 2006).

Côté and collaborators characterized the main products obtained from the glucosylation of **maltose**, the most efficient acceptor. The acceptor reaction products were purified by preparative HPLC and analyzed by NMR, methylation and enzymatic digestion. This work allowed the structural characterization of the maltose acceptor products until DP 8 and led to the model shown in Figure 30 and Table 12

Table 12. From this model, the authors suggested that ASR cannot form **consecutive α-1,3 linkages**, which is in agreement with the fact that consecutive α-1,3 linkages were never found in the polymer either. They also proposed that ASR could not synthesize more than **two consecutive α-1,6 linkages** (Côté and Sheng, 2006).

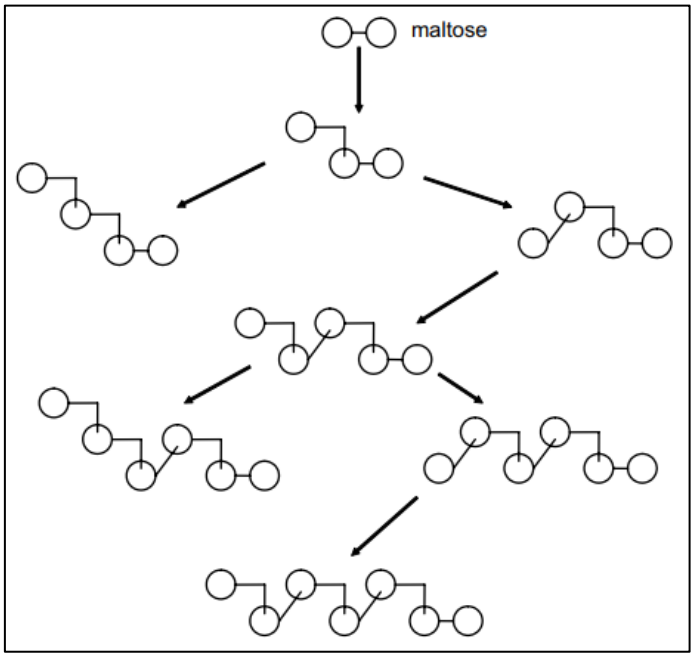


Figure 30: Model of oligoalternan formation from maltose from (Côté and Sheng, 2006). Circles: α-D-glucopyranosyl units. (/): α-1,3 linkage; (—): α-1,4 linkage; (┐): α-1,6 linkage.

Table 12: Structures corresponding to Figure 30. References: (1)= (Côté and Robyt, 1982b); (2)= (López-Munguía et al., 1993); (3)= (Côté and Sheng, 2006); (4)= (Côté et al., 2008)

Acceptor	Maltose
DP3	Panose (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc)[1,2]
DP4	α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[1,2] α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[1]
DP5	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[3]
DP6	α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[3] α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[3]
DP7	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[3]
DP8	α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[4] Three minors: α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc[4] Two not known (maybe branched)[4]

Certain acceptor products synthesized acceptor reactions with saccharides different from maltose were also characterized. The data from the different studies are sum-up in the following table and they are ranked from the most efficient acceptors to the least efficient ones (Table 13). Notably, no **consecutive α-1,3 linkage** between glucosyl residues were ever identified in the acceptor reaction products but **three consecutive α-1,6 linkages** were sometimes observed like with α-methyl-D-mannoside or α-butyl glucopyranoside acceptor reactions.

Table 13: Oligoalternan structure. References: (1)= (Côté et al., 2003); (2)= (Côté and Robyt, 1982b); (3)= (Argüello Morales et al., 2001); (4)= (Côté et al., 2008); (5)= (Côté, 2009); (6)= (Côté and Dunlap, 2003); (7)= (Côté et al., 2009) ; (8)= (Richard et al., 2003)

Alternan produced (Rel. %)[1]	**Acceptor**	**DP2**	**DP3**	**DP4**	**DP5**
18	Maltitol		Panitol[1]		
20	α-methyl-D-glucoside	α-methyl-D-isomaltoside[2]			
23	Nigerose		α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-D-Glc[2]	α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-D-Glc[2] α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-D-Glc[2]	
25	Gentiobiose		α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→6)-D-Glc[1,5]	α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→6)-D-Glc[5] α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→6)-D-Glc[5]	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→6)-D-Glc[5]
44	Isomaltose				
45	Melibiose		α-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→6)-D-Glc[1] minor product : α-D-Glc*p*-(1→4)-α-D-Gal*p*-(1→6)-D-Glc[1]		
53	D-glucose	Isomaltose[2]	α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-D-Glc[2] Isomaltotriose[2] Minor compound[2]		
60	β-methyl-D-glucoside	β-methyl-D-isomaltoside[2,6]	Methyl β-isomaltotrioside[6] methyl α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-β-D-Glc[6]		
67	Raffinose			α-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1↔2)-β-D-Fru*f*[1,7] α-D-Glc*p*-(1→4)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1↔2)-β-D-Fru*f*[1,7] α-D-Glc*p*-(1→6)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1↔2)-β-D-Fru*f*[7]	10 pentasaccharides[7]
70	α-methyl-D-mannoside	Methyl α-D-Glc*p*-(1→6)-α-D-mannoside[6] Minors:	Methyl α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-mannoside[6] Minors: 3,6-di-O-substitued mannoside[6]	Methyl α-D-Glc*p*-(1→6)-α-D-Glcp-(1→6)-α-D-Glc*p*-(1→6)-α-D-mannoside[6]	

		Methyl α-D-Glc*p*-(1→2)-α-D-mannoside[6] Methyl α-D-Glc*p*-(1→3)-α-D-mannoside[6]			
72	Cellobiose		α-D-Glc*p*-(1→2)-β-D-Glc*p*-(1→4)-D-Glc*p*[3] α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→4)-D-Glc*p*[3]	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→4)-β-D-Glc*p*[3]	α-D-Glc*p*-(1→3)-[α-D-Glc*p*-(1→6)]-α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→4)-D-Glc*p*[3] or α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→4)-D-Glc*p*[3]
73	α-methyl-D-galactoside	Major: Methyl α-D-Glc*p*-(1→4)-α-D-galactoside[1,6] Minor: Methyl α-D-Glc*p*-(1→3)-α-D-galactoside[1,6]			
78	D-fructose	Leucrose[2] Isomaltulose[4]	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→5)-D-Fru*p*[4] isomelezitose[4] isomaltotriulose (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Fru*p*)[4]		
n.d.	α-methyl-D-allopyranoside	Methyl α-D-Glc*p*-(1→6)-α- D-allopyranoside[6]	Methyl α-D-Glc*p*-(1→6)- α-D-Glc*p*-(1→6)-α-D-allopyranoside[6]		
n.d.	β-methyl-D-allopyranoside	Methyl α-D-Glc*p*-(1→6)-β-D-allopyranoside[6]	Methyl α-D-Glc*p*-(1→6)- α-D-Glc*p*-(1→6)-β-D-allopyranoside[6]		
n.d.	α-butyl glucopyranoside	α-D-Glc*p*-(1→6)-O- butyl-α-D-Glc*p*[8]	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-O- butyl-α-D-Glc*p*[8]	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-O- butyl-α-D-Glc*p*[8]	

III.2.2. LMM alternan and oligoalternan applications

Alternan could be used as **prebiotic** in certain food, feed or cosmetic preparations (Côté, 1992).

Mixture of oligoalternans produced from acceptor reaction with maltose (DP 3-7), melibiose (DP 2-4), raffinose (DP 3-5), gentiobiose or maltitol were tested for their ability to enhance the growth of **probiotic** (*e.g Bifidobacterium* sp., *Lactobacillus* sp.) or **undesirable** bacteria (*e.g. Bacteroides thetaiotaomicron, Enterobacter aerogenes, Clostridium perfringens, Salmonella cholerasuis)*. The results were comparable with those obtained with commercial well-known fructo-oligosaccharides such as those found in Neosugar® products (Côté et al., 2003). The ASR-derived oligosaccharides supported the *in vitro* growth of most of *Bifidobacterium* strains, but not that of coliforms or pathogenic bacteria. However, they were not metabolized by *Lactobacillus* sp. (Holt et al., 2005). More precisely, the specific activities of the α-galactosidase and α-glucosidase secreted by *Bifidobacterium adolescentis* were significantly enhanced when the bacteria were grown on oligoalternans compared to the level obtained on fructooligosaccharides, gentiobiose, maltitol, maltose or raffinose; thus proving an increased metabolic activity represented by α-galactosidase and α-glucosidase activities (Holt et al., 2008). Such promising results led to the patent US7182954B1 (Côté and Holt, 2007). To determine whether the oligosaccharide size influenced the prebiotic effect, the population of oligoalternans obtained from maltose acceptor reaction was fractionated into six fractions varying in their average DP. The growth of most bacteria was enhanced by the oligosaccharides of short DP and the optimal growth was obtained with the DP 3. The oligoalternans of higher DPs were more selective. The **prebiotic index** (PI), which is calculated according to the changes in key bacterial groups (*Bifidobacteria, Lactobacilli, Clostridia and Bacteroides*) during fermentation (Palframan et al., 2003) was determined for each fraction and compared to that of fructooligosaccharides (FOS). The highest the PI is, the highest is the prebiotic effect. All the fractions harbored a good PI except the mixture of highest DP (≈ DP 7.4). The best PI values were reached with fractions of DP3 and DP4 (Sanz et al., 2005a). A similar method was applied to oligoalternans produced from gentiobiose (β-D-Glc*p*-(1→6)-D-Glc) acceptor reaction and compared to gentiooligosaccharides (GEOS). The latter are potential prebiotics but have a low selectivity except for pure DP2 or DP3. ASR-glucosyltated gentiobiose showed a higher prebiotic selectivity than GEOS of similar DP and the highest PI was obtained with the fraction containing mainly DP 4 (Sanz et al., 2006). Another asset of ASR glucosylation of gentiobiose is that it removed the bitter **aftertaste** of the trisaccharide (Côté, 2009). The raffinose acceptor products were also fractionated and tested in *in vitro* fermentation using human feces. The oligosaccharides of DP 4 to DP 6 showed enhanced bifidogenic activity together with an increase in the gas volume compared to lactulose and raffinose.

However, the gas volume was comparable to that obtained with commercial inulin (Hernandez-Hernandez et al., 2011).

Of note, the oligoalternans produced from maltose acceptor reaction are one component of the commercial sweetener **Sucromalt** (Xtend, Cargill Inc., Wayzata, MN, USA) with a low-glycemic index (with leucrose and fructose) (Grysman et al., 2008; Vanschoonbeek et al., 2009). These oligosaccharides are slowly but completely digested by human and mice microbiota (Hasselwander et al., 2017). The preparation of this sweetener was patented (US8512739B2) and is commercialized by **Cargill** using (Carlson and Woo, 2013). The ASR is also involved in the patent WO2006088884A1 from Cargill again that aims to describe the methods for making low-glycemic syrups (LGS) of oligoalternans (Carlson et al., 2006).

Additionally, ASR glucosylates **more efficiently** melibiose, raffinose (three times more), methyl α-L-rhamnopyranoside and cellobiose than the dextransucrase DSR-S from *Ln. mesenteroides* NRRL B-512F (Argüello Morales et al., 2001; Champion et al., 2009; Côté et al., 2003, 2009) and usually produces oligosaccharide **mixtures** of lower DP than DSR-S because of its more promiscuous linkage specificity (for example, from panose, two products can arise: α-1,6-panose and α-1,3-panose).

III.3. Other potential application of ASR

The ASR produced by the NRRL B-23192 mutant strain was used for the glucosylation of the flavonoid **luteolin**. The conversion of 11 mM luteolin reached only 8% but both mono-, di- and tri-glucosides were obtained whereas only monoglucosides were synthesized with *Ln. mesenteroides* NRRL B-512F dextransucrase (Bertrand et al., 2006). The same enzyme was also more efficient than *Ln. mesenteroides* NRRL B-512F and *Ln. citreum* B-1299 dextransucrases to glucosylate α-**butylglucopyranoside** with 81% conversion compared to 70% and 68% conversion for the other enzymes. Moreover, the ASR was the only glucansucrase to be able to recognize the α-**octylglucopyranoside**, with 66% conversion rate (Richard et al., 2003).

The ASR from *Ln. citreum* SK24.002 was also used to perform the glucosylation of **stevioside** (13-O-β-D-sophorosyl-19-O-β-D-glucosyl-steviol), a promising non-cariogenic and low-calorigenic sweetener, with 43.7% conversion degree. Stevioside-glucosides could present lower bitter taste and better aftertaste similarly to gentiobiose glucosides (Musa et al., 2014).

ASR C-APY-del glucosylation of **naringenin**, a flavonoid, reached 27.1% of glucosylation efficiency, whereas DSR-S vardelΔ4N was not able to use it as acceptor. ASR was thus included in the patent US20170107242A1 that describes methods and uses of novel o-α-glucosylated flavonoids (Morel et al., 2017). Similarly, ASR C-APY-del was efficient to recognize 2-(hydoxy)ethyl methacrylate (**HEMA**)

and N-(hydroxyl)methylacrylamide (**NHAM**) as acceptors and produced glycosylated synthons that were used to produce **glyco-co-polymers** (André et al., 2018).

Another new and recent application was reported. ASR was used to elongate amylose chains of DP 30 and synthesize an oligoalternan, which was in turn elongated by a third enzyme the DSR-M dextransucrase; the use of ASR to link amylose to dextran enabled the production of the first a **tri-block polymer** uniquely using enzymatic technologies (Grimaud et al., 2018).

All these results demonstrate that ASR shows broad acceptor specificity and indicate that its active site should be rather large to accommodate such a **wide variety** of molecules.

PhD objectives

In conclusion of this literature review, the ASR is an old-known but not well-understood enzyme. Considering its multiple assets namely (i) its great **stability**, (ii) its particular mode of elongation that leads to **alternan** structure, (iii) the production of **HMM** alternan more soluble than dextran, offering different types of applications, (iv) its potential for **prebiotic** production and (v) its usage as **glucosylation** tool, the ASR deserves to be studied in much more details.

Consistently, **solving the enzyme 3D structure** should help for the understanding of the enzyme **linkage specificity** and allow unraveling the **mechanism** of **high and low molar mass alternan formation**. 3D structures acquisition could also in parallel provide valuable information on the ASR **stability.** These first objectives were at the heart of my PhD thesis to make the rational or semi-rational engineering of the enzyme possible, pave the way to further improvements of ASR and enlarge ASR utilization as **glucosylation tool**.

For that purposes, a multi-disciplinary approach, including enzyme structural and biochemical characterization, as well as mutagenesis studies (Figure 31) was undertaken to bring answers to the following questions:

- What are the **structural features** involved in the **specificity** of **ASR**, and in particular its ability to alternate α-1,3 and α-1,6 linkage during glucosyl polymerization?
- What are the **structural determinants** involved in **polymer elongation,** both in **ASR domain V** or in the **catalytic core**? Are those determinants specific to ASR, or common to those described in other GH70 enzymes?
- Could we identify some **structural determinants** explaining the higher **thermal stability** of ASR compared to the other GS of known 3D structure? If, yes, could we **improve** the **thermal stability** of another GS by rational design based on ASR features?
- Could we engineer the ASR, both by random evolution or rational engineering, to make it more stable and better understand the determinants of GS stability?

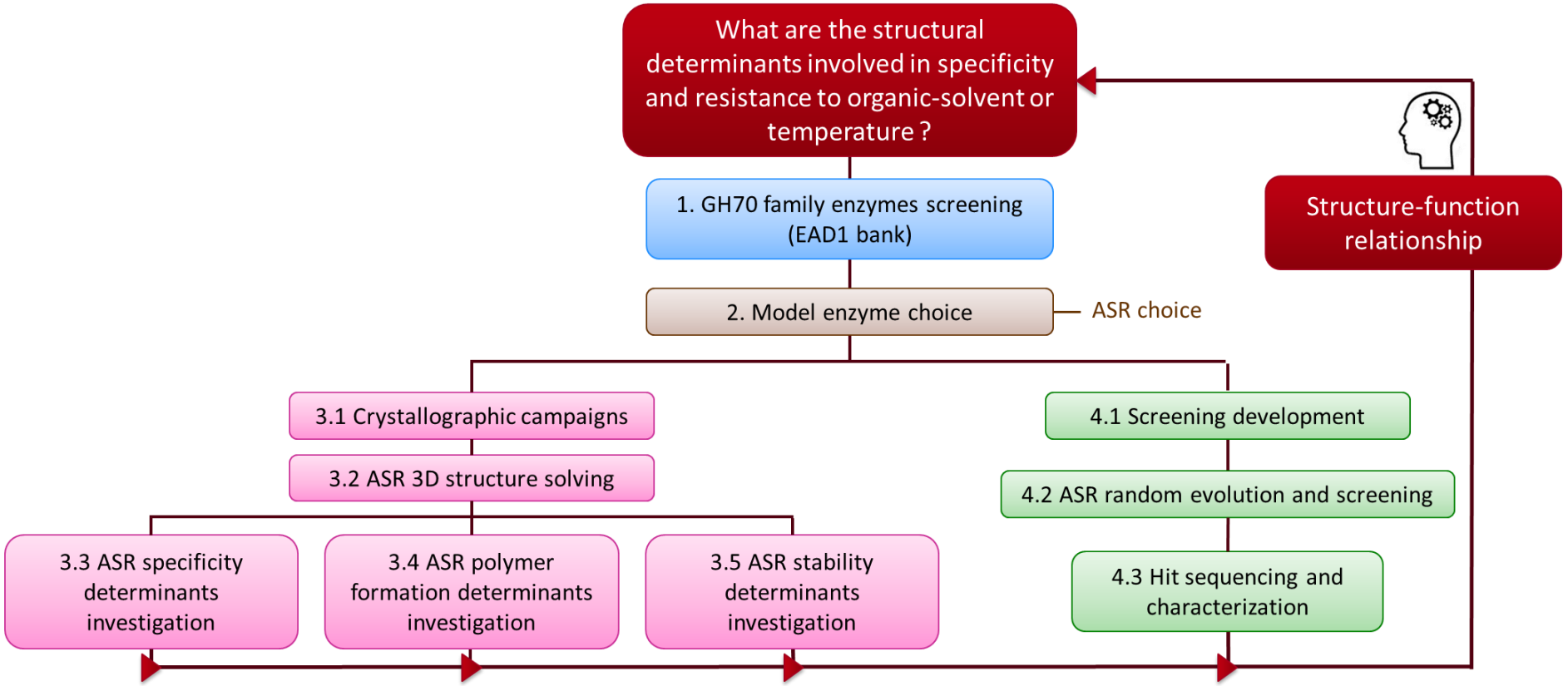


Figure 31: Workflow of the strategy used

Chapter II:

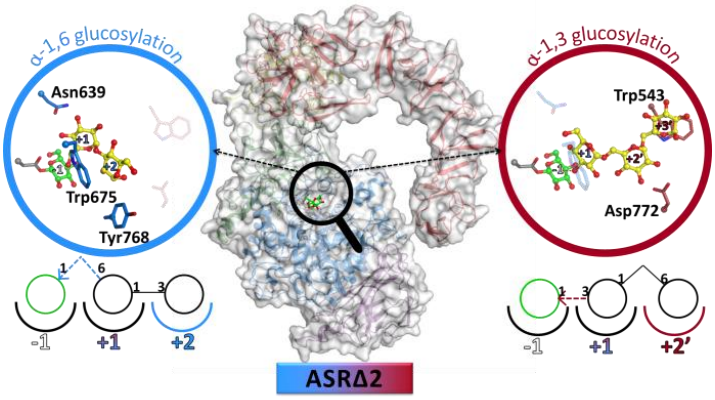
Deciphering an Undecided Enzyme: Investigations of the Structural Determinants Involved in the Linkage Specificity of Alternansucrase

Manon Molina, Claire Moulis, Nelly Monties, Sandra Pizzut-Serin, David Guieysse, Sandrine Morel, Gianluca Cioci*, Magali Remaud-Siméon*

Abstract

Understanding how polymerases catalyze the synthesis of biopolymers is a timely and important issue in generating controlled structures with well-defined properties. With this objective in mind, here we describe the 2.85Å crystal structure of a truncated version of alternansucrase (ASR) from *L. citreum* NRRL B-1355. Indeed, ASR is a striking example of α-transglucosylase among GH70 glucansucrases, capable of catalyzing high and low molar mass alternan, an α-glucan comprising alternating α-1,3 and α-1,6 linkages in its linear chain. The 3D structure sheds light on the various features involved in enzyme stability. Moreover, docking studies and biochemical characterizations of 17 single mutants and two double mutants enable the key determinants of α-1,6 or α-1,3 linkage specificity to be located and establish the structural basis of alternance. ASR displays two different acceptor subsites in the prolongation of its subsites -1 and +1. The first one is defined by Trp675, a residue of subsite +2, and orients acceptor binding exclusively toward α-1,6 linkage synthesis. The second binding site comprises Asp772 and Trp543, two residues defining the +2' and +3' subsites respectively, which are critical for α-1,3 linkage formation. It is proposed that the interplay between these two acceptor sites controls alternance. These results add to the toolbox of enzymes for the production of tailor-made polysaccharides with controlled structures.

Keywords: Glucansucrase, alternansucrase, alternan, crystal structure, GH70, α-(1→3)/α-(1→6) linkage specificity and alternance.



ACS Catal., in press (DOI 10.1021/acscatal.8b04510)

Introduction

The α-glucans constitute an important class of polymers, among which those produced using glucansucrases (GSs) from GH70 family being particularly attractive for applications in medicine, nutrition, cosmetics and materials science (Badel et al., 2011; Monsan et al., 2001; Naessens et al., 2005; Suresh Kumar et al., 2007). They are synthesized directly from sucrose, an abundant substrate available in a highly pure form. A broad variety of molecular structures can be easily obtained depending on the enzymes involved in their synthesis. Indeed, the size, type and arrangement of α-osidic linkages as well as the degree of branching can vary considerably from one polymer to another. All these factors define the physicochemical, biological, and mechanical properties of each specimen, and therefore its range of application. Improving the understanding of the molecular features at the origin of glucansucrase specificity and diversity is therefore essential for further development of α-glucans with tightly controlled structures.

In an impressive study aiming to characterize the polysaccharides produced by almost one hundred different lactic acid bacteria, Jeanes *et al.* were the first to report the presence of an α-glucan containing alternating α-1,3 and α-1,6 osidic linkages in the culture supernatant of several strains of *Leuconostoc* sp., such as *L. mesenteroides* NRRL B-1355 (recently reclassified as *L. citreum* (Bounaix et al., 2010)), NRRL B-1501 and NRRL B-1498 (Jeanes et al., 1954). Different analytical techniques (periodate oxidation, methylation, acetolysis, Smith reaction, NMR and enzymatic digestion) confirmed the presence of around 40% α-1,3 linkages and 60% α-1,6 linkages in the polymer (as determined by NMR on purified alternan), with the occurrence of alternating α-1,3 and α-1,6 linkages in the polymer chain, as well as some consecutive α-1,6 linkages and around 10% branching linkages (Côté and Robyt, 1982a; Dertli et al., 2018; Goldstein and Whelan, 1962; Hare et al., 1978; Jeanes et al., 1954; Joucla et al., 2006; Leathers et al., 2009; Misaki et al., 1980; Moulis et al., 2006; Seymour et al., 1976, 1977, 1979b, 1979a; Torii and Sakakibara, 1974). The polymer was named alternan. The enzymatic activity responsible for its synthesis was isolated and referred to as alternansucrase (ASR, EC 2.4.1.140) (Côté and Robyt, 1982a). With an optimum temperature of 45°C, ASR is one of the most stable enzymes in the GH70 family, which enabled the enzyme to be purified using a thermal treatment (López-Munguía et al., 1990; López-Munguía et al., 1993). Further, the *asr* gene from *L. citreum* NRRL B-1355 was expressed in *E. coli* (Argüello-Morales et al., 2000b). The recombinant protein was shown to produce, from sucrose, a bi-modal population of glucan, comprising a High Molar Mass (HMM) fraction of 1,700,000 $g.mol^{-1}$ as estimated by Sixe Exclusion Chromatography and a Low Molar Mass (LMM) fraction of 1,300 $g.mol^{-1}$ (Joucla et al., 2006).

Sequences sharing more than 97% identity with *L. citreum* NRRL B-1355 ASR sequence are found at present in the genome of other *L. citreum* strains (LBAE C11, KM20, EFEL 2700, ABK-1), suggesting that ASR is widespread in these species (BLASTp analysis, data not shown). To date, only one other alternansucrase from *L. citreum* ABK-1 has been produced recombinantly and characterized (Wangpaiboon et al., 2018).

Alternan polymer is more water soluble and less viscous than dextrans, making it a good substitute for Arabic gum (Côté, 1992; Leathers et al., 2009). Nanoparticles and films were recently obtained with alternan produced by *L. citreum* ABK-1 alternansucrase, thereby opening up other potential applications in nanotechnology (Wangpaiboon et al., 2018). Moreover, ASR also catalyzes transglucosylation from sucrose to many different types of sugar acceptors – methyl-α-D-glucoside, maltose, maltodextrin, maltitol, isomaltooligosaccharides (López-Munguía et al., 1993), cellobiose (Argüello Morales et al., 2001), melibiose, raffinose, gentiobiose and lactose – and produces glucosylated products showing interesting prebiotic properties (Côté et al., 2003; Holt et al., 2005; Sanz et al., 2005a). Stevioside-glucosides showing promising non-cariogenic and low-calorie properties were also synthesized (Musa et al., 2014). Finally, ASR was also recently used to elongate amylose chains of DP 30 and create a linker between amylose and linear dextran enabling the production of a tri-block polymer using enzymatic technologies alone (Grimaud et al., 2018).

Sequence analysis revealed that ASR belongs to the GH70 family. The enzyme was predicted to adopt the same fold as the other glucansucrases and the same α-retaining mechanism involving the contribution of Asp635, Glu673 and Asp767. It was suggested that these amino acids play the role of, respectively, the nucleophile, acid/base catalyst and transition state stabilizer (TSS) implicated in the formation of the β-D-glucosyl-enzyme intermediate (Argüello-Morales et al., 2000b). Unusual repeated sequences named “APY repeats” were identified at the C-terminal end of the protein (Janeček et al., 2000). Their deletion in the mutant ASR C-APY-del (covering amino acids Met1 to Gly1425) did not impact the product profile (Joucla et al., 2006) but significantly increased (five-fold) the enzyme half-life time at 50°C compared to the entire ASR (Joucla, 2003). A kinetic study conducted with ASR C-APY-del showed that both HMM and LMM alternan populations are formed in the early stage of the reaction, suggesting that ASR follows a semi-processive mechanism of polymerization (Moulis et al., 2006). In addition, a mutagenesis study of ASR C-APY-del highlighted the importance of the triplet ^{768}YDA^{770} (downstream of the putative TSS stabilizer Asp767 and found solely in ASR) in linkage specificity (Moulis et al., 2006). Replacement of this motif with the triplet SEV, which is characteristic of glucansucrase and highly specific to α-1,6 linkage synthesis, resulted in a variant synthesizing both oligoalternans and oligodextrans up to a degree of polymerization 4 (DP4) through successive glucosylations of the maltose acceptor. However, contrary to the wild type ASR,

the mutant was unable to further elongate the DP4 oligoalternan, which resulted in its accumulation. Only the oligodextrans were further elongated via α-1,6 linkage synthesis.

Despite the multiple attributes of alternansucrase from *L. citreum* NRRL B-1355 – namely (i) its thermal stability, (ii) its potential for the production of prebiotic compounds, glycoconjugates and new biomaterials and (iii) the very unique specificity of this glucansucrase for alternating α-1,3 and α-1,6 linkage synthesis – information related to the structural determinants impacting linkage specificity and stability remains scarce. Here, we explore further the structure-function relationships of this enzyme and disclose the first X-ray three-dimensional structure of a GH70 alternansucrase. With 1,278 residues, the new free 3D-structure of ASRΔ2 was obtained at 2.8Å resolution. It is the largest GH70 enzyme structure solved so far. Structural analysis combined with mutagenesis studies and biochemical characterization enabled the identification of different features and amino acids exerting a critical role in stability, specificity and ratio of LMM to HMM polymers.

Results

Design and characterization of truncated mutants

As revealed by sequence alignment, ASR from *L. citreum* NRRL B-1355 is predicted to adopt the same fold as the other GSs comprising five domains A, B, C, IV and V (Figure 1). To overcome the difficulties of recombinant ASR (Met1-Ala2057) and ASR-C-APY-del (Met1-Gly1425) crystallization, we constructed, produced and purified several truncated forms: ASRΔ2, ASRΔ3, ASRΔ4 and ASRΔ5, deleted, respectively, of the signal peptide (Met1-His38), one predicted disordered region (Met1-Asp78), two predicted disordered regions (Met1-Pro152) and the entire domain V (Met1-Ser398 and Gln1305-Gly1425) (Figure 1).

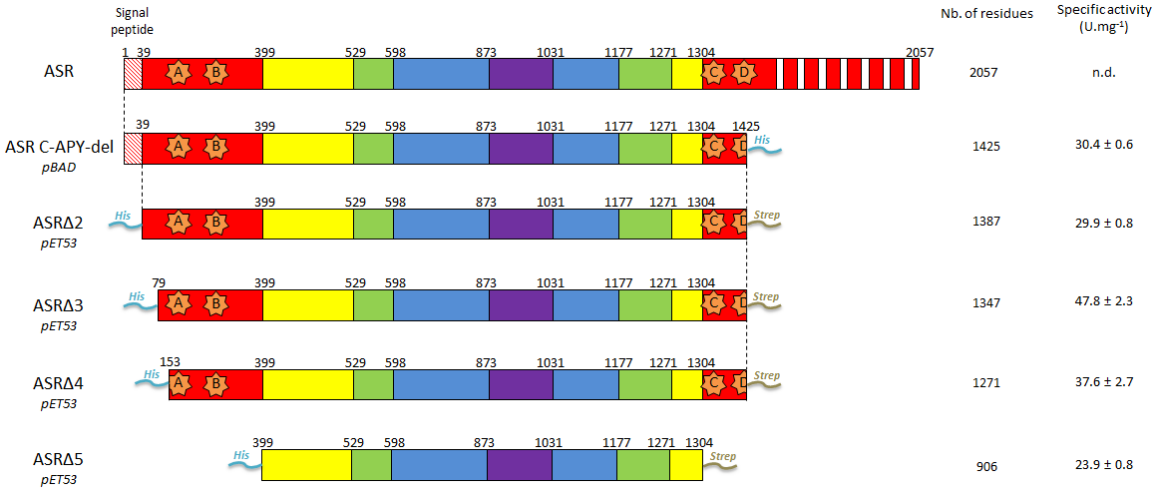


Figure 1: Schematic representation of the domain organization of ASR truncated mutants.
The upper numbering corresponds to residues delimiting the different domains (red: domain V, yellow: domain IV, green: domain B, blue: domain A, purple: domain C) predicted by alignment with DSR-M **(Claverie et al., 2017)**. White stripes represent APY repeats **(Joucla et al., 2006)**. Stars A (147-221), B (234-304), C (1324-1398) and D (1399-1425) represent the putative glucan binding pockets. His: His-tag, Strep: Strep-tag. Specific activities were determined from 292 mM sucrose in 50 mM sodium acetate buffer pH 5.75 at 30°C and with 0.05 mg.mL^{-1} of pure enzyme.

The deletions did not severely compromise the specific activity of the truncated proteins. ^{1}H NMR and ^{13}C NMR analysis of the purified polymer synthesized by ASRΔ2 revealed the presence of 39.5% α-1,3 and 60.5% α-1,6 linkages (Figure 2A). ^{1}H NMR of the non-purified glucans produced by the others variants were all similar (Figure S2). HPSEC analyses further showed that similar amounts of LMM and HMM polymers were produced with all the truncated mutants except ASRΔ5 (Figure 2B). Indeed, the deletion of domain V strongly affects the ability to synthesize HMM alternan, as only 4.5 ± 0.2% of the glucosyl units from sucrose are incorporated into HMM polymer compared with 32.4 ±

0.8% in the case of ASRΔ2. The denaturation patterns of the variants, monitored using Differential Scanning Fluorimetry (DSF), revealed two distinct transition temperatures for all the variants, one around 37°C and the other one around 55.0°C. Such profiles are often encountered for multi-domain enzymes. The melting temperatures (T_m) of ASRΔ2, Δ3 and Δ4 are in the same range. The two melting temperatures of ASRΔ5 are slightly lower, suggesting that domain V of ASR might contribute to overall enzyme stability (Figure S3).

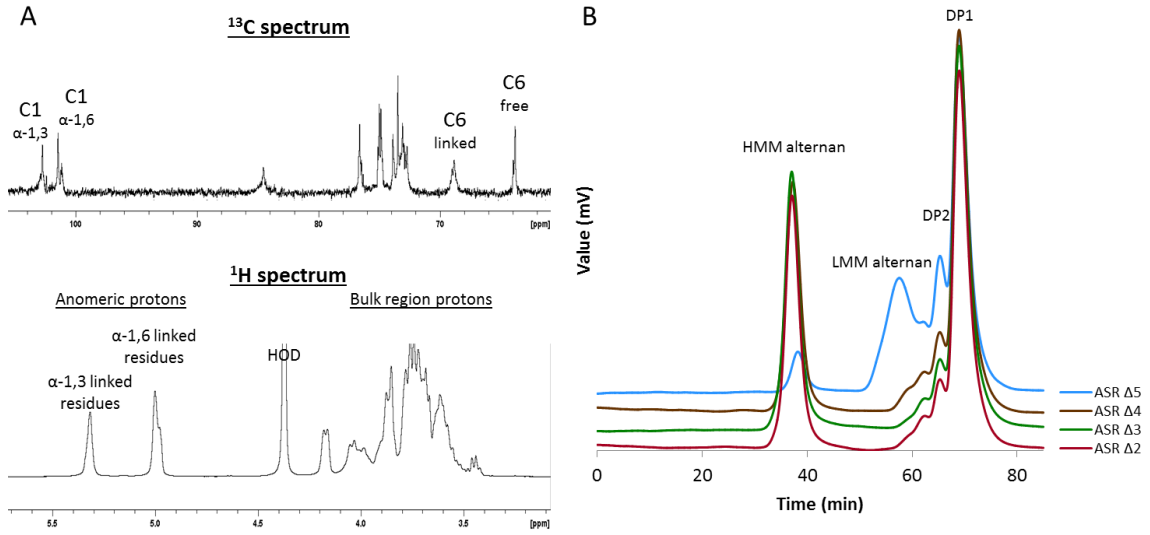


Figure 2: (A) ^{13}C and ^{1}H NMR spectra of alternan produced by ASRΔ2.
13C spectrum: the resonances at ~102 ppm and ~101 ppm are attributed to the anomeric carbons involved in α-1,3 or α-1,6 linkages respectively. At ~84.5 ppm, the signal is assigned to C-3 of a C3-substituted glucosyl unit. The resonance of the various C-2, C-3, C-4 and C-5 of the glucosyl units arises between 75 and 70 ppm. At ~68 and ~63 ppm, the signals correspond to the C-6 of the C-6 substituted glucosyl and to a free C-6 respectively. 1H spectrum: the signals at 5.32 ppm and 5 ppm are assigned to the anomeric proton of glucosyl residues involved in α-(1→3), 39.5%, or α-(1→6)-linkage, 60.5%, respectively. This result is in accordance with the values reported in the literature for alternan (Dertli et al., 2018; Joucla et al., 2006; Seymour et al., 1979b)
(B) HPSEC chromatograms of truncated mutants.
Reaction from sucrose at 30°C with 1 U.mL^{-1} of pure enzyme and sodium acetate buffer 50 mM pH 5.75.

Oligosaccharide synthesis from maltose acceptor reaction

An acceptor reaction with maltose, the best known acceptor for ASR, was carried out to ensure that the truncated mutants behaved in the same way as the full-length enzyme. First, acceptor reaction product formation was monitored using ASRΔ2 (Figure 3). Our results are in alignment with those of Côté *et al.* obtained using the native ASR (Côté et al., 2008). To summarize, maltose undergoes α-1,6 glucosylation alone to give the oligodextran of DP3 (OD3 panose: α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc), which can be further elongated at either the O6 or O3 positions of the non-reducing unit to give the structures OD4 (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc) and OA4 (α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc), respectively. The oligosaccharide OA4 was quick to

appear at the beginning of the reaction (1 min) and accumulated at a much higher level than OD4, indicating that panose is preferentially elongated with an α-1,3 linkage. Note that a very small peak of OD5 originating from OD4 only appeared towards the end of the reaction (~30 min) and was not elongated further (no OD6 was found). As soon as the OA4 starts to accumulate, it is efficiently converted to OA5 (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc). The OA5 itself can act as an acceptor for the formation of two OA6s (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc and α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc). Other products eluted at retention times close to those of OA6 were identified. They probably correspond to minor products of DP greater than DP6 and could not be determined using our LC/MS apparatus (Figure S4). All the truncated mutants behaved exactly like ASRΔ2 (Figure S5).

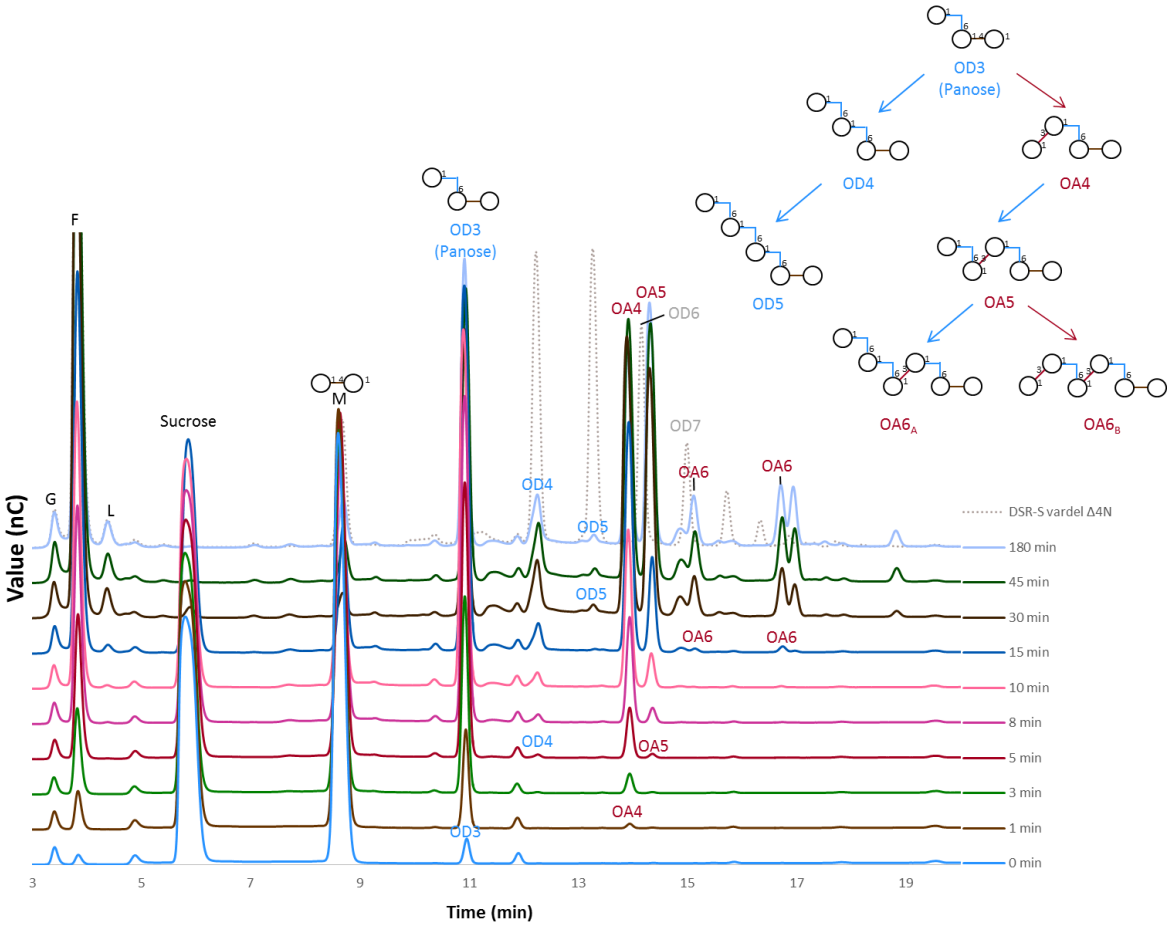


Figure 3: HPAEC monitoring of acceptor reaction of ASRΔ2 on maltose and model of product formation from maltose for DP 3 to 6, adapted with permission from reference (Côté and Sheng, 2006). Copyright 2006, Elsevier.

Reaction from 292 mM of sucrose and 146 mM of maltose with 50 mM sodium acetate buffer pH 5.75 and 1 $U.mL^{-1}$ of pure enzyme. G: Glucose, F: Fructose, L: Leucrose, M: Maltose, OD: Oligodextran, OA: Oligo-alternan.

Overall 3D structure of ASRΔ2

After extensive crystallization trials, the 3D structure of the unliganded ASRΔ2 was solved by molecular replacement using the structure of DSR-M as a template (Claverie et al., 2017) and refined at 2.8Å resolution. We did not manage to reconstruct the entire N-terminal part of the enzyme due to the poor quality of the density map in this region. We subjected the crystals to Edman sequencing, which revealed protein degradation during the crystallization process. The crystallized fragment started at Thr145 instead of Ala38, thus being an intermediate between the ASRΔ3 and ASRΔ4 constructions. The final model corresponds to the largest structure solved so far in the GH70 family and includes residues Ser147 to Ser1423 for chain A, and residues Gln248 to Ser1423 for chain B, for which the electron density was less well resolved in the N-terminal region.

Like all the other glucansucrases of solved 3D structure, ASRΔ2 comprises five distinct domains (A, B, C, IV and V) in which domains A, B, IV and V are made up by sequence fragments on either side of domain C (Figure 1, Figure 4A). This domain (Ser873-Gln1030) is composed of ten β-sheets forming a Greek key motif (Ito et al., 2011). Compared to the other glucansucrase structures, it also contains two insertions (^{882}SSGKDLKDGE^{890} and ^{913}QDNS^{916}) and one additional β-hairpin (Thr991 to Glu1005) (Figure S6). Overall, the β-strands are longer and ASR domain C displays a higher number of ionic, π-π stacking interactions and hydrophobic residues (Val, Leu, Ile, Phe) than its counterparts in the other GS structures (DSR-M, GTF-SI, GTF180, GTFA and GBD-CD2).

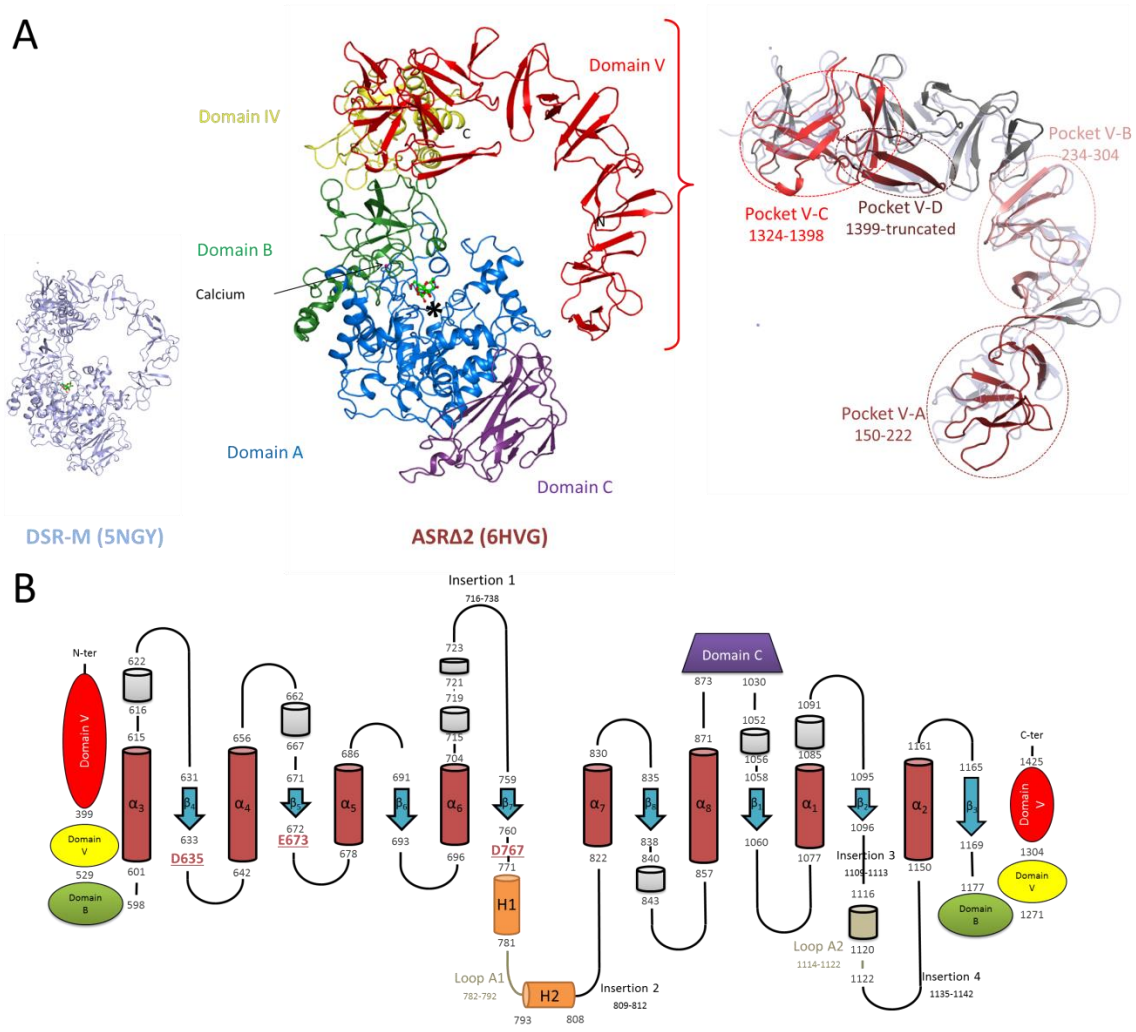


Figure 4: (A) Global view of the domain organization of ASRΔ2 (PDB ID: 6HVG, chain A) versus DSR-M (PDB ID: 5NGY, chain A) and zoom on ASRΔ2 domain V with the putative sugar binding pockets (circled). Star: active site. Red: domain V (147-398+1304-1424), yellow: domain IV (399-528+1271-1303), green: domain B (529-597+1177-1270), blue: domain A (1031-1176+598-872), purple: domain C (873-1030). Sucrose (green) was manually docked from 3HZ3 complex.

(B) Schematic representation of the structural elements of the catalytic barrel of ASR

Cylinder: α-helix, arrow: β-sheet. Pink and cyan elements belong to the catalytic barrel. Orange: helices H1/H2. Residue underlined: catalytic residues

ASRΔ2 adopts an overall "horseshoe shape" in which domain V (Ser147-Ser398 and Ala1304-Arg1424) is bending toward the catalytic domain as previously observed in DSR-M glucansucrase (Claverie et al., 2017) (Figure 4A). This is not surprising as the two enzymes share 48% and 87% of their identity between the N-terminal and C-terminal ends respectively of their domain V. This typical fold is due to the position of domain V relative to domain IV (Gln399-Asn528 and Gly1271-Asp1303) which is thought to act as a hinge between domain V and domains A/B, as previously proposed for the glucansucrases GTF180 and DSR-M (Claverie et al., 2017; Pijning et al., 2014). Notably, the ASRΔ2 domain V shows repeated units composed of three consecutive β-hairpins with a hydrophobic core

and many π-π stacking interactions reinforcing their packing (Figure S7) (Brison et al., 2016; Vujičić-Žagar et al., 2010). These super-secondary elements allow to define four putative sugar binding pockets (named V-A, V-B, V-C and V-D) in the global ASR sequence by sequence alignments and homology to the sugar binding pockets identified in DSR-E and DSR-M (Brison et al., 2016; Claverie et al., 2017) (Figure 4A).

ASRΔ2 catalytic domain

The catalytic domain A of ASRΔ2 is formed by a $(\beta/\alpha)_8$ barrel common to all enzymes of the GH-H clan, which also comprises the GH13 and GH77 families. Compared with the $(\beta/\alpha)_8$ barrel of the GH13 family, the barrel of ASRΔ2, and of all other glucansucrases, underwent a circular permutation occurring between strand β3 and helix α3 (Figure 4B). Consequently, the conserved and signature motifs I to IV are not placed in the same order along the sequence in the GH70 glucansucrases compared to those of GH13 family enzymes, motif I being downstream of motif II to IV in the sequences in the GH70 glucansucrases (Figure 5) (MacGregor et al., 1996). A manual docking of sucrose in ASRΔ2 was performed using the GTF180:sucrose complex (PDB ID: 3HZ3). The residues defining subsites -1 and +1 of ASRΔ2 and GTF180 (according to the subsite nomenclature proposed by Davies *et al*. (Davies et al., 1997)), and in interactions with the glucosyl and the fructosyl ring of sucrose respectively were similar. They were well aligned with a RMSD of 0.33 Å and confirmed that the catalytic residues of ASRΔ2 are Asp635, Glu673 and Asp767 (Figure S8) (Vujičić-Žagar et al., 2010). Compared to DSR-M and GTF180, domain A of ASRΔ2 (Asn598-Val872 and Asp1031-Tyr1176) displays several insertions in the loops emerging between i) α-helix 6 and β-strand 7 (insertion 1: Trp716-Arg738), ii) β-strand 7 and α-helix 7 (insertion 2: ^{809}NPSG^{812}), and iii) β-strand 2 and α-helix 2 (insertion 3: ^{1109}NYGGM^{1113}and 4: ^{1135}NKADGNPN^{1142}) (Figure 4B, Figure 6A).

A calcium ion is present at the interface between domain A and B in interaction with Glu589, Asp595 (motif V), Asp1173 (motif I) and Asn639 (motif II) (Figure 5), similar to the situation already described for GTF180 (Vujičić-Žagar et al., 2010). The effect of calcium on ASRΔ2 was investigated by comparing the specific activities obtained without calcium ion addition, and in the presence of 3.4 mM calcium chloride or with 5 mM EDTA. The values were 29.9 ± 1.0 U.mg^{-1}, 33.6 ± 2.6 U.mg^{-1} and 29.5 ± 1.0 U.mg^{-1} respectively, indicating that calcium ions do not significantly activate the enzyme as previously reported for the ASR of *L. citreum* ABK-1, for which a 1.2-fold increase in activity was observed in the presence of 10 mM calcium chloride at 40°C pH 5 (Wangpaiboon et al., 2018). However, the absence of calcium influenced the enzyme denaturation profile during DSF experiments and led to an almost complete elimination of the higher transition at 55°C (Figure S3).

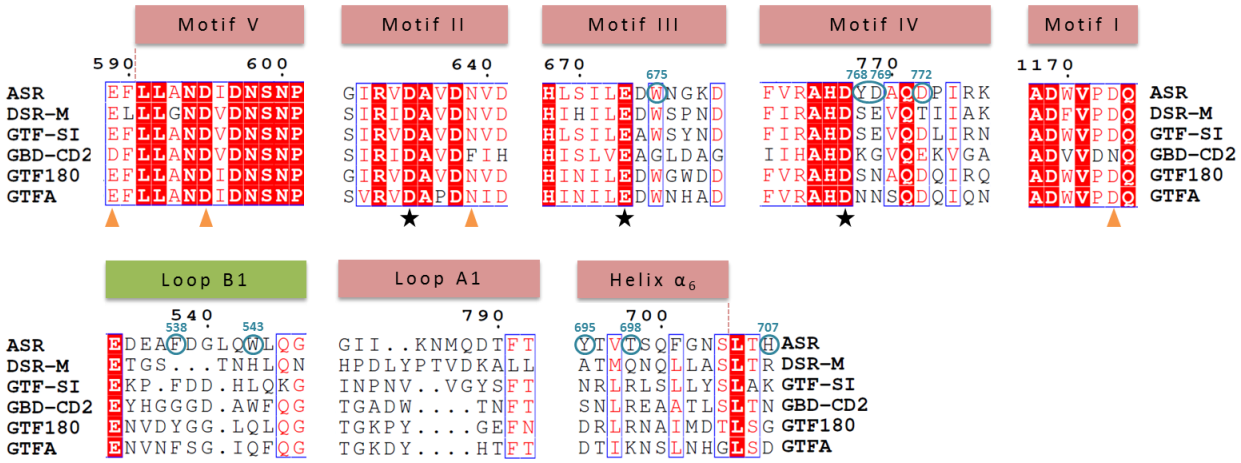


Figure 5: Partial alignment of sequences from available sucrose-active GH70 structures.
Black star: catalytic residues. Orange triangle: calcium binding site (Glu589, Asp595, Asn639 and Asp1173). Cyan circled residues: residues targeted in this study. GenBank accession numbers: CAB65910.2 (ASR), CDX66895.1 (DSR-M), AAN58706.1 (GTF-SI), CDX66820.1 (GBD-CD2), AAU08001.1 (GTF180), AAU08015.1 (GTFA). Alignment created with ENDscript 2 **(Robert and Gouet, 2014)**. A structure-based sequence alignment is provided in Figure S9.

All the secondary structural elements of the barrel superpose well with those of the other GH70 structures solved so far, except the α-helix H1 (Gln771-His781) and α-helix H2 (Phe793-Glu808) between β-strand 7 and α-helix 7, for which we observed the largest deviation. This could be due to rearrangements triggered by the presence or absence of ligand in the active site, as shown for DSR-M (Claverie et al., 2017), and/or to the flexibility of this region. Indeed, the H1/H2 (Gln771-Glu808) helices present a relatively high B-factor with an average of 66.4 Å² compared to 56.4 Å² for domain A.

Focusing on the catalytic core to investigate linkage specificity and HMM polymer formation

Two different zones surrounding the sucrose molecule in the model of ASRΔ2:Sucrose complex drew our attention (Figure 6B). The first, zone 1, comprises **Trp675**, **Tyr695**, **Thr698, His707, Tyr768** and **Asp769.** Trp675 is a conserved residue (in motif III), located between subsites +1 and +2 (Figure 6B, Figure S8). The other four residues are in the proximity of Trp675. In particular, Tyr768 and Asp769 form part of the triplet ^{768}YDA^{770} (in motif IV following the TSS), which is unique to alternansucrases and is replaced by the triplet SEV in most dextransucrases (DSR-S) and mutansucrases (GTF-I) or by the triplet NNS in reuteransucrases (GTFA). The second region of interest, zone 2, is located on the opposite side of zone 1 (Figure 6B) and comprises residues that could define a path to the sucrose binding site including **Trp543, Asp772,** and **Phe538.** The Trp543 residue is unique to ASRΔ2 and located in a small α-helix (Leu541-Gln545) of loop B1 (Glu534-Gly548). Behind this residue, Phe538 is

not strictly conserved in glucansucrases. Finally, Asp772 in α-helix H1 is facing Trp543 and points toward the active site. A total of nine positions were selected and mutated to examine their role in specific activity, linkage specificity, stability or ability to synthesize HMM polymer (Figure 6B). The results are shown in Table 1.

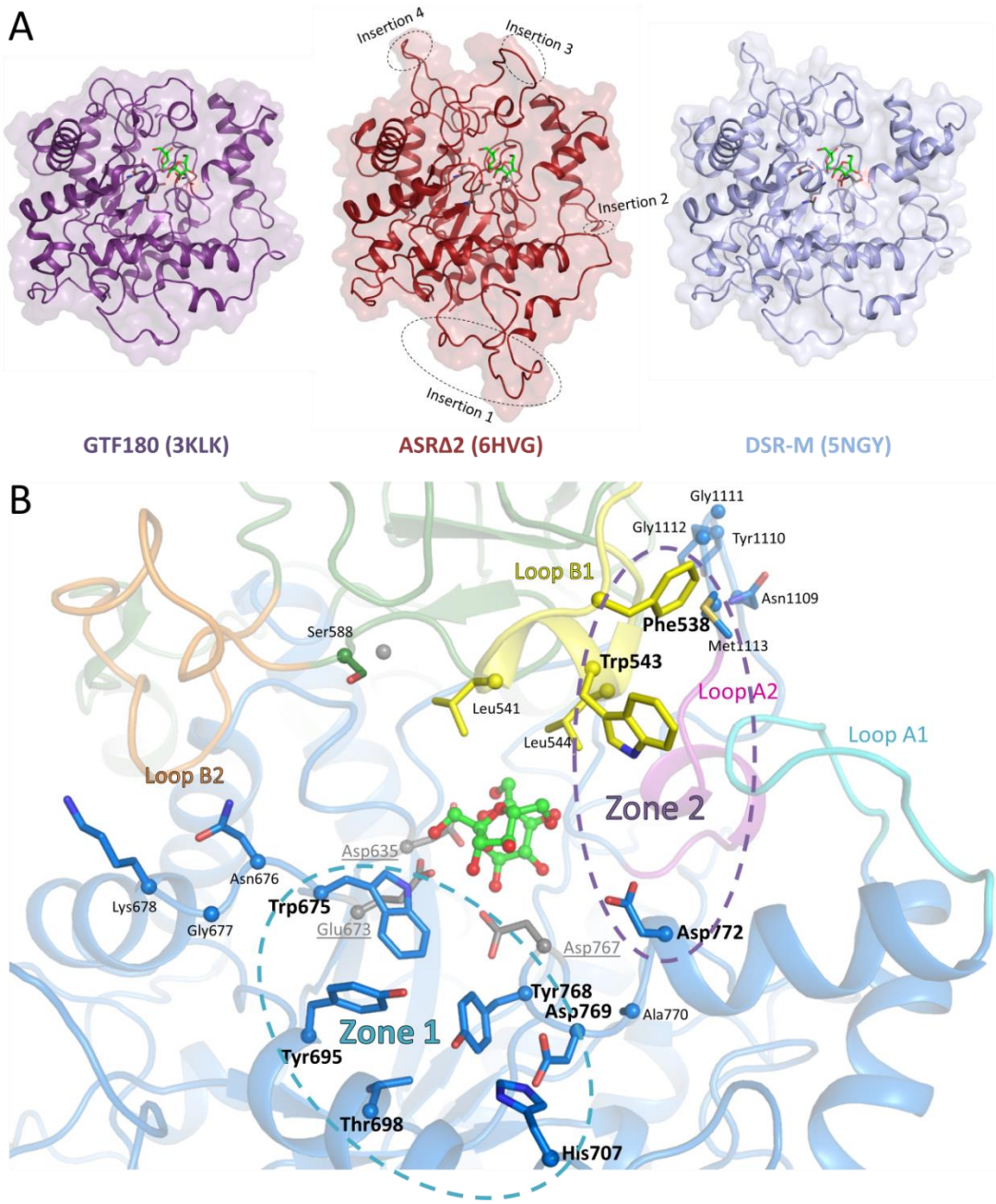


Figure 6: (A) View of the insertions in domain A of ASRΔ2 compared to GTF180 and DSR-M.
Only domain A is shown. Insertion 1: 716-738. Insertion 2: 809-812. Insertion 3: 1109-1113. Insertion 4: 1135-1142.
(B) View of the catalytic site and loop positioning in ASRΔ2.
Residues targeted are in bold. Catalytic residues are underlined in gray. Loop A1: 782-792, loop A2: 1114-1122, loop B1: 534-546, loop B2: 570-587. Navy blue: domain A. Forest: domain B. Gray sphere: calcium. Green: sucrose (docked from 3HZ3)

Table 1: Biochemical data on monomutants from the catalytic site.
Reaction from sucrose only at 30°C with 1 U.mL^{-1} of pure enzyme and sodium acetate buffer 50 mM pH 5.75 Stars: residues specific to alternansucrases. Specific activity of ASRΔ2: 29.9 ± 1.0 U.mg^{-1}. Four different wild types ASRΔ2 were characterized. Specific activity was determined in triplicate for the same sample. T_m was determined by DSF.

	Localization	Residual specific activity (%)	ΔTm with the wild type enzyme (°C)	α-1,3 linkages (%, NMR)*	α-1,6 linkages (%, NMR)*	HMM polymer (%,HPSEC area)	Hydrolysis (%)
Wild type							
ASRΔ2 WT		100 ± 3.3	0	35	65	32.4 ± 0.8	4.4 ± 0.5
Zone 1							
ASRΔ2 W675A	Motif III 2nd residue downstream of the acid/base	10.0 ± 0.77	-3.3	31	69	5.7 ± 0.3	44.9
ASRΔ2 W675F		32.4 ± 0.54	-16.8	39	61	31.0 ± 0.2	4.8
ASRΔ2 W675Y		35.1 ± 1.2	-1.8	28	72	8.6 ± 0.2	6.1
ASRΔ2 W675H		2.6 ± 0.06	-2.6	39	61	11.3 ± 0.2	12.3
ASRΔ2 H707*A	Helix α_6	120.4 ± 9.0	-0.1	35	65	34.1 ± 0.3	4.3
ASRΔ2 T698*A		85.3 ± 0.67	-2.1	36	64	34.5 ± 0.2	4.4
ASRΔ2 Y695*A		61.9 ± 2.5	0.8	33	67	27.3 ± 0.1	5.2
ASRΔ2 Y768*A	Motif IV "YDA" sequence downstream of the TSS	54.2 ± 2.9	-1.9	29	71	21.0 ± 0.1	6.5
ASRΔ2 Y768*F		110.4 ± 2.5	0.4	35	65	33.6 ± 0.2	6.1
ASRΔ2 Y768*W		52.8 ± 1.6	-0.8	29	71	11.9 ± 0.2	5.2
ASRΔ2 D769*A		110.4 ± 2.0	-0.8	29	71	26.0 ± 0.2	5.8
ASRΔ2 Y695*A+ Y768*A		67.2 ± 3.0	-0.3	27	73	18.3 ± 0.04	6.1
Zone 2							
ASRΔ2 F538A	Loop B1	95.3 ± 5.4	-4.7	37	63	31.9 ± 0.2	4.5
ASRΔ2 W543*A		78.5 ± 7.7	1.3	24	76	14.8 ± 0.1	7.2
ASRΔ2 D772A	Helix H1	36.5 ± 2.5	0.4	5	95	3.2 ± 0.1	5.1
ASRΔ2 D772E		71.9 ± 5.7	-1.8	7	93	10.9 ± 0.1	5.4
ASRΔ2 D772Y		36.1 ± 1.7	0	6	94	3.9 ± 0.1	5.1
ASRΔ2 W543*A+D772A		63.9 ± 2.6	0	6	94	2.1 ± 0.04	5.4

*: NMR was performed on crude reaction medium.

Mutations in zone 1: impact on specificity and HMM polymer synthesis

Trp675 was swapped with Ala, Phe, Tyr or His residues. All the mutants were impacted in their specific activity, in particular the Ala and His mutants (10% and 2.6% residual activity). Except for Trp675Phe, all of them produced much lower amounts of HMM polymer and higher amounts of glucose, showing that the mutation favored the hydrolysis reaction (Figure 7A). Mutation Trp675Phe was highly detrimental to the enzyme's stability, as reflected by a ΔT_m of -17°C compared to the wild type. Overall, the linkage specificity was not affected by the mutations at this position, suggesting

that this zone does not contribute to α-1,3 glucosylation. Indeed, the Trp675Ala mutation did not disfavor the formation of OA4 (Figure S10).

Mutants Tyr695Ala, Thr698Ala and His707Ala synthesized equivalent amounts of HMM polymer and α-1,3 linkages to the wild type. Individually, these positions do not seem to be highly critical for polymer elongation and linkage specificity. Considering the tyrosine 768 of YDA motif, though, our results indicate that it is different. Replacing Tyr768 with a Phe results in an enzyme that acts similarly to the parental enzyme, with the main function of the Tyr residue probably being maintained. In contrast, the Tyr768Ala mutant as well as the Tyr695Ala-Tyr768Ala double mutant induce a slight increase in α-1,6 linkage synthesis (+6% and 8%, respectively), suggesting that this zone might have a subtle but specific role in the formation of the α-1,6 linkages, also observed in the increased formation of OD oligosaccharides (oligodextran serie) compared with the wild type enzyme (Figure 8). HMM polymer formation decreases from 33% for the wild type to 20% and 12% for the Tyr768Ala and Tyr768Trp mutants, respectively.

Mutations in zone 2: impact on specificity and HMM polymer synthesis

The most spectacular effect concerns the mutations introduced at position 772 of helix H1 and to a lesser extent at position 543 of loop B1 (Figure 7B). The specific activity values of mutants Asp772 Ala/Glu/Tyr decreased, representing 36.5, 71.9 and 36.1% of that of the wild type enzyme, respectively. They all produced much lower quantities of HMM polymer although the hydrolysis activity is almost unchanged compared to wild type ASRΔ2. The synthesis of α-1,6 linked oligomers is clearly favored and the ability to synthesize α-1,3 linkages is almost lost (5%). In comparison, mutant Trp543Ala followed the same trend, even if the reduction of α-1,3 linkage synthesis and the impact on HMM polymer formation is less dramatic. The Trp543Ala-Asp772Ala double mutant behaves like the Asp772 single mutant, suggesting that the most important residue for α-1,3 glucosylation is Asp772. Finally, mutant Phe538Ala acts similarly to the wild type. The HPAEC chromatograms of the products of the maltose acceptor reaction also shed similar light. All the Asp772 mutants produce more OD oligosaccharides than the wild type. To illustrate, in Figure 8 we provide the most striking chromatographic profiles of the Tyr768Ala mutant and the Asp772Ala-Trp543Ala double mutant. Remarkably, it can be seen that the double mutant is even able to produce a small amount of OD6.

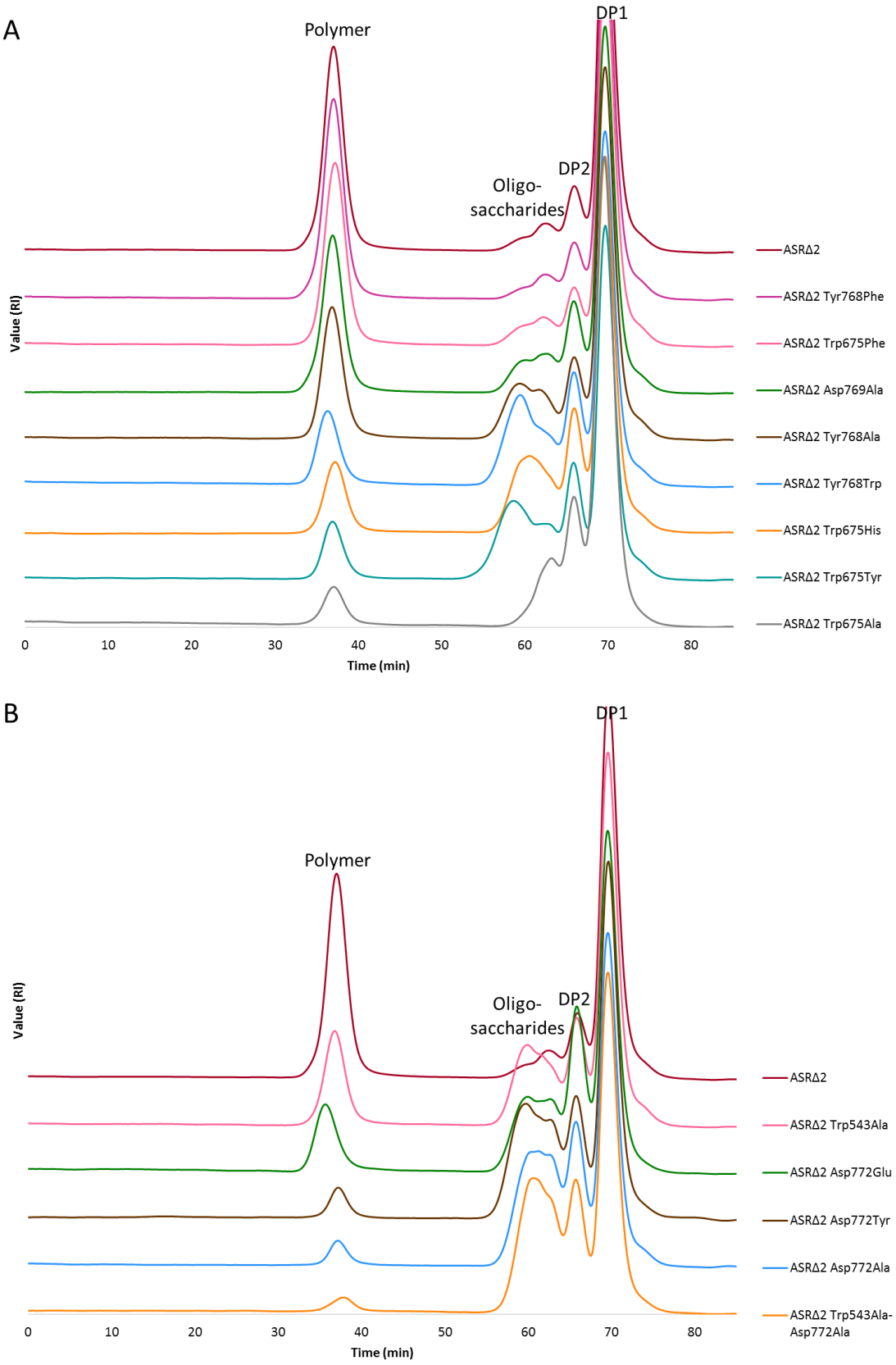


Figure 7: (A) HPSEC chromatogram of zone 1 mutants.
Mutants with slight effect are not depicted (Tyr695Ala, Tyr695Ala-Tyr768Ala, His707Ala, Thr698Ala)
(B) HPSEC chromatogram of zone 2 mutants.

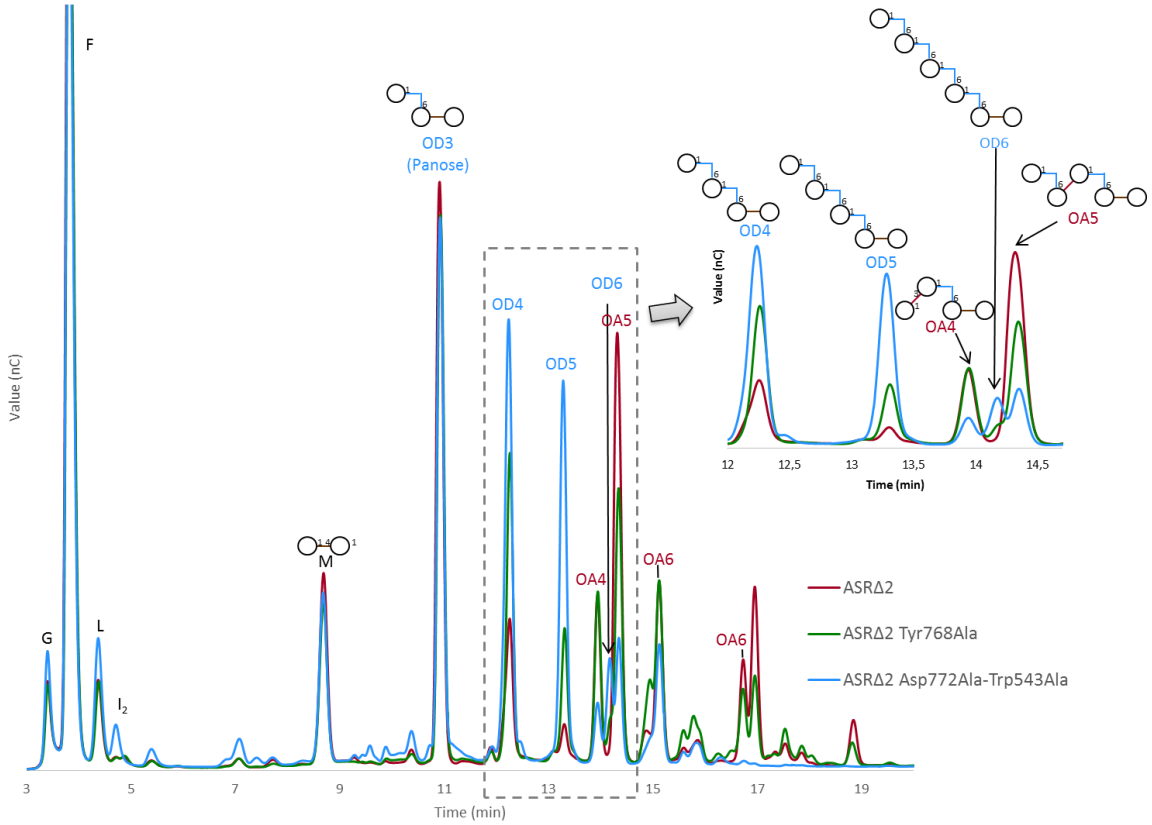


Figure 8: HPAEC chromatogram of ASRΔ2 compared to mutant Tyr768Ala and double mutant Asp772Ala-Trp543Ala. Reaction from 292 mM sucrose and 146 mM maltose with 50 mM sodium acetate buffer pH 5.75 and 1 U.mL^{-1} of pure enzyme. G: Glucose, F: Fructose, L: Leucrose, M: Maltose, OD: Oligodextran, OA: Oligo-alternan of different DP.

Discussion

Alternansucrase from *L. citreum* B-1355 has always been regarded as an intriguing glucansucrase, first because of its unusual ability to synthesize alternating α-1,3 and α-1,6 in the polymer chain and also because of its thermal stability. By solving the structure of this enzyme, we intended to bring new insights into the structural determinants influencing those highly unusual traits.

Contribution of different structural features influencing ASR stability

The X-ray structure of ASRΔ2 is typical of GH70 glucansucrases. Its closest counterpart, in terms of sequence and 3D structure similarities, is DSR-M, which also exhibits a horseshoe shape with a domain V in close proximity to the catalytic domain A/B. The stability of the entire edifice is clearly enhanced by calcium coordination that likely reinforces the interaction between domains A and B. Domain V is another contributor to stability as ASRΔ5 is the least stable enzyme of the truncated forms generated in our study. Although it is counterintuitive, the insertions (1 to 4) increasing the length of several loops on the enzyme surface could also contribute to enzyme stability. Indeed, it is generally thought that long loops in proteins decrease stability by increasing local flexibility and propensity to aggregate (Nagi and Regan, 1997). In ASRΔ2, the shape of these loops may prevent solvent penetration. Insertion 3 belonging to domain A (^{1109}NYGGM^{1113}) appears ideally located to protect some of the residues underneath that form the side wall of the catalytic pocket (Tyr1176, Leu592) (Figure 6B). Finally, the domain C of ASRΔ2 displays structural features (longer β-strands, higher number of hydrophobic residues, π-π stacking and ionic interactions) that could reinforce its role of "pedestal" at the bottom of the U-shaped fold and provide a contribution to enzyme stability. This warrants further investigation.

HMM alternan formation is controlled by catalytic core residues and by domain V

The Trp675 residue belonging to zone 1 is important for the formation of HMM alternan synthesis. It provides a critical stacking platform defining subsite +2 and accommodates sugar acceptors as previously suggested for Trp1065, the corresponding residue in GTF180. Indeed, ten mutants of Trp1065 were shown, in GTF180, to totally lose their ability to synthesize HMM polymer. Only the mutant Trp1065Phe still produced HMM polymer, but in very small amounts (just 2.2% compared with 16.5% for wild type GTF180) (Meng et al., 2017). In ASRΔ2, the Trp675Phe mutation was much less detrimental to HMM polymer formation than in GTF180. Overall, and even although the number of α-1,3 linkages obtained with mutant Trp675Tyr decreases slightly, mutations of Trp675 does not greatly impact enzyme linkage specificity. However, the factor with greatest influence on HMM

alternan formation seems to be the presence of domain V as the ASRΔ5 truncated form produces very small amounts of polymer. We can submit that the putative sugar binding pockets identified contribute to polymer elongation as proposed for DSR-M (Claverie et al., 2017). Future biochemical studies should confirm the functionalities of these pockets.

Understanding α-1,6 linkage specificity through the recognition of different DP2 acceptors

To gain a better understanding of the efficiency of maltose, nigerose and isomaltose as acceptors, here ranked in decreasing order of efficiency (Côté and Robyt, 1982b; López-Munguía et al., 1993), we docked each disaccharide in the ASR active site (Figure 9). The superposition of the GTF180-maltose complex on ASRΔ2 (Figure 9A) shows that the site formed by Asn639 and Trp675 is perfectly conserved and the maltose O6 hydroxyl (at the non-reducing end) is oriented towards the covalent intermediate, thus indicating that the mechanism already proposed for GTF180 (Vujičić-Žagar et al., 2010) to explain the α-1,6 specificity is very likely to operate for ASR. This is in alignment with previous observations (Côté and Robyt, 1982b; López-Munguía et al., 1993) and with our own observations that, starting from maltose, only panose can be formed (Figure 8). Interestingly, our docking simulation found a binding pose for nigerose (Figure 9B) that closely mimics the GTF180 maltose-complex, thus confirming that nigerose is also a good acceptor and can be efficiently elongated with an α-1,6 linkage from this subsite. In contrast, the docking failed to identify a similar binding pose for isomaltose (IM2) in this subsite. If we assume that the stacking interaction with Trp675 is a requisite for acceptor binding, the lack of a binding pose may be explained by the presence of the bulky Tyr768 next to the Trp675, which could lower the affinity of this site for IM2. This hypothesis is confirmed by the Tyr768Ala mutant, which indeed produces higher amounts of oligodextrans from the maltose acceptor reaction (Figure 8), suggesting that the Tyr768Ala mutation has opened the way for the accommodation and synthesis of α-1,6 isomaltooligosaccharides. Also in accordance with this suggestion, the content of α-1,6 linkage in HMM glucan also increased by 6% for this mutant and by 8% for the Tyr768Ala-Tyr695Ala double mutant. All our data are indeed in accordance with the previous finding that isomaltose is not as good an acceptor as nigerose and maltose (Côté and Robyt, 1982b; López-Munguía et al., 1993). Moreover, comparison of ASR with GTF180 and DSR-M (both being more specific for α-1,6 linkage synthesis) shows that Tyr768 is replaced by a serine. Furthermore, in both GTF180 and in DSR-M a second Trp residue is present in close proximity to the first stacking tryptophan. Given the importance of carbohydrate-aromatic interactions, we could hypothesize that this additional Trp might form an extended binding platform, not therefore present in the ASR, for oligosaccharide binding in this area (Figure 10).

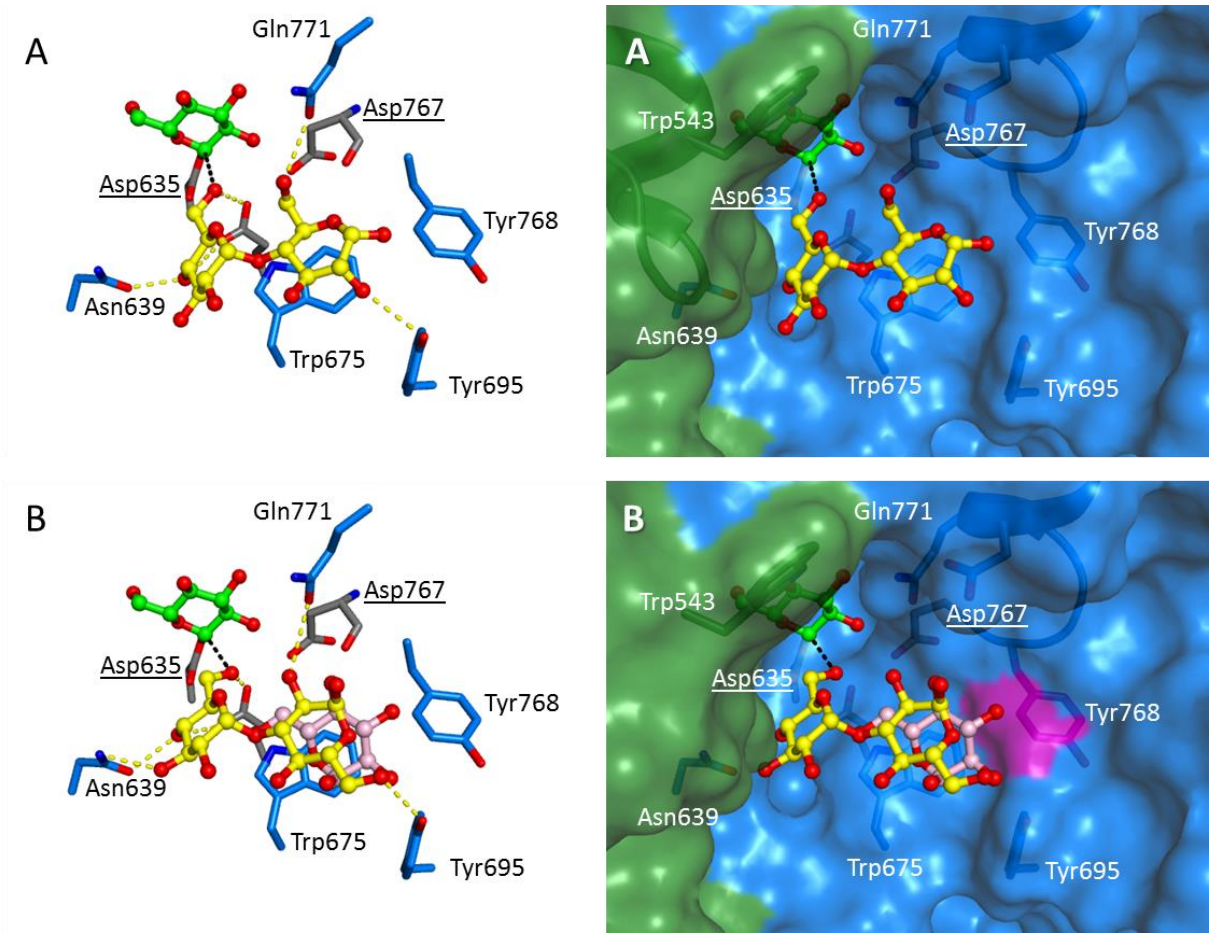


Figure 9: (A) Maltose from GTF180 (3kll) manually superposed on the catalytic cleft of the ASR. (B) Nigerose is docked automatically (yellow) and isomaltose is manually superposed (light pink) to illustrate the steric clash with Tyr768 (magenta). Two different representations are shown, the catalytic triad is colored in gray and the covalent intermediate is colored green. The direction of attack of the acceptor is indicated by a dashed black line and all the possible polar contacts are shown by dashed yellow lines.

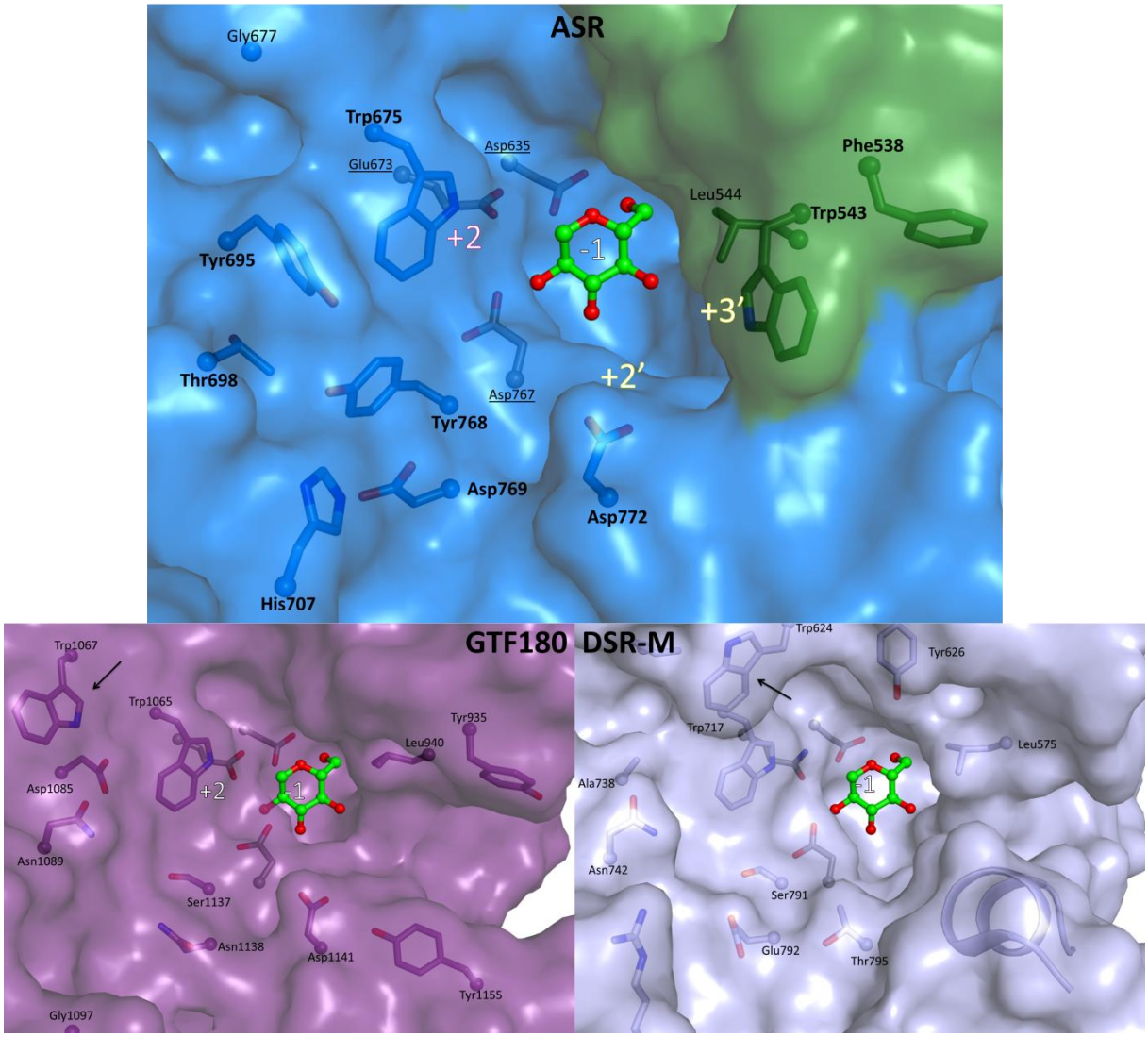


Figure 10: Comparison of the catalytic cleft and subsites of ASRΔ2, DSR-M, and GTF180.
Residues targeted in this study in ASR are shown in bold.

The control of α-1,3 linkage introduction is mediated by another subsite

If we assume that only α-1,6 glucosylation is possible from the first acceptor site (Asn639, Trp675), another mechanism should occur in a different location in the catalytic cleft that could explain the α-1,3 glucosylation. Indeed, the mutations operated in zone 2 and particularly on Asp772 and Trp543 led to a drastic decrease in α-1,3 linkage formation in both the HMM polymer obtained from sucrose and in the oligosaccharides formed by the acceptor reaction with maltose. The phenomenon was more severe for mutation on Asp772. To investigate this possibility, we performed docking of α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-D-Glc and isomaltotriose (IM3) into the catalytic cleft of ASR with the aim of identifying the best poses that oriented the O3 hydroxyl of the non-reducing glucosyl at a distance conducive to nucleophilic attack on the covalent intermediate. For IM3, the most representative acceptor pose is ensured as a result of interaction of the O6 of non-reducing terminal glucosyl with Asn639 and Glu673 (the acid/base), Asp767 (the transition state stabilizer), the middle

glucosyl ring with Gln771 and with Asp772 (defining +2' subsite), further reinforced by a stacking interaction with Trp543 (defining +3' subsite) (Figure 11A). Based on this mode of interaction we can submit that Trp543 has a role in capturing the acceptors, bringing them into the proximity of the active site (elongation mainly affected by mutation on Trp543), while the correct positioning of the acceptor is ensured by Asp772 and a network of interactions (Figure 11). A very similar positioning is obtained for α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-D-Glc, showing that acceptors with alternated linkages are accommodated well in subsite +2' (Figure 11B). It is noteworthy that our mutation Asp772Glu has the same global effect on linkage specificity as an Ala substitution (near-complete loss of α-1,3). Only an Asp sidechain, which is shorter and less flexible than a Glu sidechain, can control this very precise accommodation leading to α-1,3 glucosylation.

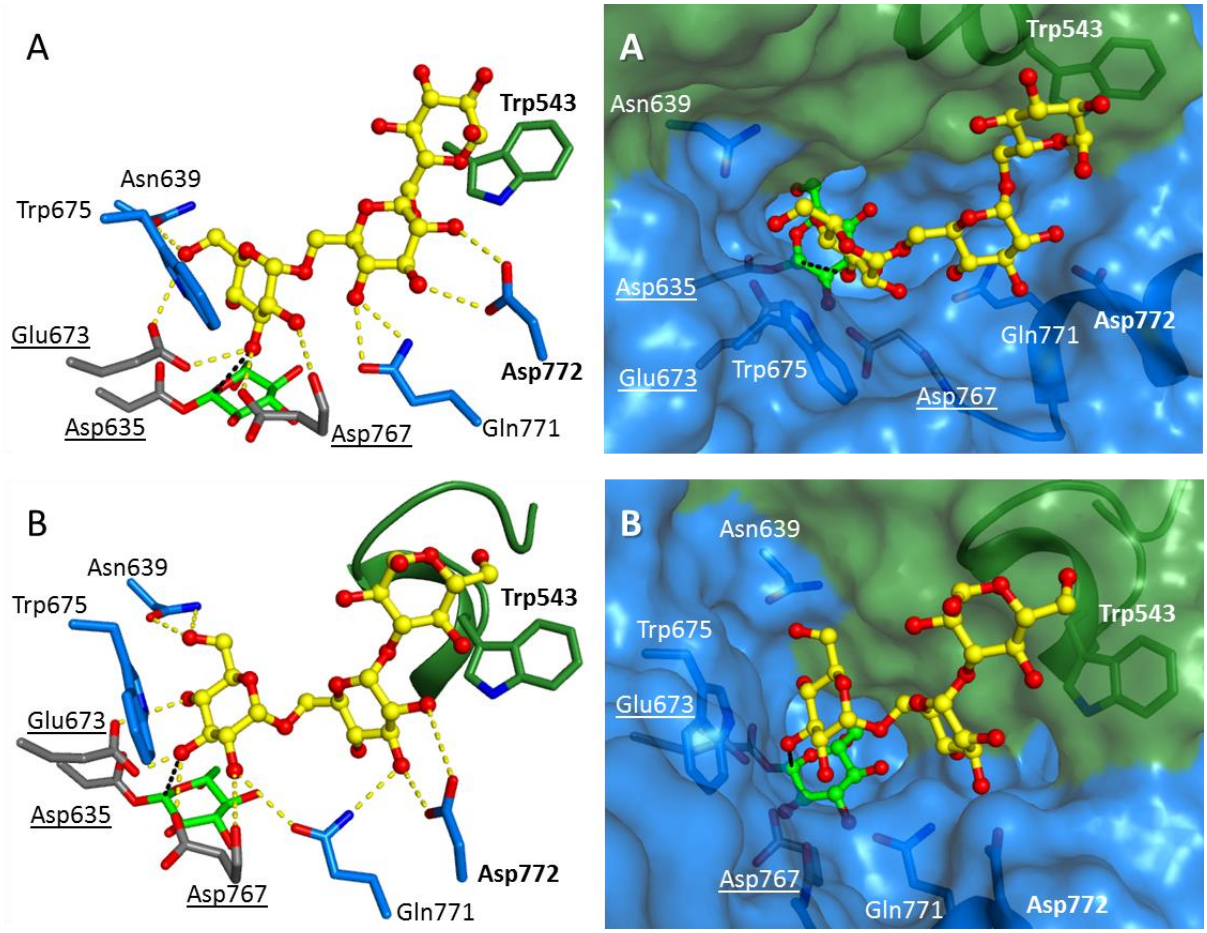


Figure 11: Isomaltotriose (IM3) (A) and α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-D-Glc (B) are automatically docked to illustrate α-1,3 glucosylation. Two different representations are shown, the catalytic triad is colored in gray and the covalent intermediate is colored in green. The direction of attack of the acceptor is indicated by a dashed black line and all the possible polar contacts are shown as dashed yellow lines.

Interestingly, Asp772 is conserved in 18 out of 63 sequences of biochemically characterized GH70 enzymes, and essentially in enzymes synthesizing polymers mainly composed of α-1,3 linkages in their linear chains of insoluble α-1,3 linked mutant type. In most of the other sequences, a Thr is

found at this position. Replacement of this aspartate with a Thr in mutant Asp567Thr GtfB from *S. mutans* GS5 or Asp569Thr GTF-I from *S. downei* MFe28 resulted in an increase in soluble glucan formation presumably containing greater numbers of α-1,6 linkages (Monchois et al., 2000b). Conversely, mutations Thr589Asp and Thr589Glu in GtfD from *S. mutans* GS5 (which synthesizes a soluble α-1,6-linked glucan) reduced soluble glucan production from 86% to 15% and 2%, respectively (Shimamura et al., 1994). From our results, we can submit that the Asp residue in these proteins plays a similar role to that suggested in ASR. The role of Trp543 cannot be directly related to other studies as the residue is unique to ASR. However, our mode of acceptor binding in the +2' subsite is close to what was already proposed by the docking calculation to explain the role of Leu940 in GTF180, the equivalent of Leu544 in ASR (Meng et al., 2014, p. 940). The Leu940Trp mutation completely abolishes α-1,3 glucosylation, showing that this region is clearly important for GTF180 linkage specificity. Additionally, dextransucrase DSR-M displays a threonine in place of the Asp772 and moreover, the presence of a short helix in this area prevents the existence of the equivalent of subsites +2' and +3' for this enzyme (Figure 10). All these observations suggest that the catalytic cleft of each GH70 enzyme has its own identity, expressed through an extremely precise interplay between the different residues and loops surrounding the active site.

ASR is designed to alternate α-1,6 and α-1,3 linkages

Nigerotriose has never been described in alternan or in oligosaccharides formed from maltose (Côté and Robyt, 1982b; Goldstein and Whelan, 1962) meaning that ASR is unable to α-1,3 elongate a terminal nigerose moiety. Docking of nigerose or α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc failed to identify acceptable binding poses in the +2' subsite that could lead to α-1,3 glucosylation. The formation of two consecutive α-1,3 linkages may be disfavored for the following reasons: (i) the specific interaction with Asp772 will necessarily be disrupted by the presence of an O3 substituent and (ii) the shortness and rigidity of an α-1,3 linkage (compared to an α-1,6) will position the non-reducing glucosyl either too far from the covalent intermediate or in a poor orientation, which will result in a non-productive complex. We cannot totally exclude the possibility (although to a limited extent) of consecutive α-1,6 linkage formation from the +2' subsite due to the flexibility of the α-1,6 linkage. In addition, we do not explicitly address the question of branch formation. However, the active site is wide and the presence of both subsites +2 and +2' bordering subsite +1 should allow for alternative accommodations of isomaltotriosyl units and explain the presence of 3,6 di-substituted glucose in alternan (Figure S11).

Bringing together all our data, we submit that α-1,3 linkage formation is governed by the +2' subsite with the help of Asp772, only if the acceptor is previously α-1,6 linked. The formation of consecutive

α-1,3 linkages is prevented and further elongation of an α-1,3 linked moiety will continue only with an α-1,6 linkage. Comparison of the binding modes leading to either α-1,6 or α-1,3 linkages reveals that the accommodation of the glucosyl unit in subsite +1 (Asn639) is indeed split between two different binding modes: the proper orientation leading to α-1,6 glucosylation is ensured by Trp675 (subsite +2), whereas the orientation leading to α-1,3 glucosylation is controlled by Asp772 (subsite +2'), assisted by Trp543 (subsite +3'). As such, subsite +1 does not control linkage specificity and rather, appears to be "undecided" with regard to the linkage promiscuity of ASR. It is the positioning of the glucosyl unit in subsite +2 or +2' that determines linkage specificity and eventually resolve the ASR linkage hesitation, as was previously proposed by Moulis *et al.* (Moulis et al., 2006). It can be assumed that the mechanism will operate during the synthesis of longer oligosaccharides as the enzyme catalytic cleft will only "see" the last 2 or 3 glucosyl units of the growing chain. If we extrapolate this for the synthesis of an HMM alternan, the "round-trip" mechanism between the two sites would result in an accumulation of alternating α-1,6 and α-1,3 linkages in the polymer, with the only variable being the number of consecutive α-1,6 linkages. This is not only in accordance with the α-1,6/α-1,3 ratio observed in the HMM alternan (roughly 2:1) but also corroborates the fact that linkage alternation is kinetically determined (Côté et al., 2008).

Conclusion

We disclosed herein the first crystal structure of ASR from *L. citreum* NRRL B-1355, the largest structure of a glucansucrase solved so far. Structural analysis showed that the enzyme is stabilized thanks to its calcium binding site and an entire domain V. In addition, the presence of longer loops in domain A and extended hydrophobic packing in domain C have been identified as traits specific to ASR that could influence the enhanced enzyme's stability. Concerning catalytic activity, we showed that domain V participates in polymer elongation, especially to produce HMM alternan. Focusing on the catalytic core of the enzyme, we selected several mutation targets and discerned their effect on alternan or oligoalternans formation using a combination of acceptor reaction and HPLC/MS. Combining these data with molecular docking and structural analysis, we identified two different acceptor subsites, +2 and +2', that are specific to the synthesis of α-1,6 and α-1,3 linkages respectively. Our data set allowed us to propose the mechanism by which the ASR catalytic cleft differentiates the incoming acceptors based on the nature of the terminal glucosidic linkage. This mechanism leads to the formation of alternated α-1,3 or α-1,6 linkages in the alternan polymer. Our results shed new light on a long known but very interesting enzyme that could pave the way for further engineering efforts with the aim of producing tailor-made α-glucan polysaccharides.

Experimental procedures

Truncated mutant design and construction

SignalP (Petersen et al., 2011), DisEMBL (Linding et al., 2003), RONN (Yang et al., 2005) and PSIPRED (Buchan et al., 2013) web servers were used to design N-terminal truncations.

Genes of *asr*-Δ2, *asr*-Δ3, *asr*-Δ4 and *asr*-Δ5 were amplified by PCR using the *asr*-Cdel gene as a template, Phusion polymerase® (NEB), and the following primers: *asr*-Δ2 forward primer CAC-CGC-GGA-TAC-AAA-TTC-G, *asr*-Δ3 forward primer CAC-CGG-TTT-TTG-GTA-TGA-TTC-AG, *asr*-Δ4 forward primer CAC-CAT-CAC-TGG-GGG-TCA-C, *asr*-Δ2/Δ3/Δ4 reverse primer CCC-TCG-AGA-CAT-AGT-CCC-ATC, *asr*-Δ5 forward primer CAC-CCA-AAG-TAA-TGA-AAA-TAC-TCC and *asr*-Δ5 reverse primer CGC-ATC-TTT-ATT-CTG-CAA-CTG. A "CACC" sequence was added at the beginning of each forward primer for cloning in pENTR D-TOPO vector before being recombined in pET53-DEST vector (Gateway system, Thermo Fisher Scientific). The recombined product was transformed into competent *E. coli* TOP10 (Invitrogen). Each gene in the pENTR plasmid was verified by sequencing (GATC Biotech).

Production and purification of truncated enzymes

The *E. coli* BL21 DE3* strain was used for enzyme production. A preculture of transformed *E. coli* BL21 DE3* in LB medium supplemented with ampicillin 100 $\mu g.mL^{-1}$ was used to inoculate a culture at an OD_{600nm} of 0.05 in ZYM-5052 auto-inducible medium (Studier, 2005) modified by supplementation with 100 $\mu g.mL^{-1}$ ampicillin, 1% (w/v) α-lactose, and 1% (w/v) glycerol for pET53 enzyme production. After 26 hours of growing at 21°C, cells were harvested by centrifugation and resuspended in Binding buffer containing 20 mM phosphate buffer, 20 mM Imidazole (Merck Millipore), and 500 mM NaCl, pH 7.4 supplemented with EDTA-free anti-protease tablets (Roche). Cells were disrupted by sonication and debris was removed by a centrifugation step at 45,000 g for 30 minutes at 8°C. Purification was performed with the ÄKTA Xpress system (GE Healthcare). Two-step purification was performed in a cold chamber at 8°C using (i) a HisTrap HP 1mL column (GE Healthcare) for the affinity step and (ii) a Superose12 16/60 (GE Healthcare) for the size exclusion step, or a HiPrep desalting 26/10 column (GE Healthcare) for desalting. The size exclusion step was performed upstream of crystallization trials and Differential Scanning Fluorimetry assays, and protein was eluted in MES buffer pH 6.5 at 30 mM with 100 mM NaCl and 0.05 $g.L^{-1}$ $CaCl_2$. The desalting step was performed for biochemical characterization, for which protein was eluted in 50 mM sodium acetate buffer pH 5.75. Purified fractions were pooled together and concentrated using AmiconUltra-15 with a cut-off of 50 KDa to 10-15 $mg.mL^{-1}$. Purification was checked by SDS-PAGE electrophoresis

using NuPAGE 3-8% Tris-Acetate protein gels (Invitrogen), and protein concentration was assessed by spectroscopy at 280 nm using a NanoDrop instrument. The theoretical molecular weight and molar extinction coefficient of the enzyme were calculated using the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

Mutagenesis study

Mutants were constructed by inverse PCR using the pET53-*asr*-Δ2 gene as a template, Phusion® polymerase (NEB), and the primers described in Table S1. Following overnight *Dpn*I (NEB) digestion, the PCR product was transformed into competent *E. coli* DH5α and clones were selected on solid LB medium supplemented with ampicillin 100 µg.mL^{-1}. Plasmids were extracted with the QIAGEN spin miniprep kit and mutated *asr* genes were checked by sequencing (GATC Biotech). Mutants were produced and purified as described above.

Activity measurement

Activity was determined in triplicate at 30°C in a Thermomixer (Eppendorf) using the 3,5-dinitrosalicylic acid method (Miller, 1959). 50 mM sodium acetate buffer pH 5.75, 292 mM sucrose and 0.05 mg.mL^{-1} of pure enzyme were used. One unit of activity is defined as the amount of enzyme that hydrolyzes 1 µmol of sucrose per minute. To evaluate the calcium effect, initial activity was measured with 3.4 mM of calcium chloride, 5 mM EDTA, or without additives under the same conditions as above.

Enzymatic reactions and product characterization

Polymer productions were performed using 1 U.mL^{-1} of pure enzyme with 292 mM sucrose in 50 mM NaAc buffer pH 5.75 at 30°C over a period of 24 hours. The products were analyzed using High Pressure Size Exclusion Chromatography (HPSEC) with Shodex OH-Pak 805 and 802.5 columns in series in a 70°C oven with a flow rate of 0.250 mL.min^{-1} connected to RI detector. The eluent was 50 mM sodium acetate, 0.45 M sodium nitrate and 1% (v/v) ethylene glycol. The ratios of HMM polymer to LMM polymer produced were calculated using the area of each peak divided by the sum of the areas. The same sample was analyzed in triplicate in HPSEC.

Acceptor reactions were set up in the presence of maltose (sucrose:maltose mass ratio 2:1) in the same conditions. The products were analyzed by High Pressure Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) using a CarboPac TM PA100 guard column upstream of a CarboPac TM PA100 analytical column (2 mm x 250 mm) at a flow rate of 0.250 mL.min^{-1}. The eluents were A: 150 mM NaOH and B: 500 mM sodium acetate with 150 mM NaOH.

Sugars were eluted with an increasing 0 to 60% gradient of eluent B for 30 minutes. Quantification was performed using standards of glucose and sucrose at 5, 10, 15 and 20 $mg.L^{-1}$. The hydrolysis percentage was calculated by dividing the final molar concentration of glucose by the initial molar concentration of sucrose. The degree of polymerization (DP) of the oligosaccharides produced was determined using HPAEC coupled with a simple quadrupole mass spectrometer (ISQ EC, Thermo Scientific) and sugars were eluted under the same conditions as above with a 0 to 30% gradient of eluent B for 45 minutes. A mass spectrometer fitted with an ESI ion source was used with the following parameters: Collision Induced Dissociation (CID) of 10 V, a vaporizer temperature of 289°C, an ion transfer tube temperature of 300°C and a source voltage of 3kV in positive mode. The mass spectrometry range was *m/z* 100 to 1250 (maximum DP: 7). To identify the peaks, an acceptor reaction was set up under the same conditions as above with the dextransucrase DSR-S vardel Δ4N, which elongates maltose only through α-1,6 linkages. These peaks correspond to oligodextran (OD) and only contain linear α-1,6 linkages joined to maltose. The ASRΔ2 and DSR-S vardel Δ4N oligosaccharides were co-eluted to ensure peak attribution (Figure S1). The other main oligosaccharides obtained with the ASRΔ2 reaction from maltose were named oligoalternan (OA) of DP3 to 6. Oligoalternans contain both α-1,6 and α-1,3 linkages and their structure is well defined in the literature (Côté and Robyt, 1982b; Côté and Sheng, 2006; López-Munguía et al., 1993).

NMR samples were prepared by dissolving 10 mg of the total products from sucrose in 0.5 mL D_2O. Deuterium oxide was used as the solvent, and sodium 2,2,3,3-tetradeuterio-3-trimethylsilylpropanoate ($TSPD_4$) was selected as the internal standard (δ^1H = 0 ppm, $\delta^{13}C$ = 0 ppm). ^1H and ^{13}C NMR spectra were recorded on a Bruker Avance 500-MHz spectrometer operating at 500.13 MHz for ^1H NMR and 125.75 MHz for ^{13}C using a 5-mm z-gradient TBI probe. The data were processed using TopSpin 3.0 software. 1D ^1H NMR spectra were acquired by using a zgpr pulse sequence (with water suppression). Spectra were performed at 298 K with no purification step, for all mutants. Spectrum was performed at 343 K for the ASRΔ2 alternan polymer after purification with dialysis using 14 kDa cut-off cellulose dialysis tubing (Sigma-Aldrich) in water.

For the wild type enzyme and all the variants (truncated forms and single and double mutants), Differential Scanning Fluorimetry was performed with 7 µM of pure enzyme in 50 mM sodium acetate buffer pH 5.75 supplemented with 0.5 $g.L^{-1}$ of calcium chloride and 10 X of SYPRO orange (Life Technologies). A ramp from 20 to 80°C was applied with 0.3°C increments at the rate of 0.3°C per second on a C100 Thermal Cycler.

Crystallization and Data collection

The initial conditions were screened using JCSG+ and PACT screens (QIAGEN). A Mosquito robot was used to make sitting drops by mixing 0.2 µL of enzyme solution with 0.2 µL of reservoir solution. The initial hits were reproduced and diffraction quality crystals of ASRΔ2 appeared after three months at 12°C using 17% (w/v) PEG 3350, $NaNO_3$ 0.5 M as a precipitant and an enzyme concentration of 8 $mg.ml^{-1}$ with the presence of 10 mM isomaltohexaose (IM6). Crystal was cryoprotected in the reservoir solution supplemented with 15% (v/v) ethylene glycol and cryo-cooled directly in liquid nitrogen. Diffraction data were collected on beamline ID30A-3 of the European Synchrotron Radiation Facility (Grenoble, France).

Structure solving and refinement

Images were integrated using XDS (Kabsch, 2010) and converted to structure factors using CCP4 programs (Winn et al., 2011). The structure was solved by molecular replacement using PHASER and the N-terminally truncated DSR-M as a search model (5LFC). To complete the model, cycles of manual rebuilding using COOT were alternated with automatic rebuilding cycles using BUCCANEER and refined using REFMAC5 (Murshudov et al., 2011). The final model was evaluated using the WHATIF (Vriend, 1990) and MOLPROBITY (Chen et al., 2010) web servers and deposited in the PDB under the accession code 6HVG. To analyze the network of interactions in the structure, the RING (Piovesan et al., 2016) web server was used. Data collection and refinement statistics are shown in Table S2.

Structure analysis, manual and automated docking

The ASRΔ2 and DSR-M (5NGY) domain V structures were aligned to manually dock the isomaltotetraose into the putative sugar binding pocket V-A of ASRΔ2. The ASRΔ2 and GTF180 (3HZ3 or 3KLL) structures were aligned to manually dock the sucrose or the maltose into the active site of ASRΔ2. For the docking calculations, a model of the ASRΔ2 glucosyl-enzyme intermediate was constructed based on the high resolution structure of the GH13 covalent intermediate (PDB 1S46). Oligosaccharide structures were built using the Glycam server (Woods Group. (2005-2018) GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA. (http://glycam.org)). Receptor and ligand structures were prepared with Autodock Tools and docked using Vina-Carb (Nivedha et al., 2016). Representative structures were selected from all the resulting conformers to meet a distance criterion of ≤ 3.5Å from the O3 or O6 hydroxyl groups of the non-reducing terminal glucose to both the C1 atom of the glucosyl-enzyme intermediate and the side chain of the acid/base catalyst (E673).

Acknowledgements

We are grateful to MetaSys, the Metabolomics & Fluxomics Center at the Laboratory of Biological Systems and Process Engineering (Toulouse, France), for the NMR experiments. Our special thanks go to the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and the Structural Biophysics team of the Institute of Pharmacology and Structural Biology (IPBS, Toulouse, France) for access to the crystallization facility and help in synchrotron data collection. We wish to thank the ICEO facility, which is part of the Integrated Screening Platform of Toulouse (PICT), for providing access to HPLC and protein purification equipment. We are very grateful to Florent Grimaud for his technical assistance.

Supporting information

Table S1: Primer used for site-directed mutagenesis

Mutant	Forward primer (5' -> 3')	Reverse primer (5' -> 3')
Phe538Ala	GAAGCTGCCGATGGTTTGC	CCATCGGCAGCTTCATCTTCTG
Trp543Ala	GGTTTGCAGGCGCTTCAAGGGGGATTC	CCTTGAAGCGCCTGCAAACCATCGAAAG
Trp675Ala	TTTAGAAGACGCGAACGGTAAAGATC	TTTACCGTTCGCGTCTTCTAAAATAGAC
Trp675Phe	GAAGACTTCAACGGTAAAGATCCTCAG	ACCGTTGAAGTCTTCTAAAATAGACAAATGC
Trp675His	TTAGAAGACCATAACGGTAAAGATCC	CTTTACCGTTATGGTCTTCTAAAATAGAC
Trp675Tyr	TTAGAAGACCATAACGGTAAAGATCC	GATCTTTACCGTTATAGTCTTCTAAAATAGAC
Tyr695Ala	ATGGATGCCACAGTTACTTCACAGTTTG	AACTGTGGCATCCATTGTTAATTGC
Thr698Ala	TACACAGTTGCTTCACAGTTTGGC	CAAACTGTGAAGCAACTGTGTAATCC
Tyr768Ala	CACATGATGCCGATGCTCAAGATCCAATTAG	TCTTGAGCATCGGCATCATGTGCTCTAAC
Tyr768Trp	CACATGATTGGGATGCTCAAGATCCAATTAG	TCTTGAGCATCCCAATCATGTGCTCTAAC
Tyr768Phe	CACATGATTTCGATGCTCAAGATCCAATTAG	CTTGAGCATCGAAATCATGTGCTCTAAC
Asp769Ala	CATTGATTACGCTGCTCAAGATCCAATTAG	TTGAGCAGCGTAATCATGTGCTCTAAC
Asp772Ala	CGATGCTCAAGCTCCAATTAGAAAAGC	TCTAATTGGAGCTTGAGCATCGTAATC
Asp772Glu	CGATGCTCAAGAACCAATTAGAAAAGC	TTTCTAATTGGTTCTTGAGCATCGTAATCATG

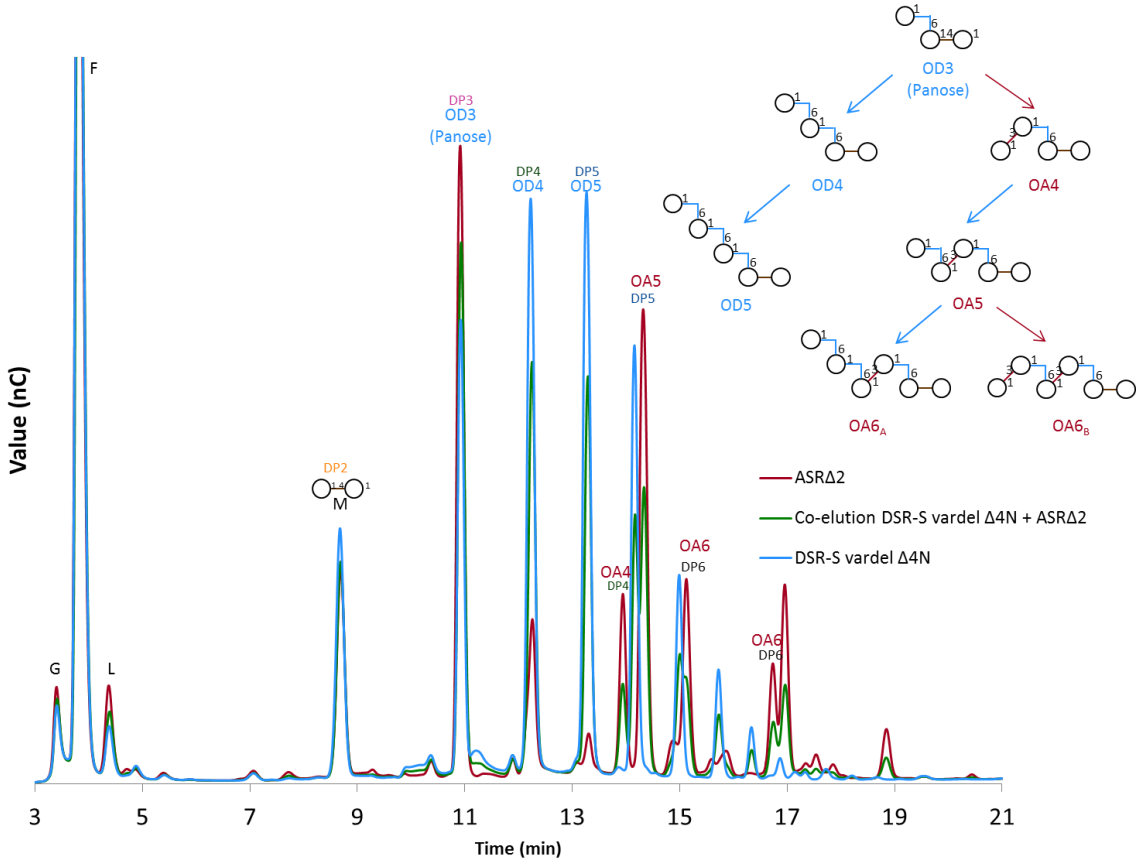


Figure S1: Superposition and co-elution of acceptor reaction of DSR-S on maltose and model of product formation for DP 3 to 6, adapted with permission from reference (Côté and Sheng, 2006, p. 206). Copyright 2006, Elsevier

Reaction from 292 mM sucrose and 146 mM maltose with 50 mM sodium acetate buffer pH 5.75 and 1 $U.mL^{-1}$ of pure enzyme. G: Glucose, F: Fructose, L: Leucrose, M: Maltose ,OD: Oligodextran (identified using a superposition with DSR-S acceptor reaction in the same condition), OA: Oligo-alternan.

Table S2: Data collection of the ASR crystal. Values in parenthesis refer to high resolution shell.

Data collection		
Wavelenght (Å)		0.9677
Space group		$P2_12_12_1$
Molecules per asymmetric unit		2
Cell constants a, b, c (Å)		101.23, 134.80, 237.03
α, β, γ (°)		90.00, 90.00, 90.00
Resolution (Å)		47.82-2.80 (2.95-2.80)
Measured reflections		551987 (81876)
Unique reflections		80120 (11606)
Data completeness %		99.4 (99.9)
Data redundancy		6.9 (7.1)
Rmerge		0.088 (0.571)
< I/s (I) >		7.3 (1.4)
CC1/2		0.99 (0.87)
Wilson B-factor (Å²)		60.3
Refinement statistics		
R , R_{free}		0.207 , 0.237
Number of		
	protein atoms	10013 (chain A) 9180 (chain B)
	water molecules	279
	calcium atoms	2
Average B, main chains (Å²)		62.0 (chain A) 70.1 (chain B)
Average B, side chains (Å²)		63.8(chain A) 71.2 (chain B)
RMS Deviation		
	bond lengths (Å)	0.009
	bond angles (°)	1.265
Ramachandran plot (%)		
	favored	96
	allowed	4
	outliers	0
Molprobity Score	5.5 (100^{th} percentile)	

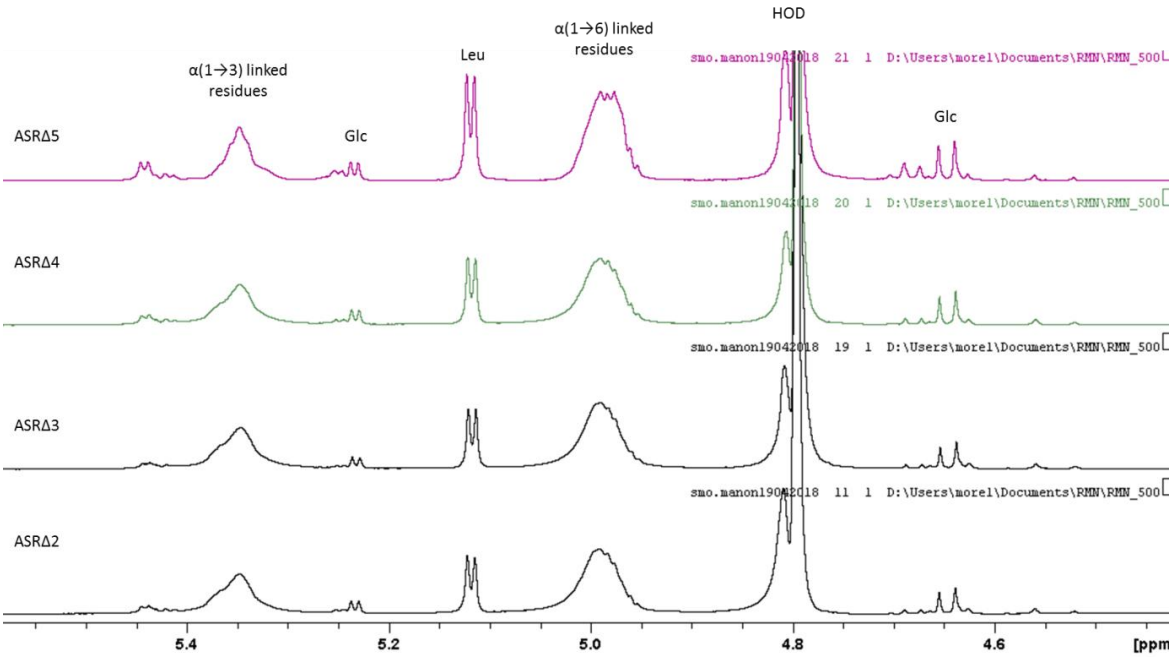


Figure S2: Proton NMR of alternan produced by truncated mutants ASRΔ2 to Δ5.
Glc: Glucose, Leu: Leucrose.

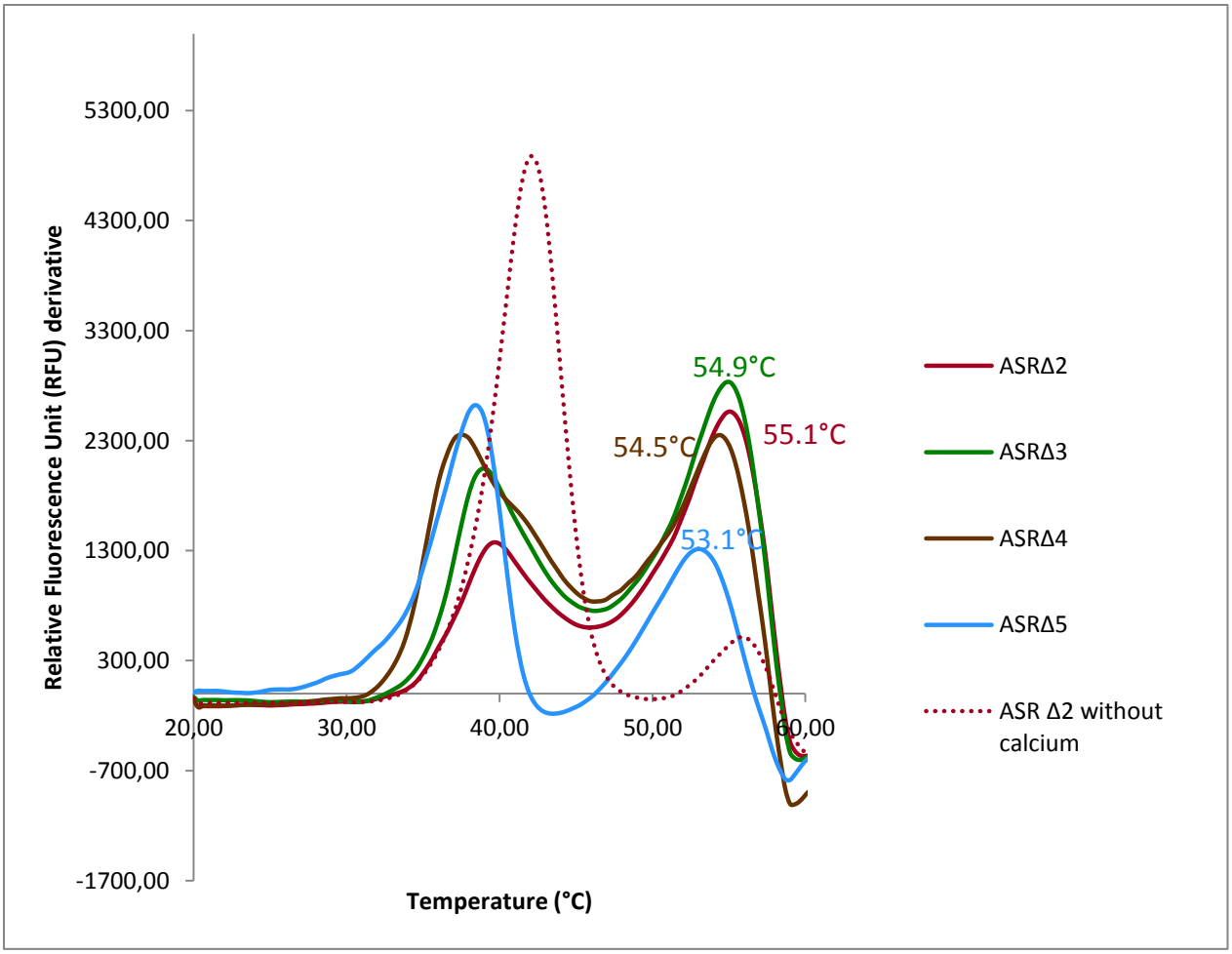
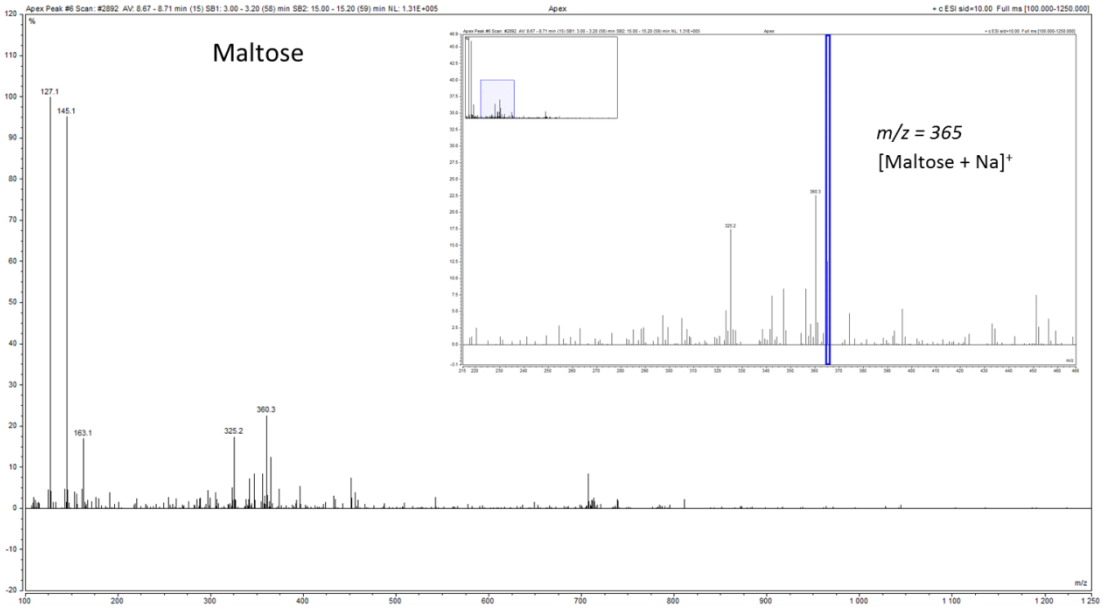
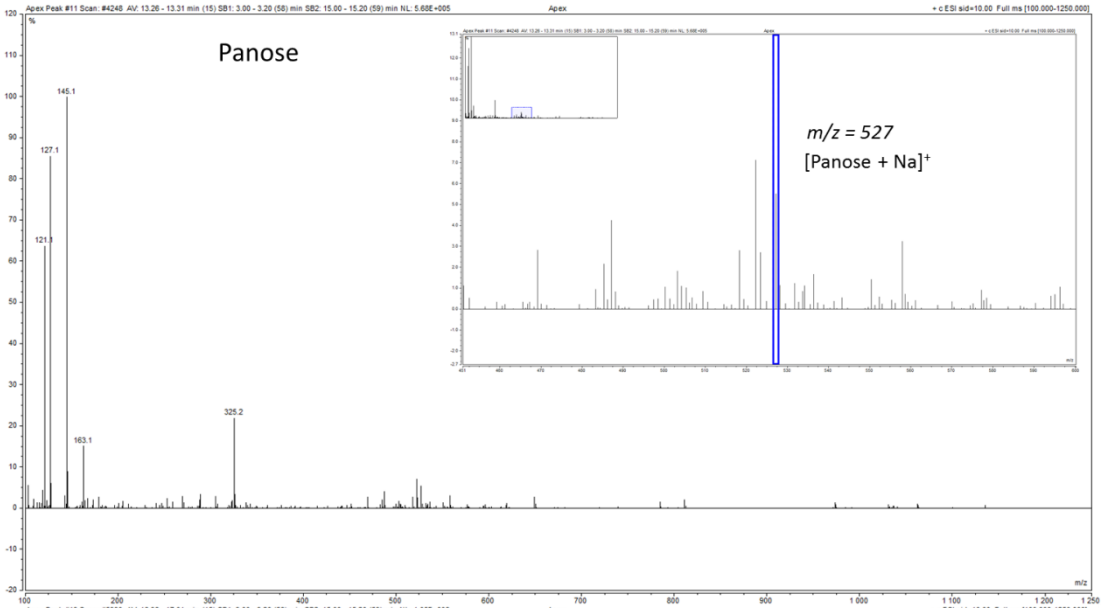
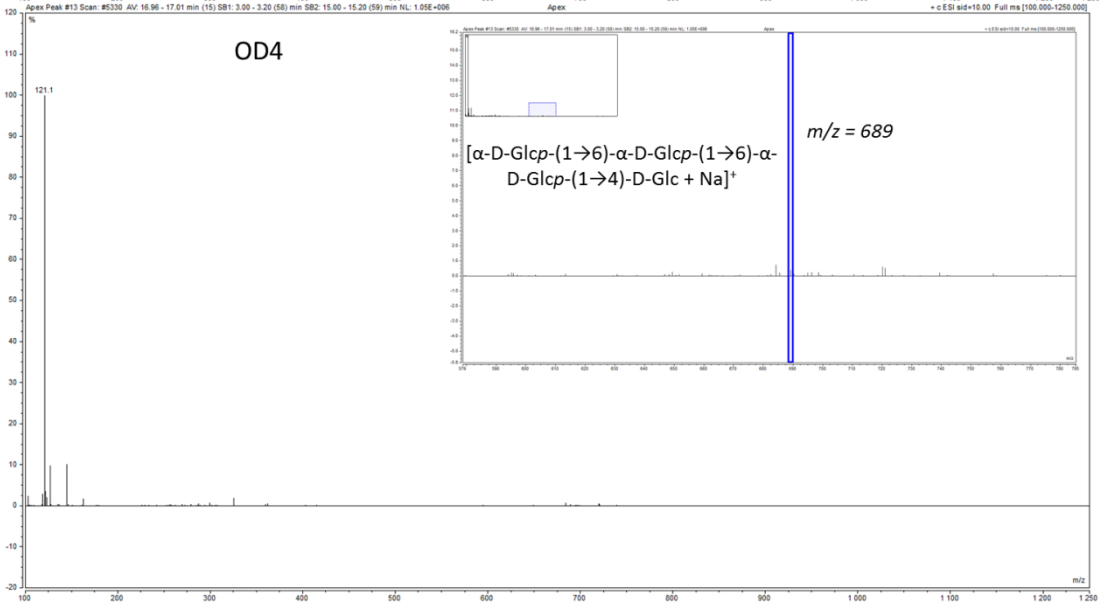
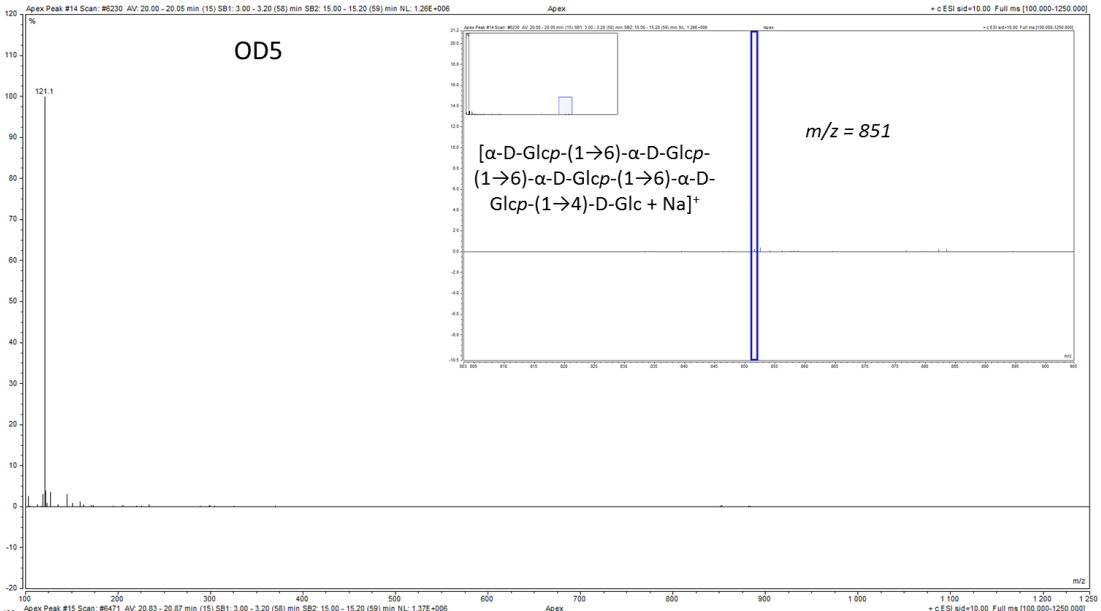


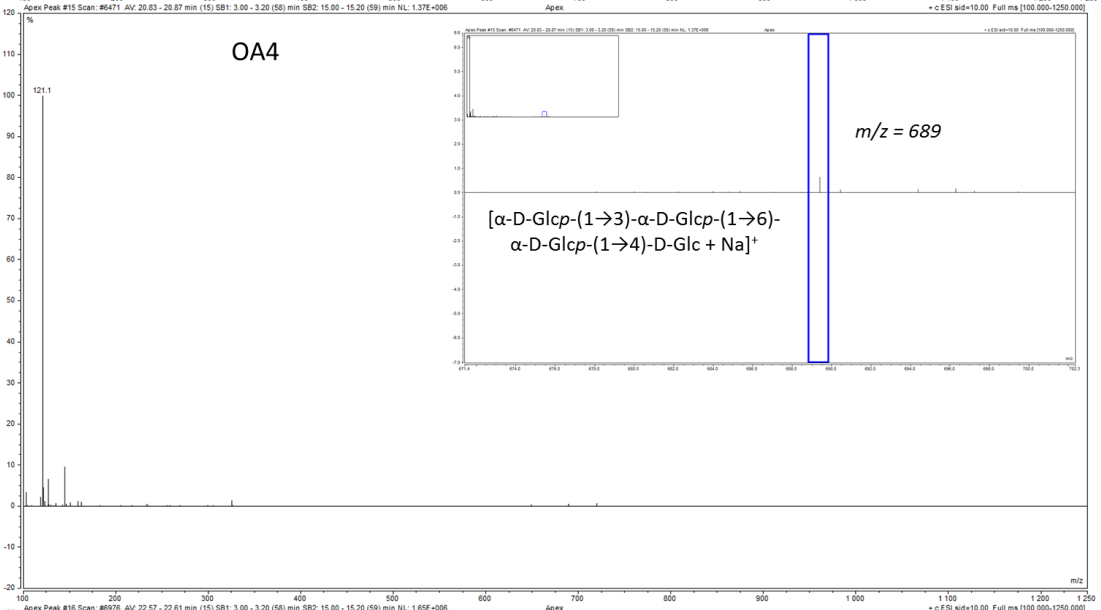
Figure S3: Differential Scanning Fluorimetry (DSF) experiment of the truncated mutants with values of the second transition observed. Comparison with the curve without calcium. Buffer used: NaAc 50 mM pH 5.75 with or without 3.4 mM $CaCl_2$. 7µM of pure enzyme.

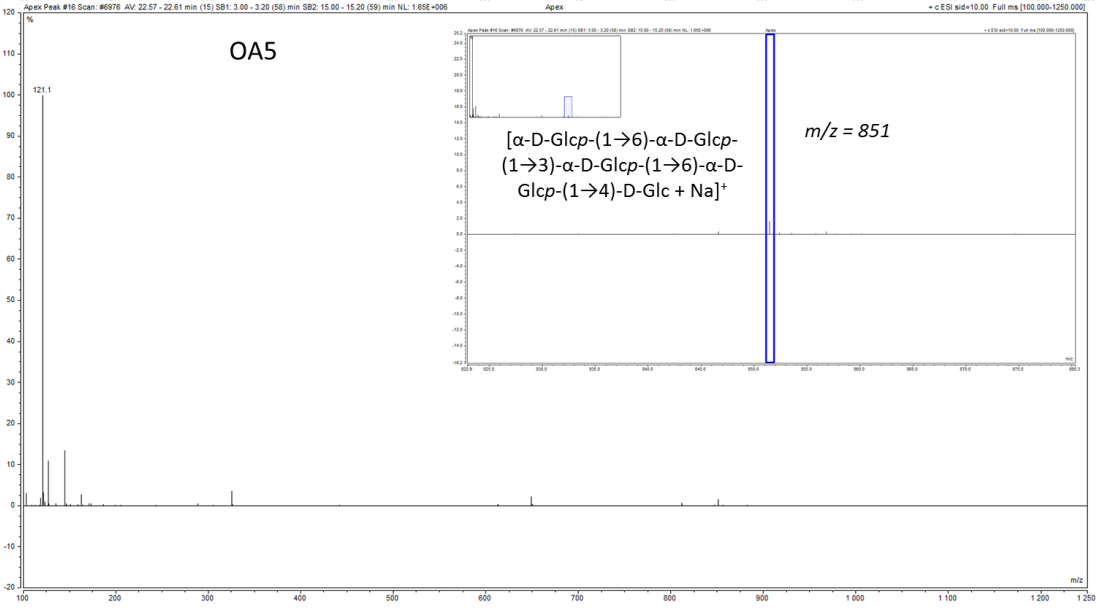

Maltose
m/z = 365
[Maltose + Na]+
127.1
145.1
163.1
325.2
360.3


Panose
m/z = 527
[Panose + Na]+
145.1
127.1
121.1
163.1
325.2


OD4
m/z = 689
[α-D-Glcp-(1→6)-α-D-Glcp-(1→6)-α-D-Glcp-(1→4)-D-Glc + Na]+
121.1


OD5
121.1
m/z = 851
[α-D-Glcp-(1→6)-α-D-Glcp-(1→6)-α-D-Glcp-(1→6)-α-D-Glcp-(1→4)-D-Glc + Na]+
m/z


OA4
121.1
m/z = 689
[α-D-Glcp-(1→3)-α-D-Glcp-(1→6)-α-D-Glcp-(1→4)-D-Glc + Na]+
m/z


OA5
121.1
m/z = 851
[α-D-Glcp-(1→6)-α-D-Glcp-(1→3)-α-D-Glcp-(1→6)-α-D-Glcp-(1→4)-D-Glc + Na]+
m/z

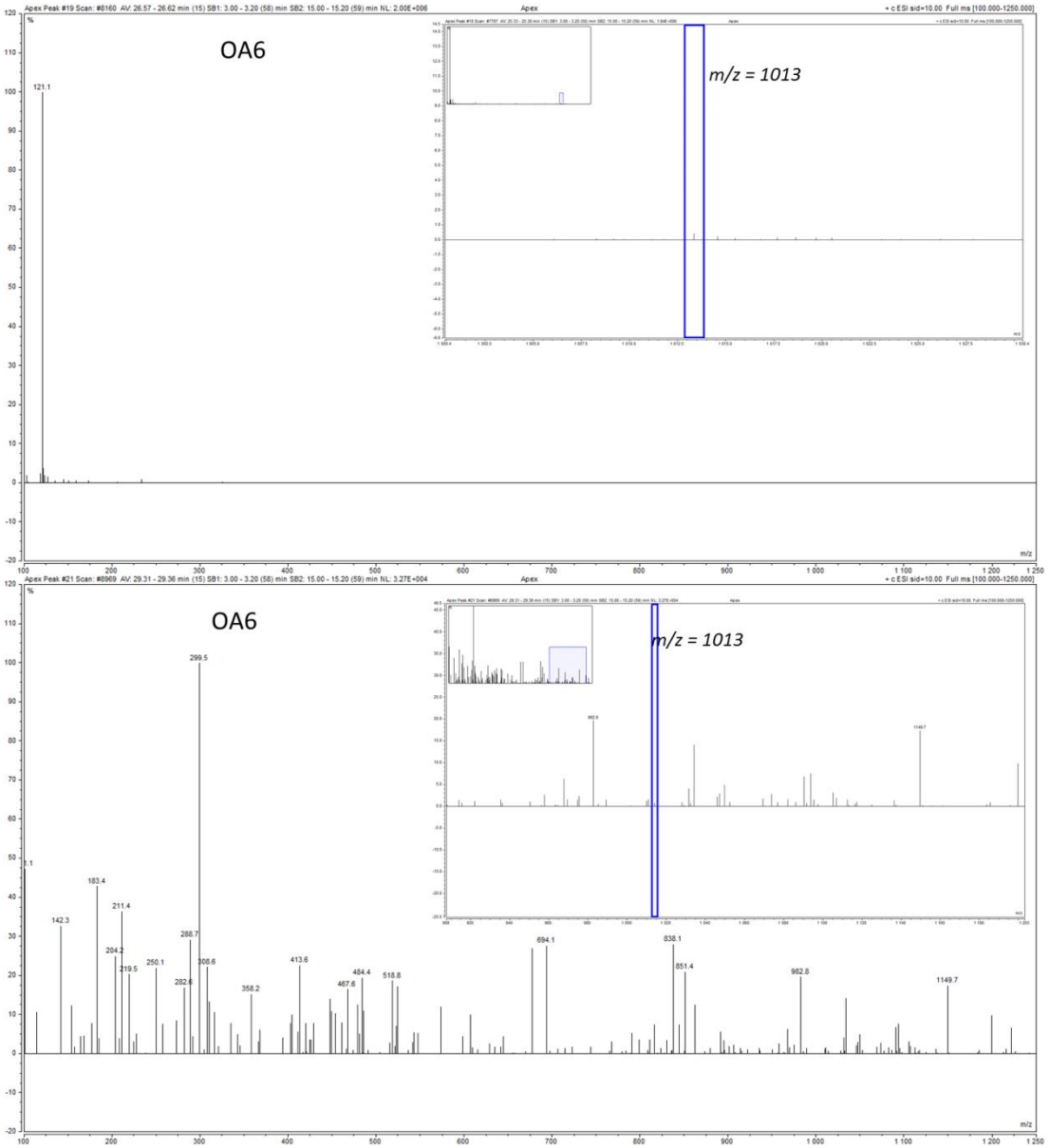


Figure S4: Mass spectrum of acceptor reaction from maltose products with ASRΔ2

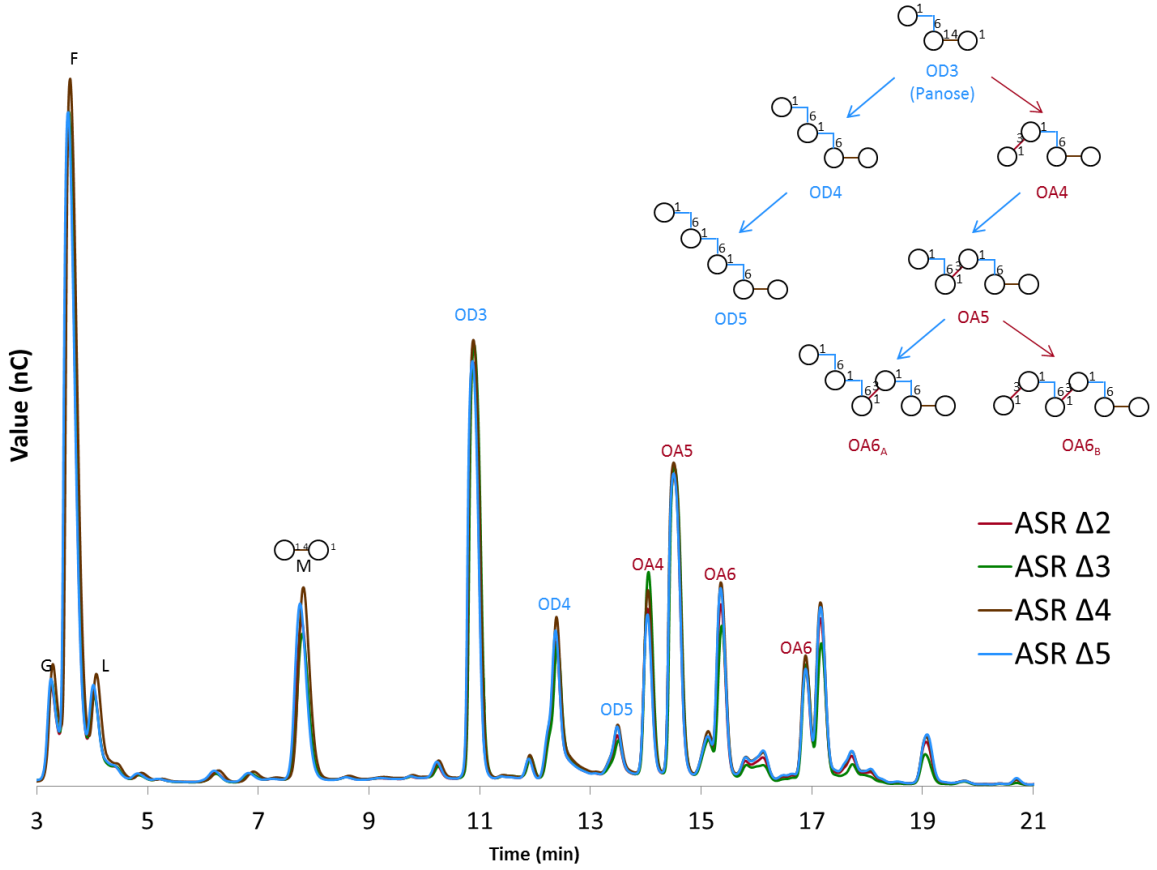


Figure S5: HPAEC-PAD chromatograms of truncated mutants ASRΔ2 to Δ5 and model of product formation from maltose for DP 3 to 6, adapted with permission from reference (Côté and Sheng, 2006). Copyright 2006, Elsevier.

Reaction from 292 mM of sucrose and 146 mM of maltose with 50 mM sodium acetate buffer pH 5.75 and 1 $U.mL^{-1}$ of pure enzyme. G: Glucose, F: Fructose, L: Leucrose, M: Maltose ,OD: oligodextran (identified using a superposition with DSR-S acceptor reaction in the same condition), OA: Oligo-alternan.

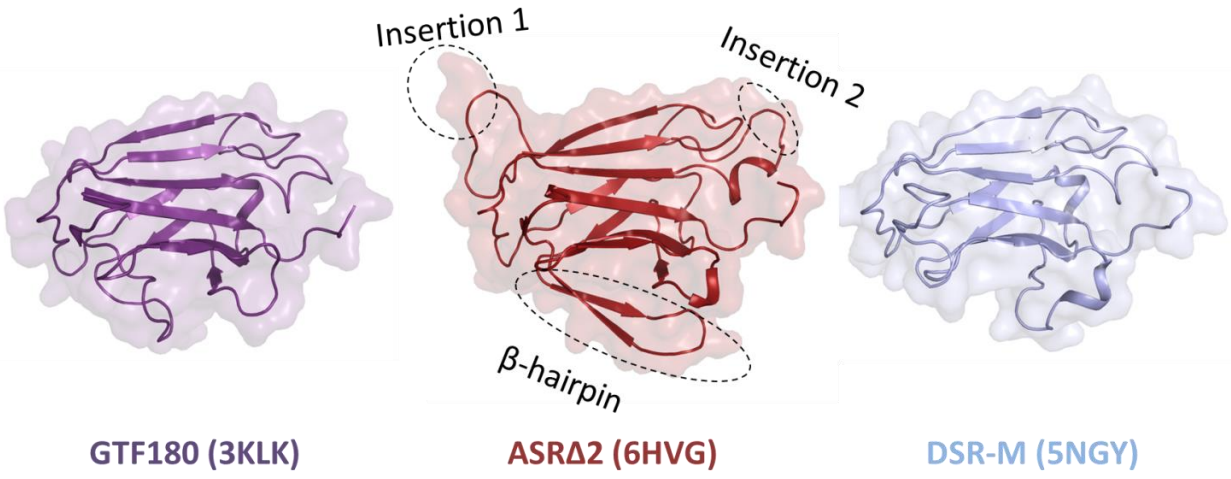


Figure S6: Comparison of the domain C of GTF180, ASRΔ2 and DSR-M.

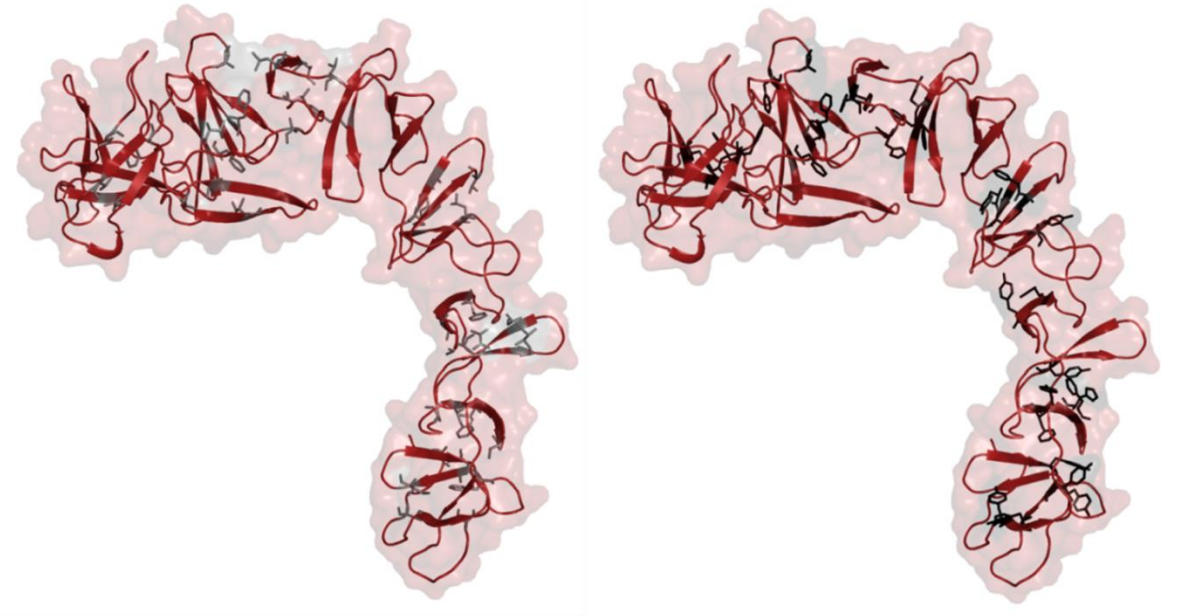


Figure S7: Hydrophobic packaging of ASR.
Grey: hydrophobic residue forming local small hydrophobic core. Black: residues forming π-π stacking interactions (predicted by RING).

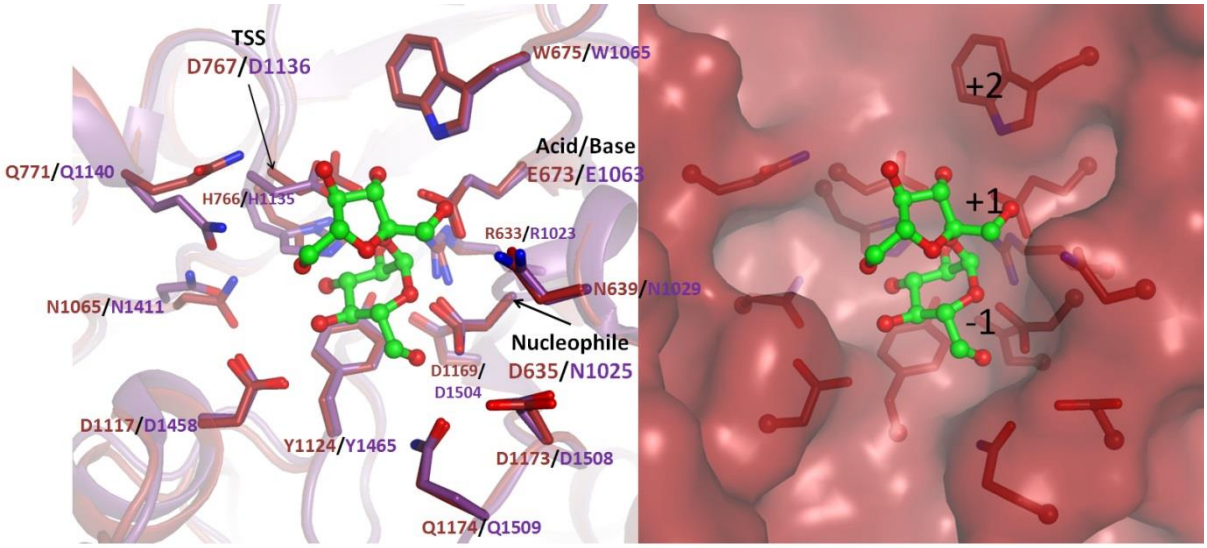


Figure S8: View of the catalytic site of ASR with sucrose docking from 3HZ3.
Red: ASR, Purple: GTF180, Green: sucrose.

```
Conservation:
6HVG.pdb_chainA_s001 1 SNLSDPITGGHYENHNGYFVYIDASGKQVTGLQN-IDGNLQYFDD-NGYQVKGSFRDVNGKHIYFDSVTG 68
5ngy_chainA_p002 1 --VPETITGGRYSLKDGYYVYLDKQGKQVVGPKN-IDNHLQYFDETTGKQVKGDFRSVNGKRIYFNANLG 67
3klk_chainA_p006 ------------------------------------------------------------------------
4amc_chainA_p005 ------------------------------------------------------------------------
3aib_chainB_p003 ------------------------------------------------------------------------
3ttq_chainA_p004 1 ------------------------------GLRQDSNGKLRYFDLTTGIQAKGQFVTIGQETYYFSKDHG 40
Consensus_aa: ............................................................................
Consensus_ss: ee e eeee eeeeeeee eeeee

Conservation:  96
6HVG.pdb_chainA_s001 69 KASSNVDIVNGKAQGYDAQGNQLKKSYVADSSGQTYYFDGNGQPLIGLQTIDGNLQYFNQ-QGVQIKGGF 137
5ngy_chainA_p002 68 YADDYTTDVAGKLVGYDSNGNQVKAGYVTNSQGKTYYFNNQGEAIIGLKTDNNKTQYFGP-DGAQVKGAF 136
3klk_chainA_p006 1 ----------------------------------------------------------QY---------- 2
4amc_chainA_p005 1 ----------------------------------------------------------QY---------- 2
3aib_chainB_p003 ------------------------------------------------------------------------
3ttq_chainA_p004 41 DAQLLPMV---------------TEGHYGTITTAWVYRDQNNTILKGLQNINGTLQFFDPYTGEQLKGGV 95
Consensus_aa: ..........................................................Q@............
Consensus_ss: eeeeeeee ee eeee eeeeeeee eeeee eee e

Conservation: 96 69 6 6 6
6HVG.pdb_chainA_s001 138 QDVNNKRIYFAPNTGNAVANTEIINGKLQGRDANGNQVKNAFSKDVAGNTFYFDANGVMLTGLQTISGKT 207
5ngy_chainA_p002 137 QQVNGKNIYFDAQTGYARQNVGFLDGTAKGFDEQGNQIKSGIATDLSGNVYYFDASGKMLTGVQNIDGKK 206
3klk_chainA_p006 3 --------YIDPTTGQPRKNFLLQNG---------------------------------------------- 20
4amc_chainA_p005 3 --------YIDPTTGQPRKNFLLQSG---------------------------------------------- 20
3aib_chainB_p003 ------------------------------------------------------------------------
3ttq_chainA_p004 96 AKYDDKLFYFESGKGNLVSTVAGDYQDGHYISQD-------------GQTRYADKQNQLVKGLVTVNGAL 152
Consensus_aa: ........Yh-s.pG...psh.....................................................
Consensus_ss: eee eeeee eeeeeeeee eeee eeee eeeeeeee ee

Conservation: 6 996 6 6 6 6 66 9 6 6 9 6 96
6HVG.pdb_chainA_s001 208 YYLD-EQGHLRKNYAGTFNNQFMYFDADTGAGKTAIEYQFDQGLVSQSNENTPHNAAKSYDKSS-FENVD 275
5ngy_chainA_p002 207 YYFD-EQGHRRRNYAGVFNNEFIYFGLD-GVGQSAIEYQFEKGLTSQNSVATSHNAAKSYDTKS-FTNVD 273
3klk_chainA_p006 21 --------------------NDWIYFDKDTGAGTNALKLQFDKGTISADEQYRRGNEAYSYDDKS-IENVN 70
4amc_chainA_p005 21 --------------------NNWIYFDSDTGVGTNALELQFAKGTVSSNEQYRNGNAAYSYDDKS-IENVN 70
3aib_chainB_p003 1 ----------------------------------------------------AQYNQVYSTDAAN-FEHVD 18
3ttq_chainA_p004 153 QYFDNATGNQIKNQQVIVDGKTYYFDDKGNGEYLFTN-TLDM---STNAFSTKNVAFNHHDSSSFDHTVD 218
Consensus_aa: ...................sphhYFs.c.ss...hhp.ph-b...Sps...pp.N.hbS@DspS.hcsVs
Consensus_ss: eee eee eeeee eeee eeeee eeeeeeee eeee

Conservation: 66999 699996 99 69669 96669699666699966 6966999696 6
6HVG.pdb_chainA_s001 276 GYLTADTWYRPTDILKNGDTWTASTETDMRPLLMTWWPDKQTQANYLNFMSSKG---LG--ITTTYTAAT 340
5ngy_chainA_p002 274 GFLTANSWYRPTDILRNGTKWEPSTETDFRPLLMTWWPDKEVQANYLNYMSALG---LG--DQKIYTGAS 338
3klk_chainA_p006 71 GYLTADTWYRPKQILKDGTTWTDSKETDMRPILMVWWPNTVTQAYYLNYMKQYGNLLPA--SLPSFSTDA 138
4amc_chainA_p005 71 GYLTADTWYRPKQILKDGTTWTDSKETDMRPILMVWWPNTLTQAYYLNYMKQHGNLLPS--ALPFFNADA 138
3aib_chainB_p003 19 HYLTAESWYRP------------------------------------------------------------ 29
3ttq_chainA_p004 219 GFLTADTWYRPKSILANGTTWRDSTDKDMRPLITVWWPNKNVQVNYLNFMKAN----GLLTTAAQYTLHS 284
Consensus_aa: G@LTA-oWYRPppIL.sGTpWpsSp-pDhRPllhhWWPsp.hQh.YLN@Mp.........s...@o..t
Consensus_ss: eeee eee eee eeeee eeeeee eeee
```


Motif V
Motif II
Motif III

Motif IV

```
Conservation: 6 6 66 6 6 9696666999 9 6 66
6HVG.pdb_chainA_s001 589 DKNRPNSGT-----LVNR-IANSGDTKVIPNYSFVRAHDYDLKDLRKAM----------IDHGIIKNMQ 642
5ngy_chainA_p002 589 -KSFTKDAL-----LVHR-TADNSYNQAVPNYSFIRAHDSEVQTIIAKIISDKHPDLYPTVDKALLAKDS 651
3klk_chainA_p006 363 --NKLITQS-----LVNR-ANDNTENAVIPSYNFVRAHDSNAQDQIRQAI------------QAATGKPY 412
4amc_chainA_p005 359 --DAIVHQS-----LADR-ENNSTENVVIPNYSFVRAHDNNSQDQIQNAI------------RDVTGKDY 408
3aib_chainB_p003 220 --NPLITNS-----LVNR-TDDNAETAAVPSYSFIRAHDSEVQDLIRDII------------KAEINPNV 269
3ttq_chainA_p004 500 --TNMLQDVDGGTLITDHTTQNSTENQATPNYSIIHAHDKGVQEKVGAAI------------TDATGADW 555
Consensus_aa: ..s.h.p.s.....LhsR.hssss-s.hlPsYSFlRAHDpphQ-bI.phI............pshhs.s.
Consensus_ss: ee eeeee hhhhhhhhhh h
```

Motif VII

```
Conservation: 6 6 9 6 9666 69 6 96 699996669969999 6 6 96996 6
6HVG.pdb_chainA_s001 643 ---DTFTFDQLAQGMEFYYKDQENPSGFKKYNDYNLPSAYAMLLTNKDTVPRVYYGDMYLEGGQYMEKGT 709
5ngy_chainA_p002 652 ---ALYD-----EAFTEYNADMQKISSQKQYTHNNMPSAYAILLTNKDTVPRVYYGDLFTDNGEYMANKT 713
3klk_chainA_p006 413 ---GEFNLDDEKKGMEAYINDQN--STNKKWNLYNMPSAYTILLTNKDSVPRVYYGDLYQDGGQYMEHKT 477
4amc_chainA_p005 409 ---HTFTFEDEQKGIDAYIQDQN--STVKKYNLYNIPASYAILLTNKDTIPRVYYGDLYTDGGQYMEHQT 473
3aib_chainB_p003 270 VGYSFTME-EIKKAFEIYNKDLL--ATEKKYTHYNTALSYALLLTNKSSVPRVYYGDMFTDDGQYMAHKT 336
3ttq_chainA_p004 556 ---TNFTDEQLKAGLELFYKDQR--ATNKKYNSYNIPSIYALMLTNKDTVPRMYYGDMYQDDGQYMANKS 620
Consensus_aa: ...s.@s..pbppth-hY.pDbp..to.KKYs.YNhPttYAlLLTNKDoVPRVYYGDh@pDsGQYM.ppT
Consensus_ss: hhhhhhhhhhhhhhhh h hhhhhhhhhh eeee
```

```
Conservation: 6 6 966 9 669 6999 6 66699699666 6 66 6 666 9
6HVG.pdb_chainA_s001 710 IYNPVISALLKARIKYV-SGGQTMATDSSGKDLKDGETDLLTSVRFGKGIMTSDQTTTQDNSQDYKNQGI 778
5ngy_chainA_p002 714 PYYDAITSLLTARTKFV-SGGQSLSVD---------KNDVLTSVRYGKGALSA----TDNGSSDTRNQGI 769
3klk_chainA_p006 478 RYFDTITNLLKTRVKYV-AGGQTMSVD---------KNGILTNVRFGKGAMNA----TDTGTDETRTEGI 533
4amc_chainA_p005 474 RYYDTLTNLLKSRVKYV-AGGQSMQTMSVGG-----NNNILTSVRYGKGAMTA----TDTGTDETRTQGI 533
3aib_chainB_p003 337 INYEAIETLLKARIKYV-SGGQAMRNQQVGN------SEIITSVRYGKGALKA----TDTGDRITRTSGV 395
3ttq_chainA_p004 621 IYYDALVSLMTARKSYVSSGGQTMSVD---------NHGLLKSVRFGKDAMTA----NDLGTSATRTEGL 677
Consensus_aa: .Y@-hlssLLptRhKYV.tGGQoMphp.........psslTSVR@GKGAhsA....TDsGop.TRspGl
Consensus_ss: hhhhhhhhhhhhhh eeeee eeeeeee eee
```

```
Conservation: 696 69 6 9 9 666 99 69669 996 966 9 6 9 66 96 9 9 6
6HVG.pdb_chainA_s001 779 GVIVGNNPDLKLNNDKTITLHMGKAHKNQLYRALVLSNDSGIDVYDSDDKAP---TLRTNDNGDLIFHKT 845
5ngy_chainA_p002 770 GVIVSNNPNLDLNNDK-VTLSMGISHAHQAYRPLLLTNSQGIVAYATDSEVPQNLYKTTNDKGELTFDAS 838
3klk_chainA_p006 534 GVVISNNTNLKLNDGESVVLHMGAAHKNQKYRAVILTTEDGVKNYTNDTDAP---VAYTDANGDLHFTNT 600
4amc_chainA_p005 534 GVVVSNTPNLKLGANDKVVLHMGAAHKNQQYRAAVLTTTDGVINYTSDQGAP---VAMTDENGDLYLSSH 600
3aib_chainB_p003 396 VVIEGNNPSLRLKASDRVVVNMGAAHKNQAYRPLLLTTDNGIKAYHSDQEAAGL-VRYTNDRGELIFTAA 464
3ttq_chainA_p004 678 GVIIGNDPKLQLNDSDKVTLDMGAAHKNQKYRAVILTTRDGLATFNSDQ-AP---TAWTNDQGTLTFSNQ 743
Consensus_aa: GVlltNsPsLcLssscpVhLpMGhAHKNQ.YRsllLTspsGl.sYsoDpcAP...h.hTs-pG-LhFssp
Consensus_ss: eeee eeeeee eeeeee eeee eeee eeee
```

Figure S9: Structural alignment of the sucrose-active GH70 enzymes with available 3D structure. Alignment created with PROMALS3D (Pei and Grishin, 2014).

Red: domain V, yellow: domain IV, green: domain B, blue: domain A, purple: domain C. Star: catalytic residues. Black boxed residues: residues targeted in this study. Pink boxed residues: conserved motifs.

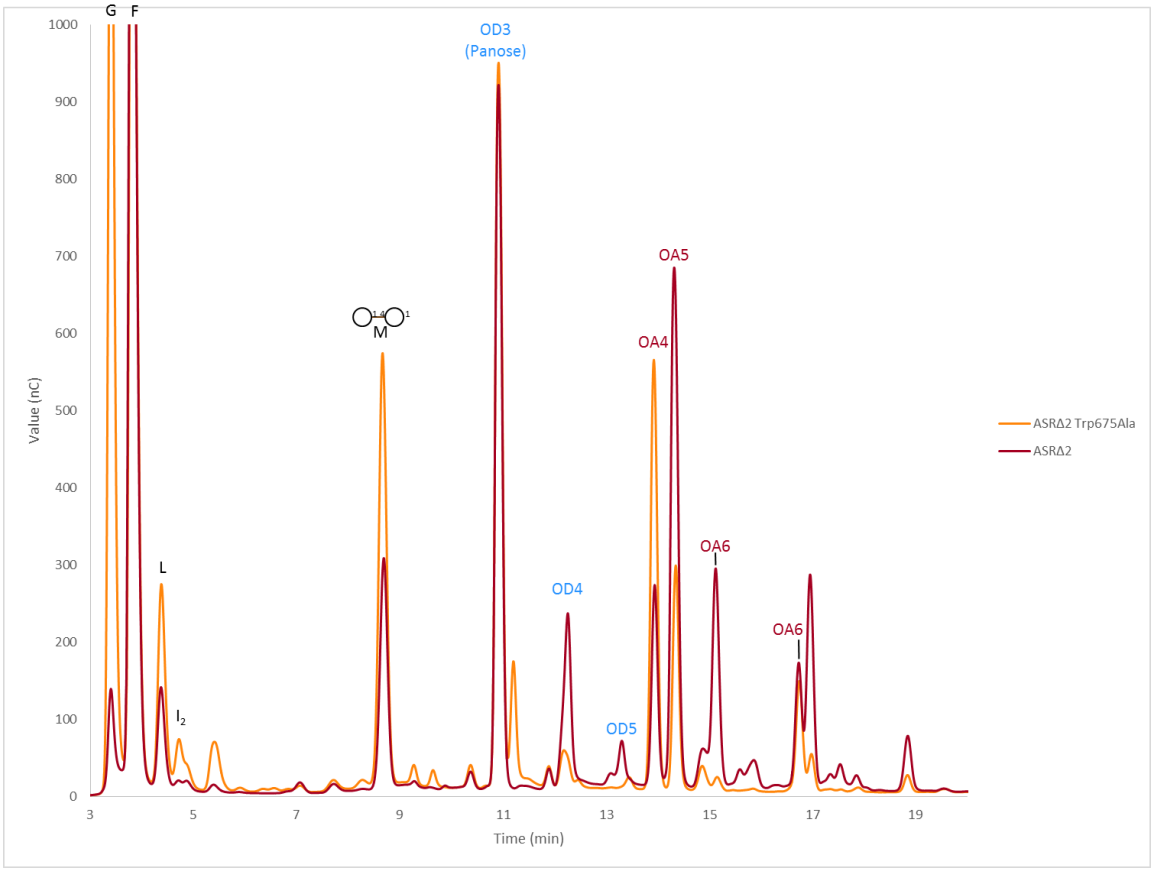


Figure S10: HPAEC-PAD chromatograms of Trp675Ala mutant compared to the wild type ASRΔ2

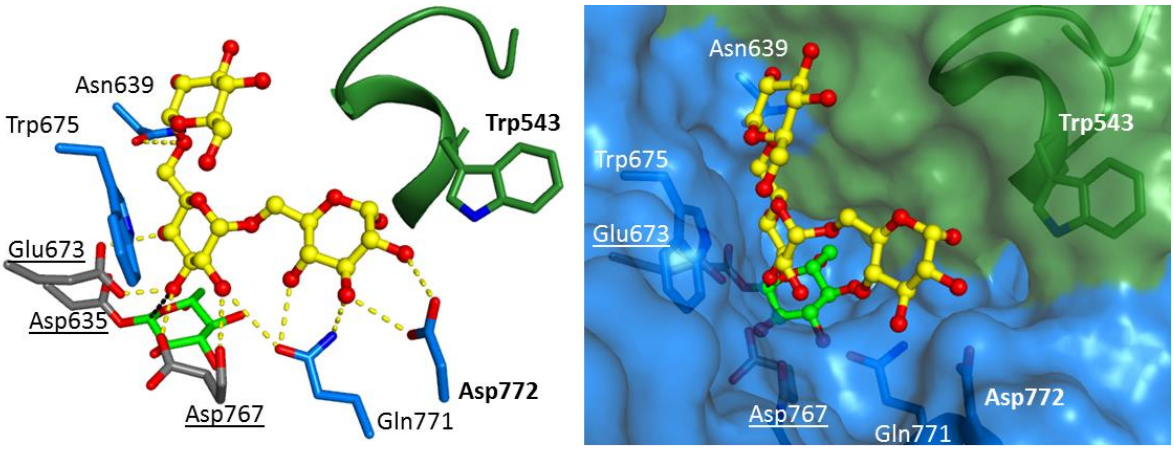


Figure S11: Isomaltotriose (IM3) is automatically docked to illustrate α-1,3 branching glucosylation.
The glucose at the reducing end is positioned in +2' subsite and the one at the non-reducing end +2 subsite.

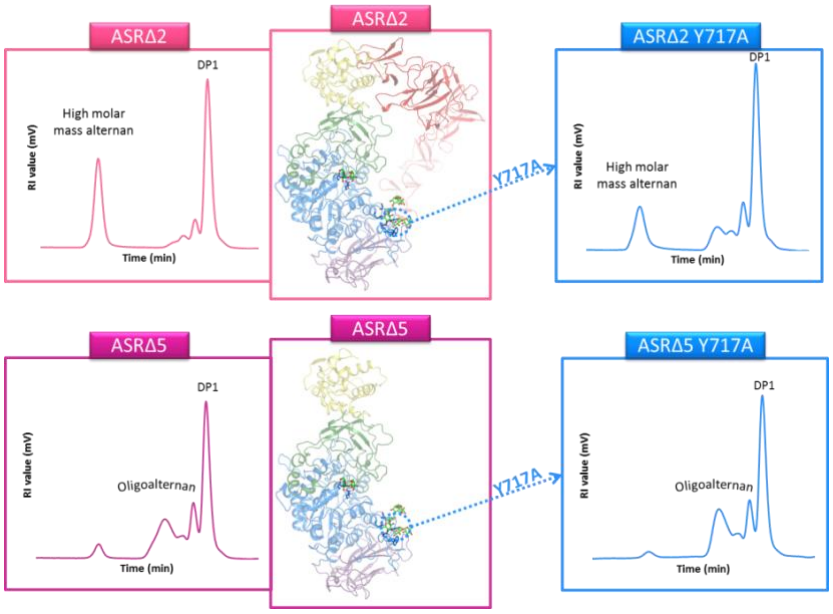
Chapter III:

A specific oligosaccharide binding site in domain A of alternansucrase is involved in alternan elongation

Abstract

Microbial polymers, and particularly α-glucans produced by glucansucrases from GH70 family, are gaining in importance due to their smooth conditions of synthesis from sucrose and their high range of current and potential applications. Focusing on the alternansucrase (ASR), an old-known glucansucrase catalyzing the synthesis of both high molar mass (>1,700,000 g.mol-1) and low molar mass alternans (1,300 g.mol-1), we describe here five crystal structures of a truncated version of alternansucrase (ASRΔ2) in complex with different oligosaccharides (isomaltose, isomaltotriose, isomaltononaose, panose and oligoalternan). Structural analysis of the complexes pinpointed residues in interaction with these various sugars in both the domains A and V of ASR. More precisely, the sugar binding site A1 (SBS-A1), unique to ASR, was identified in domain A relatively remote from the catalytic core. The biochemical characterization of nine single mutants and seven double mutants indicated that the ASR domain V has a role for HMM alternan formation and binding. The involvement of the SBS-A1 subsite in HMM alternan formation was also highlighted and the participation of residues Gln700 and Tyr717 was demonstrated. Indeed, the mutation of these residues to alanine dropped by a third the HMM alternan yield without significantly impacting the enzyme stability or specificity. We suggest the SBS-A1 to be an anchoring region, which could act as a bridge between the active site and the domain V sugar binding pockets to contribute to a processive elongation of the alternan chain.

Keywords: glucansucrase, crystal structure, high molar mass, alternan, processive, GH70



In preparation

Introduction

Since the last decades, microbial polysaccharides have gained attention as promising bio-sourced polymers that can be regularly supplied and are less sensitive to market and climate fluctuations than plant polymers. In addition, progresses in structure-function studies and engineering of polymerases allow today to better control their structures and by extension their physico-chemical properties. In this field, the glucansucrases are very interesting enzymes. These α-transglucosylases catalyze the formation of high molar mass (HMM) homopolysaccharides of D-glucosyl units from sucrose, a low-cost and abundant substrate. The panel of polymers produced by glucansucrases varies a lot in terms of size, type and arrangement of α-osidic linkages and degree of branching, all these features defining polymer structural properties and consequently, the range of ongoing or potential applications. An illustration of such a broad polymer diversity is found in the pioneer work of Jeanes *et al.* who isolated and characterized the α-glucans produced by 96 strains of Lactic Acid Bacteria (Jeanes et al., 1954). One of these polymers distinguished from the others by its intriguing structure made of alternating α-1,3 and α-1,6 linkages in the main chain and was thus renamed « alternan » (Côté and Robyt, 1982a). This polymer was notably produced by *Leuconostoc mesenteroides* NRRL B-1355 strain, recently reclassified as *Ln. citreum* (Bounaix et al., 2010). The enzymatic activity responsible for alternan production is the alternansucrase (ASR). From sucrose, ASR produces a bi-modal population of α-glucans, comprising a High Molar Mass (HMM) alternan of around 1,700,000 $g.mol^{-1}$ and a Low Molar Mass (LMM) alternan fraction of 1,300 $g.mol^{-1}$ as estimated by Size Exclusion Chromatography (Joucla et al., 2006). HMM alternan is more soluble in water and less viscous than dextrans, making it a good substitute of gum arabic (Côté, 1992; Leathers et al., 2009). ASR is also a promising tool for the production of prebiotics from sugar acceptor reaction (Côté et al., 2003; Hernandez-Hernandez et al., 2011; Holt et al., 2005; Sanz et al., 2005a).

The *asr* gene from *Ln. citreum* NRRL B-1355 was cloned and expressed in *E. coli* (Argüello-Morales et al., 2000b). Sequence analysis revealed that ASR belongs to the family 70 of glycoside hydrolases (GH70) according to the CAZy classification (Lombard et al., 2014). Residues Asp635, Glu673 and Asp767 were proposed to respectively play the role of the nucleophile, acid/base catalyst and transition state stabilizer (TSS) for the formation of the β-D-glucosyl intermediate (Argüello-Morales et al., 2000b) likely to the other glucansucrases that all use the classical Koshland α-retaining mechanism (Leemhuis et al., 2013). This was recently confirmed by the 3D structure resolution of a truncated variant of ASR, named ASRΔ2, and comprising residues 39 to 1425 (PDB ID: 6HVG). Notably, the deletion of the signal peptide at the N-terminal extremity and seven APY repeats at the

C-terminal end of the protein did not affect significantly the enzyme specificity in terms of linkage type and polymer size distribution (Joucla et al., 2006; Molina et al., 2019). Similar results were obtained with the alternansucrase of *L. citreum* ABK-1 deleted of APY repeats (Wangpaiboon et al., 2019). A previous kinetic study showed that both HMM and LMM alternan populations were formed during the early stage of the reaction, suggesting that ASR follows a semi-processive mechanism of polymerization involving polymer anchoring regions in the protein to facilitate HMM polymer formation (Moulis et al., 2006).

The recently solved 3D structure of ASR showed that the protein adopts a U-shaped fold made up by five domains (A, B, C IV, V) like the other GH70 sucrose-active enzymes (Bai et al., 2017; Brison et al., 2016; Claverie et al., 2017; Ito et al., 2011; Pijning et al., 2012; Vujičić-Žagar et al., 2010). Mutagenesis study combined to molecular docking highlighted the presence of two different acceptor subsites in the extension of subsites -1 and +1. The first one is defined by Trp675 (subsite +2) and directs glucosylation towards α-1,6 linkage formation whereas the second one defined by Asp772 (subsite +2') is critical for the α-1,3 glucosylation and alternance of α-1,3 or α-1,6 linkages. Bringing together these results, the mechanism which leads to the formation of alternan polymer was suggested to be governed by the differentiation of the terminal glucosidic linkage of the incoming acceptor in the catalytic cleft (Molina et al., 2019). We have also shown that several residues of the active site pocket were critical for polymer formation. In particular, the mutations of the conserved tryptophan in motif III (domain A, Trp675) resulted in a 82% decrease of HMM polymer synthesis with a concomitant increase of hydrolysis reaction (Molina et al., 2019), similarly to what was obtained for the corresponding mutants of Trp1065 in GTF180 and Trp717 in DSR-M (Claverie et al., 2019a; Meng et al., 2017). Additionally, a unique anchoring region more distant from the catalytic core (subsite +3') and involving the residue Trp543 was identified in the loop B1 of domain B, its replacement with an Ala residue resulted in a 54% loss of HMM polymer (Molina et al., 2019). This region corresponds to the one formed by residues from the H1/H2 subdomain described to be in interaction with a glucan chain in GTF-SI according to MD simulations (Osorio et al., 2019). Four putative sugar binding pockets (V-A, V-B, V-C and V-D), homologous to those found in GBD-CD2 and DSR-M GH70 enzymes (Brison et al., 2016; Claverie et al., 2017), were also identified in the domain V of ASR. Notably, the deletion of the entire domain V did not completely abolished polymer formation but reduced it by 86% (without increase of hydrolysis) (Molina et al., 2019).

The control of HMM polymer formation by ASR is thus complex. In this process, the contribution of the sugar binding pockets of domain V is strongly suspected and could be comparable to those described for DSR-M or DSR-OK (Claverie et al., 2017, 2019b). Other determinants located in the domain A, B or IV could also be important. In other polymerases like DSR-M and DSR-S, residues near

the catalytic core (subsite +3) such as Trp624 and Phe504 located in loop B2 of DSR-M and DSR-S respectively (domain B) were shown to be involved in polymer elongation (Claverie et al., 2019a; Irague et al., 2013). Similar or different residues could play a comparable role in ASR. Their identification and location are highly challenged by the difficulty encountered for obtaining complexes with oligosaccharides.

To get deeper insight in the mechanism of HMM polymer formation, we placed our effort in the resolution of several ASR complexes. To this end, ASR crystals were soaked with different sugars varying in terms of degree of polymerization (DPs) and osidic linkages (panose, nigerose, isomaltooligosaccharides of DP 2 to 12, oligoalternans). Several complexes were obtained with different types of oligosaccharides bound in the domains A or V. Combined with mutant characterization, the study highlighted the contribution of the domain V in HMM alternan formation and shed light on the importance of a new sugar binding site, relatively remote from the catalytic core (*i.e.* not corresponding to +1, +2 or +3 subsites), defined by residues Tyr717 and Gln700 of domain A.

Results

New complexes and identification of two sugar binding sites: Pocket V-B in domain V and SBS-A1 in domain A

To get new insight in alternansucrase structure-function study, ASRΔ2 was soaked with different acceptors varying in size and structures. The idea was to obtain complexes allowing the identification of sugar binding sites close to or remote from the active site and investigate their possible role. The following carbohydrate molecules were tested: isomaltose (I2), nigerose, isomaltotriose (I3), panose, isomaltononaose (I9), isomaltododecaose (I12) and a mixture of oligoalternans (OA) of DP around 9 that did not contain α-1,6 linked oligodextrans (Figure S2). Of note, in all the complexes, the glucosyl units of each oligosaccharide will be numbered by ascending order from their reducing end.

For the complexes ASRΔ2:I2, ASRΔ2:I3, ASRΔ2:panose and ASRΔ2:OA, we unexpectedly observed a clear density in the domain A, in a site that we named sugar binding site A1 (SBS-A1) and that was never described before in GH70 enzymes (Figure 1). For all the ligands, another clear density was obtained in the domain V at a position corresponding to the pocket V-B (Gly234-Thr304) (Figure 1) previously proposed to be a putative sugar binding pocket (Molina et al., 2019). One, two or a maximum of three glucosyl units were visible in this pocket depending on the complex (Table 2). In contrast, no electron density was found in the pocket referred as V-A, also previously predicted to be a sugar-binding pocket.

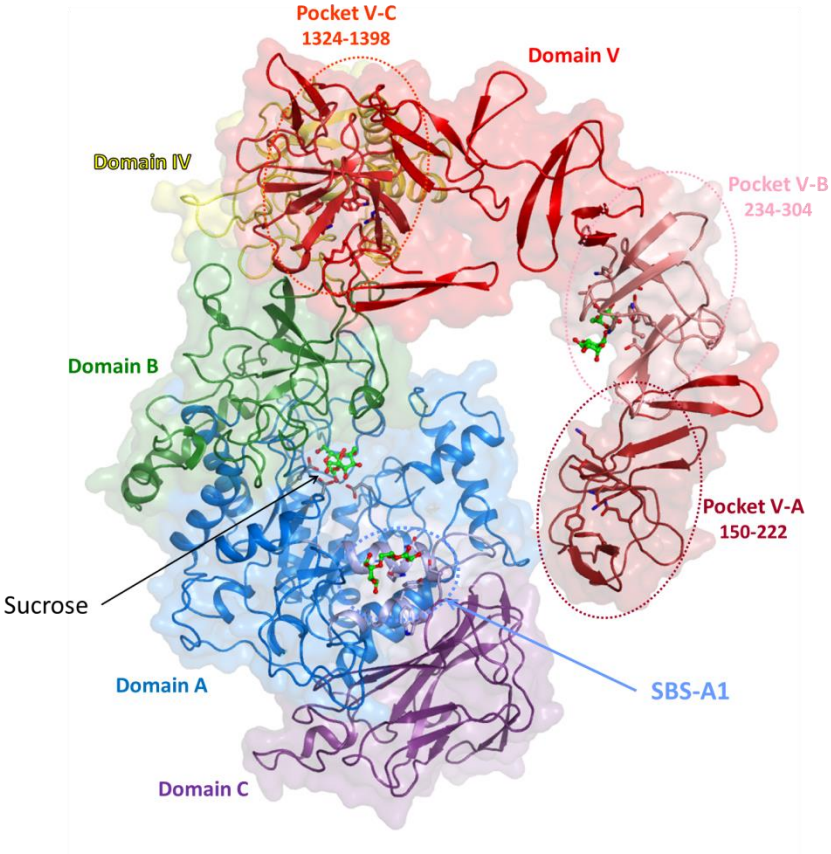


Figure 1: ASR sugar binding sites identified in ASR:I2 complex
Sucrose was manually docked from GTF180:sucrose complex (PDB ID: 3HZ3).

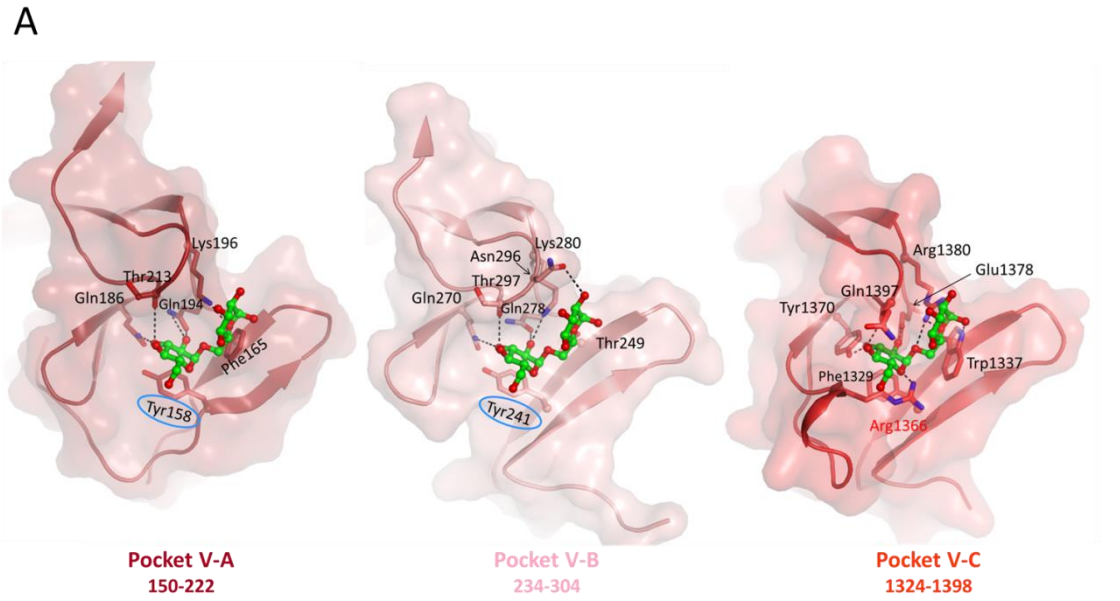
Table 2: Density observed with each ligand.
See Figure S3 for electron density maps in domain V

Ligand	Structure	Domain A: SBS-A1		Domain V: Pocket V-B	
		Density?	Nb. of glucosyl residues	Density?	Nb. of glucosyl residues
I2	α-D-Glc*p*-(1→6)-D-Glc	YES	2	YES	2
I3	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc)	YES	3	YES	3
Panose	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc	YES	3	YES	3
Nigerose	α-D-Glc*p*-(1→3)-D-Glc	No		YES	1
I9	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc	No		YES	3
I12	α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc	No		YES	*n.d.*
OA	*DP~10*	YES	6	YES	3

Sugar binding pockets description and comparison

We used the I2 complex for which we obtained one of the highest resolution to model the isomaltose bound in pocket V-B (Figure 1, Table 1). Our model was based on the high resolution crystal structure of ΔN_{123}-GBD-CD2 in complex with I3 (1.85Å, PDB ID: 4TVC) (Brison et al., 2016). The Glcp_2 is in CH-π stacking interaction with Tyr241 (ASR numbering) and forms hydrogen bonds with Gln278 and Gln270 through O2 and O3 hydroxyls, respectively, and with the main chain oxygen of Thr297 through O3 (Figure 2A and S3). The Glcp_1 mainly interacts with Thr249 through its O5 and with Lys280 through O6 and O5 hydroxyls. Interestingly, the other complexes obtained in pocket V-B (I3, I9, OA) show similar interaction networks despite different linkage specificities (Figure S3). To note, two protein monomers are arranged in the crystal around a pseudo 2-fold axis in the asymmetric unit and the domain V of each unit is intertwined with its equivalent of the second chain. For example in the I9 complex, the oligosaccharide is found at the interface of the two pockets V-B of each chain, probably maintaining the domains V stacked together during the crystallization process (Figure S4).

An isomaltose molecule was manually docked in pocket V-A and V-C of ASRΔ2 using the model from pocket V-B (Figure 2A). The structural comparison of pockets V-A and V-B revealed similarities. In particular, Gln270, Tyr241 and Lys280 of pocket V-B are well aligned with Gln186, Tyr158 and Lys196 of pocket V-A whereas different residues are found at these positions in pocket V-C (Figure 2B). There are however subtle differences between pocket A and B as the distance between residues is not exactly the same (Figure S5). To note, we did not observe any clear electron density in pocket V-A in our complexes with the exception of I3 and I2 where a small blob of density could be observed but too weak to place a ligand. The absence of oligosaccharides could be due to the effect of crystal packing and/or to the fact that pocket V-A is less exposed and accessible than pocket V-B. In dextransucrases DSR-E and DSR-M, the QxK motif associated with a Tyr residue at the bottom of the pocket was suggested to be a signature of sugar binding pocket functionality (Brison et al., 2016; Claverie et al., 2017). On this basis, we assumed the pockets V-A and V-B to be functional. In contrast, the pocket V-C is likely to be non-functional due to the absence of the QxK motif (replaced by ^{1378}ELR^{1380}) and the presence of a small loop, ^{1364}NTR^{1366}, from which Arg1366 emerges, blocking the access to the pocket (Figure 2A).



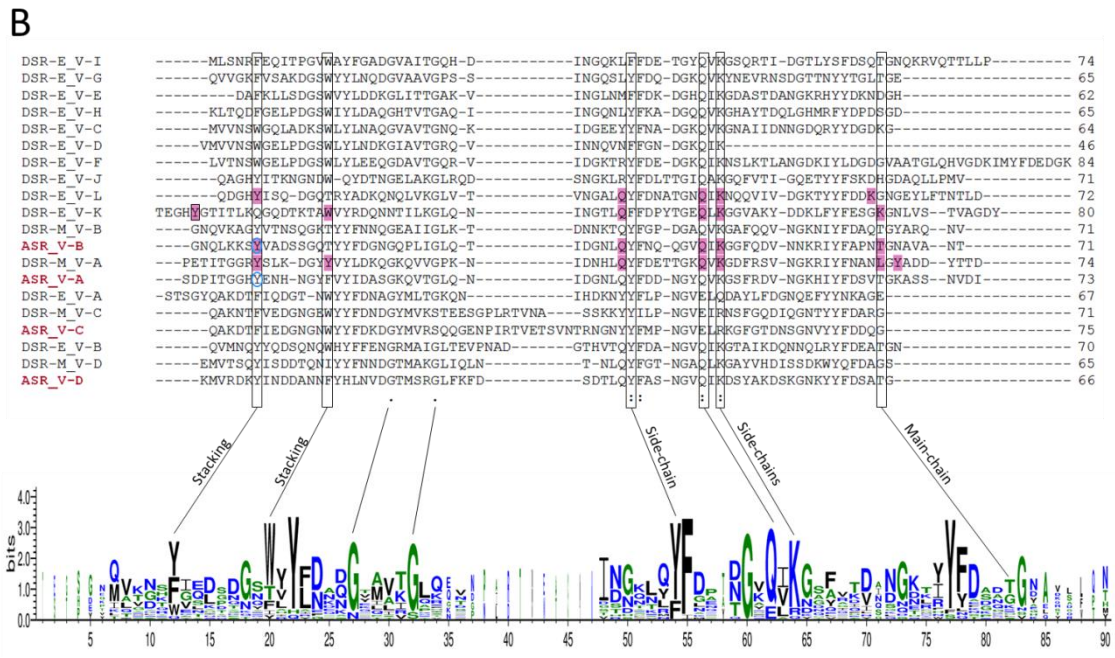


Figure 2: (A) Isomaltose binding in pocket V-B and manual docking in pockets V-A and V-C
The residues interacting with the glucosyl rings are shown in sticks.
(B) Sequence alignment of the sugar binding pockets identified in ASR, DSR-M and DSR-E (Brison et al., 2016; Claverie et al., 2017; Molina et al., 2019).
Pink highlighted residues were shown to directly interact with sugar ligands in 3D structures. Blue circled residues have been mutated to Ala in this study.

Mutation of the conserved tyrosine of sugar binding pockets

To further investigate the role of V-A and V-B pockets, the central stacking residues, Tyr158 and Tyr241 respectively, were replaced by an alanine. The mutations resulted in a slight decrease of the HMM polymer yield, from 31.5% for the wild type to 27.2%, 28% and 27.6% for the Tyr158Ala mutant, the Tyr241Ala mutant and the Tyr158Ala+Tyr241Ala double mutant, respectively (Figure 3A, Table 3). Enzyme specific activity, specificity and melting temperature were not significantly affected for these three mutants (Table 3). The mutations were also performed combined to the Asp772Ala mutation, found to be critical for α-1,3 formation in our previous study (more than 95% of α-1,6 linkages formed) (Molina et al, 2019). The HMM polymer yield decreases from 3.2% for the Asp772Ala single mutant to 2.1%, 1.7% and 0.9% for the Asp772Ala+Tyr158Ala double mutant, the Asp772Ala+Tyr241Ala double mutant and the Asp772Ala single mutant devoted of the whole domain V (ASRΔ5 Asp772Ala), respectively (Figure 3B, Table 3). Compared to the Asp772Ala mutant, enzyme melting temperature was not affected whereas relative activity decreased from 36.5% to 23.7% and 26.1% for the Asp772Ala+Tyr158Ala and Asp772Ala+Tyr241Ala double mutants respectively (Table 3).

To explore the affinity of the enzyme with glucan, we performed affinity gel electrophoresis of ASRΔ2, DSR-MΔ1 and ΔN_{123}-GBD-CD2. The two latter enzymes also possess sugar binding pockets in which oligosaccharides were experimentally shown to bind. Logically, the migration of the three enzymes is delayed in the presence dextran. In contrast, only ASR was slightly retained by the presence of alternan (Figure 4A). Focusing on ASR sugar binding pockets, both single and double mutations Tyr158Ala and Tyr241Ala affected the binding ability of the enzyme with dextran and alternan. In particular, a subtle difference is observed between the migration of Tyr158Ala and Tyr241Ala mutants in presence of alternan as the Tyr241Ala mutant is almost not delayed (Figure 4B).

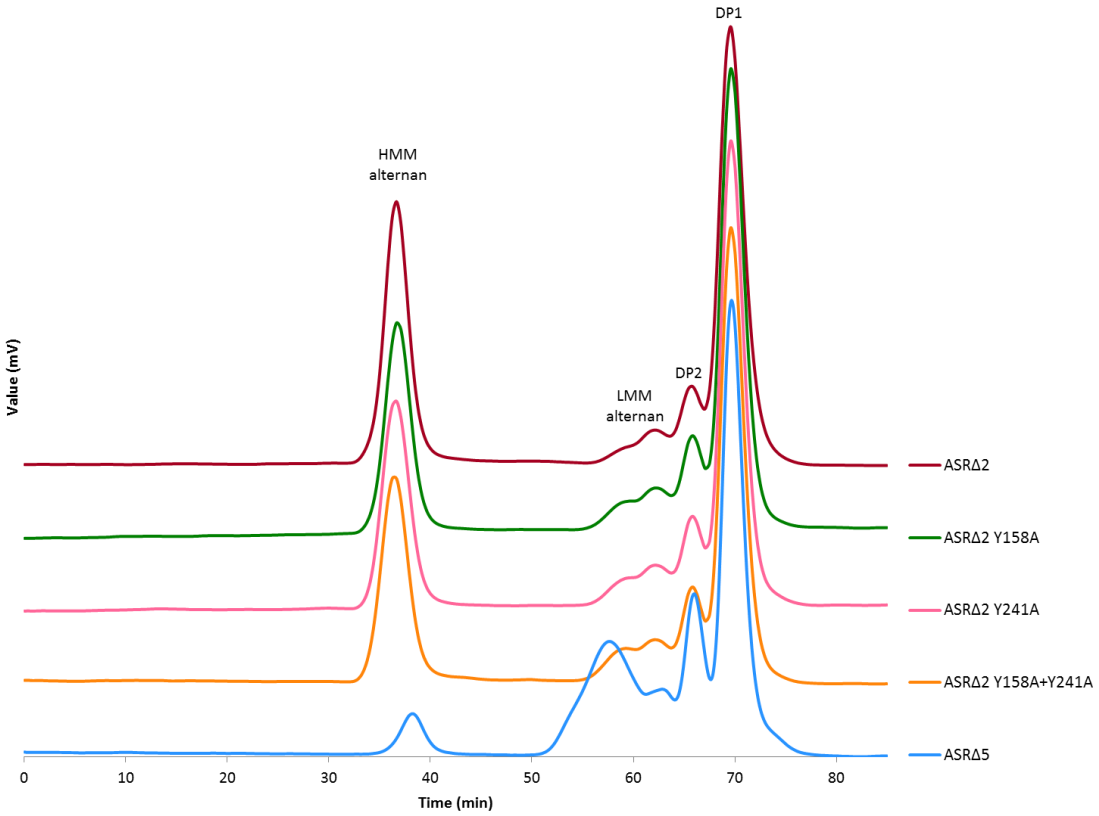


Figure 3: HPSEC chromatograms of the mutants generated in sugar binding pockets
Reaction from sucrose at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75.

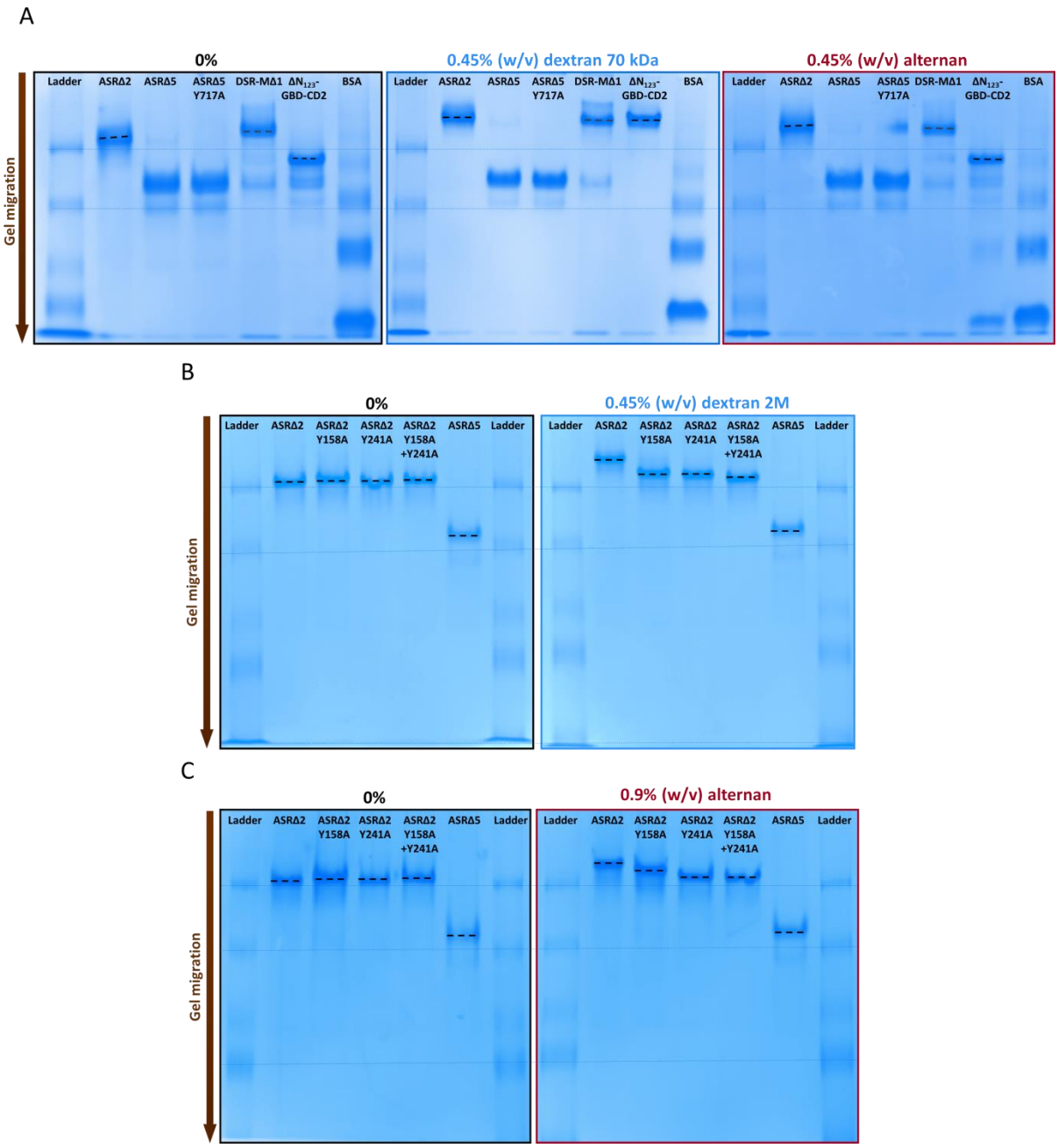


Figure 4: (A) Affinity gel electrophoresis of ASRΔ2, ASRΔ5, ASRΔ5-Tyr717Ala.

DSRM-Δ1 and ΔN_{123}-GBD-CD2 (branching sucrase) are used as a positive control (Brison et al., 2016; Claverie et al., 2017). BSA 1% and Protein standard (ladder) are used as negative controls. Gels were made in the presence or absence of 0.45% (w/v) dextran 70,000 $g.mol^{-1}$ or alternan

(B) and (C) Affinity gel electrophoresis of ASRΔ2, ASRΔ2-Tyr158Ala, ASRΔ2-Tyr241Ala, ASRΔ2-Tyr158Ala+Tyr241Ala and ASRΔ5.

Protein standard (ladder) is used as negative control. Gels were made in the presence or absence of 0.45% (w/v) dextran 2,000,000 $g.mol^{-1}$ or 0.9% (w/v) alternan

Description of the Sugar Binding Site (SBS) identified in the domain A: SBS-A1

The SBS-A1 site was never described before for any other GH70 enzymes. This site is defined by residues from the α-helix 6 of the catalytic $(\beta/\alpha)_8$ barrel and from a long loop specific to ASR and previously referred to as insertion 1 (Trp716 to Arg738) (Molina et al, 2019).

- **Fitting the glucosyl residues in the density map**

To model the various ligands found in SBS-A1, we first searched for the most probable direction of the sugar chain in the binding site. To do so, I2 was fitted in the electron density considering two possible orientations with either Glcp_1 or Glcp_2 in stacking interaction with Tyr717. The most likely model corresponds to Glcp_1 in stacking interaction with Tyr717 with its O1 pointing in the opposite direction of the Tyr717 stacking platform (Figure 5). This positioning allows the preferred interaction between alpha-D sugars and aromatic platforms (Asensio et al., 2013; Hudson et al., 2015). A very similar density was observed around the Tyr717 for both the isomaltotriose (I3) and the panose complexes suggesting that I2, I3 and panose all interact with their isomaltose moiety (Figure 5). Concerning the oligoalternan complex and assuming that this molecule is composed of alternative α-1,6 and α-1,3 linkages, we modeled it as an hexasaccharide and succeeded to fit it in the electron density with the Glcp_3 ring in stacking interaction with Tyr717 (Figure 5). Lastly and interestingly, no electron density could be observed for the longer isomaltooligogaccharides (I6, I9, I12) that have been tested.

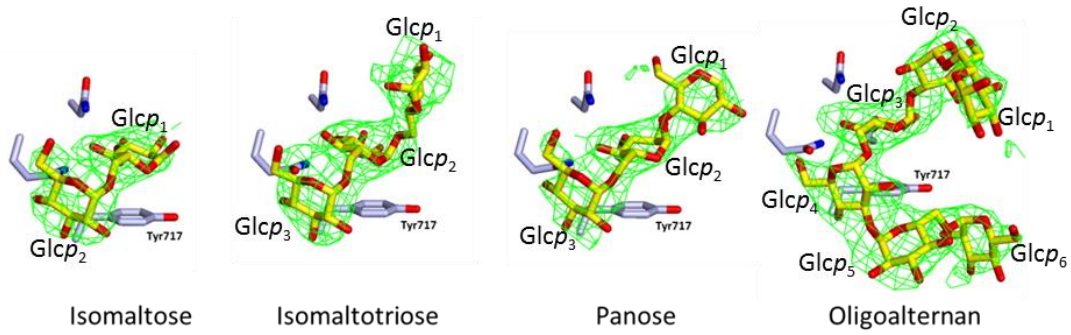


Figure 5: Difference electron density maps in the SBS-A1 site.
The difference electron density map (*Fo – Fc)* around carbohydrates was calculated directly after the molecular replacement and contoured at 2.5σ. Final structures of the oligosaccharides are shown in yellow sticks.

- **Isomaltotriose, isomaltose and panose adopt a similar positioning**

Four residues are found in interaction with I3 (Figure 6A). Tyr717 is in CH-π stacking with Glcp_2 (parallel configuration) and Glcp_3 (T-shaped configuration). In addition, the O3 and O4 of Glcp_3 form a hydrogen bond with the carbonyl of Tyr717 and the side chain of Gln700. The O4 of Glcp_2 forms hydrogen bonds with Gln700 and Asn703, the O3 with Gln700, Ser713 and Asn703 and the O2 with Ser713 only. Glcp_1 is not stabilized by any interactions with the protein. Glcp_1 and Glcp_2 of I2 superimpose well with Glcp_2 and Glcp_3 of I3 and are bound through the same network of interactions (Figure S6). Panose binding involves the same four residues as those described for I2 or I3 complexes. Similarly, Glcp_1 of panose is not maintained by any interactions (Figure 6B). Noteworthy, we also attempted crystal soaking with nigerose but could not obtain any complex with nigerose bound in SBS-A1 (Table 2). Hence, we suggest that at least two α-1,6 linked glucosyl units are required for a correct positioning around Tyr717, one in parallel and the other in T-shaped stacking interaction.

- **Oligoalternan wraps around Tyr717**

The glucosyl units Glcp_2, Glcp_3, Glcp_4 of the DP 6 OA are in the same region as the three glucosyl residues of I3 and panose (Figure 6C). The units Glcp_1 and Glcp_2 are unbound. The O3 of Glcp_3 interacts with Ser713 and Gln700 and the O4 is coordinated with Gln700 and Asn703. There is a CH-π stacking interaction between Glcp_3 and Tyr717 and also a T-shaped stacking interaction with Glcp_4. The latter is also hydrogen bonded with Gln700 through O6 and Tyr717 carbonyl through O4. The unit Glcp_5 interacts with residues not identified previously with its O2 bound to Asp720 and the main chain of Gly722. Finally, the O2 of Glcp_6 interacts only with the hydroxyl group of Tyr717 side chain. For Glcp_5 and Glcp_6, there may also be a parallel-displaced stacking interaction with Trp716. Noteworthy, Tyr717 is in interaction with Glcp_3, Glcp_4 and Glcp_6. The OA literally wraps around this amino acid making it a pillar residue of SBS-A1 active site.

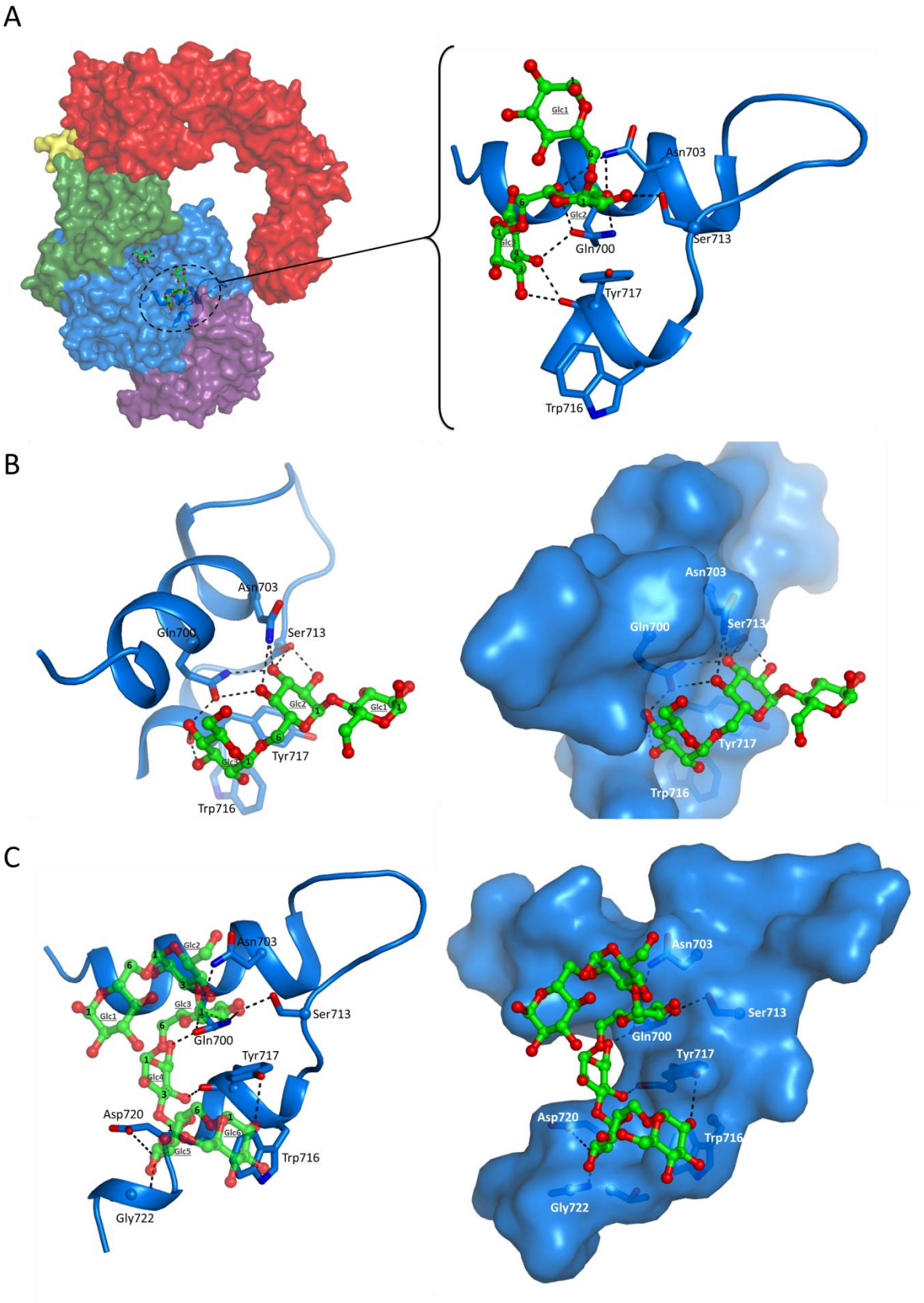


Figure 6: Models of (A) I3, (B) panose and (C) OA binding in sugar binding site A1 based on the I3-bound, panose-bound and OA-bound crystal structures respectively.
Sucrose was manually docked from 3HZ3.

The Sugar Binding Site A1, SBS-A1, is a mediator of HMM alternan formation

The five residues described in interaction with both I3 and OA (Gln700, Asn703, Ser713, Trp716 and Tyr717) were all replaced by an alanine to evaluate their importance for enzyme specificity and polymer size distribution. The mutations did not affect enzyme melting temperature, linkage specificity, or hydrolysis rate (Table 3). The specific activity was also globally well-conserved, all mutants keeping between 76.4% and 95.8% of residual activity compared to ASRΔ2. The product profile was unchanged for the Asn703Ala, Ser713Ala, Trp716Ala mutants (Figure S7). In contrast, the polymerization process was clearly affected by the mutations of Gln700 and Tyr717 confirming the importance of these residues for oligosaccharide binding. Indeed, the amount of HMM polymer decreased from 31.5% for the wild type enzyme to 20.4%, 18.8% and 18.3%, respectively, for the Gln700Ala, Tyr717Ala mutants and the Gln700Ala-Tyr717Ala double mutant (Table 3, Figure 7A). The peak-APEX of HMM polymer formed with the mutants is also slightly displaced towards higher masses. This may reflect a possible variation of the polymer structural organization in water as suggested by DLS assay (data not shown) and could be due to variation of branching length (among other possible reasons) even if the global percentage of α-1,3 linkages was unchanged compared to the wild-type alternan. Additional analyses would be required to investigate the polymer structure in more details and conclude.

HPAEC-PAD analysis of the sucrose reaction products obtained with ASRΔ2 and the Tyr717Ala mutant confirmed that the amount of oligosaccharides formed with the mutant is more abundant, hence corroborating the results of HPSEC (Figure 7B). However, when performing the reaction with maltose acceptor, the chromatograms of the reaction products from DP 2 to DP around 7/8 were perfectly stackable for all the seven mutants, as shown as example for Tyr717Ala mutant (Figure 7C). These products result from maltose glucosylation and correspond to previously characterized oligoalternans with a maltose unit at the reducing end (Côté and Sheng, 2006; Molina et al., 2019).

To assess binding interactions with alternan or glucans, we performed affinity gel electrophoresis of ASRΔ5 or mutant ASRΔ5 Y717A with dextran or alternan to check whether SBS-A1 site could confer affinity for dextran in the absence of domain V. We did not observe any differences between the ASRΔ5 and ASRΔ5-Y717A migration in the presence of dextran or alternan indicating that the contribution of SBS-A1 to polymer binding remains weaker than that conferred by domain V (Figure 4A).

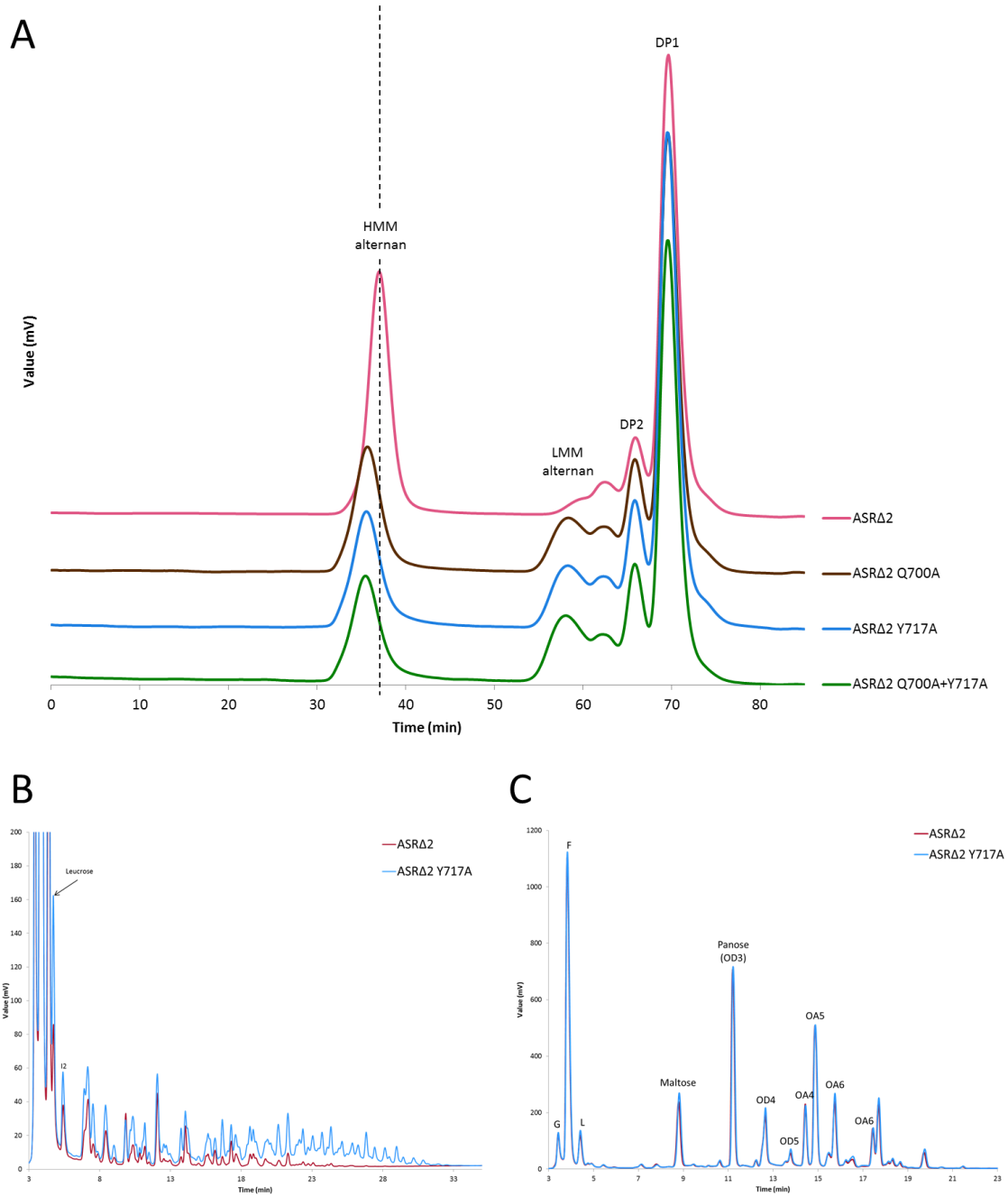


Figure 7: (A) HPSEC chromatograms of SBS-A1 mutants.
(B) HPAEC-PAD chromatogram of the oligoalternans produced from sucrose.
Reaction from sucrose at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75.
(C) HPAEC-PAD chromatogram of the acceptor reaction products from maltose.
Reaction from sucrose and maltose with sucrose:maltose mass ratio 2:1 at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75. G: glucose, F: fructose, L: leucrose. For detailed structures, see Experimental procedures.

The Tyr717Ala mutation was combined with mutations of the stacking residues in the sugar binding pocket of the domain V, Tyr158Ala and Tyr241Ala for pocket V-A and V-B, respectively. The resulting HMM alternan yield was reduced from 18.8% for the Tyr717Ala single mutant to 13.3% and 8.5% for the Tyr158Ala+Tyr717Ala and Tyr241Ala+Tyr717Ala double mutants, respectively (Figure 8A). The mutations did not impact significantly the specificity and residual activity was higher than 60% for both mutants (Table 3). HMM polymer formation starts earlier with ASRΔ2 than with the mutant Tyr717Ala (around 5 minutes and 10 minutes, respectively) and the production rate is almost twice faster (0.50 $g.L^{-1}.min^{-1}$ and 0.27 $g.L^{-1}.min^{-1}$, respectively) (Figure 8B, Figure S8). In parallel, oligoalternans is enhanced with the mutant Tyr717Ala, which in accordance with a reduced formation of HMM polymer (Figure S9). Notably, HMM polymer formation rate is even lower for the double mutants Tyr717Ala+Tyr158Ala mutant (pocket V-A) and Tyr717Ala+Tyr241Ala mutant (pocket V-B). Mutation in pocket V-B affects more the kinetics and yield of HMM polymer formation than that in pocket V-A (Figure 8B).

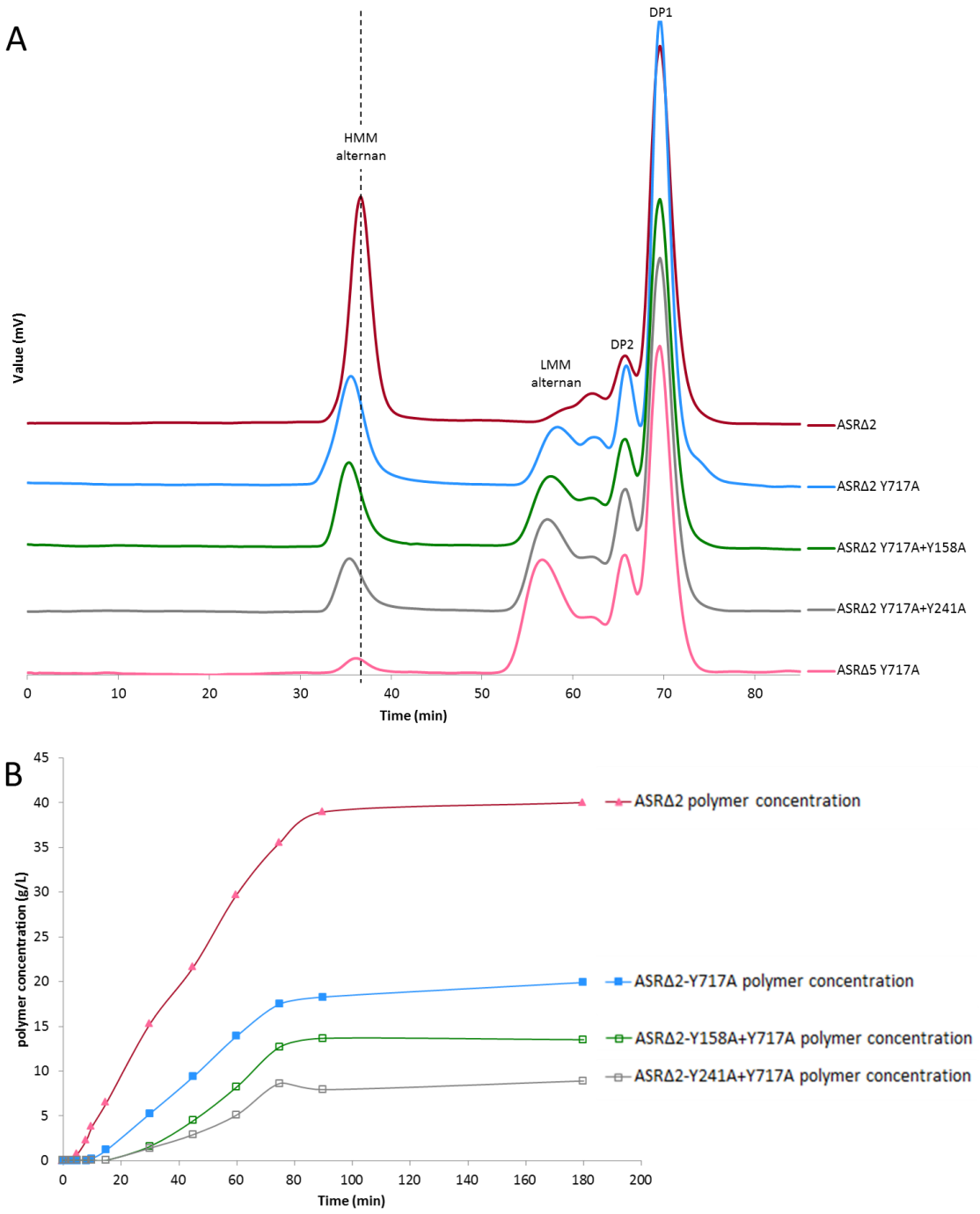


Figure 8: (A) HPSEC chromatograms of alternan populations produced with Tyr717Ala and the double mutants Tyr717Ala+Tyr158Ala mutant (pocket V-A) and Tyr717Ala+Tyr241Ala mutant (pocket V-B).
Reaction from sucrose at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75.
(B) Monitoring of polymer formation with time.
Reaction from sucrose at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75. Production rate was calculated from 5 minutes to 75 minutes (R^2 of 0.997) and from 10 minutes to 75 minutes (R^2 of 0.999) for ASRΔ2 and ASRΔ2 Y717A, respectively.

Table 3: Biochemical data of the characterized mutants

Reaction from sucrose only at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75 Specific activity of ASRΔ2: 30.2 ± 1.0 $U.mg^{-1}$. Specific activity was determined in triplicate. T_m was determined by DSF.

	Residual activity (%)	ΔTm with the wild-type enzyme (°C)	% of α-1,3 linkages (NMR)*	% of α-1,6 linkages (NMR)*	% of polymer formed (area HPSEC)	t_R polymer (min, HPSEC)	Hydrolysis (%)
	Wild-type						
ASRΔ2	100 ± 3.3	0	35	65	31.5 ± 1.6	37 ± 0.2	4.4 ± 0.4
	Mutations in Domain V						
ASRΔ2 Y158A (pocket V-A)	80.2 ± 3.0	+0.1	34	66	27.2	36.8	4.7
ASRΔ2 Y241A (pocket V-B)	78.4 ± 3.2	+0.2	35	65	28.0	36.6	4.7
ASRΔ2 Y158A+Y241A (pockets V-A+V-B)	77.5 ± 2.4	+0.2	33	67	27.6	36.5	4.7
ASRΔ5[1]	79.1 ± 2.7	-2	30	70	4.5	38.2	5.8
	Mutations in SBS-A1						
ASRΔ2 Q700A	80.5 ± 2.8	-0.2	35	65	20.4	35.7	5.6
ASRΔ2 N703A	95.3 ± 6.0	-0.1	35	65	33.5	36.9	4.3
ASRΔ2 S713A	95.8 ± 3.3	-0.3	35	65	31.7	36.6	4.3
ASRΔ2 W716A	76.4 ± 7.0	-1.0	35	65	30.0	36.4	5.0
ASRΔ2 Y717A	80.2 ± 2.4	-0.1	33	67	18.8	35.6	5.2
ASRΔ2 Q700+Y717A	26.1 ± 7.7	-0.9	35	65	18.3	35.5	5.8
	Mutations in Domain V + SBS-A1						
ASRΔ2 Y158A+Y717A	62.9 ± 3.9	-0.1	32	68	13.3	35.3	5.5
ASRΔ2 Y241A+Y717A	78.3 ± 5.8	-0.2	31	69	8.5	35.4	5.9
ASRΔ5 Y717A	63.1 ± 3.6	-2.4	30	70	2.3	36.4	6.2

*: NMR was performed on crude reaction medium.

[1] (Molina et al., 2019)

Discussion

We disclose here six crystal structures of alternansucrase in complex with different oligosaccharides. They are the first complexes ever obtained with this enzyme, a transglucosylase showing a unique linkage specificity among the GH70 glucansucrases. These structures enabled us to locate several oligosaccharide binding sites in the protein. Two of them were located in the domain V and a new site was identified in the domain A. We have investigated their role on both the linkage specificity, stability and ability to synthesize high molar mass alternans. It is important to emphasize on the fact that this enzyme naturally catalyzes the synthesis of both high molar mass and low molar mass alternans. The idea behind this work was to identify structural parameters that could allow the design of alternasucrase strictly specific for either HMM or LMM polymer synthesis in the future. What have we learned?

Involvement of ASR domain V and the sugar binding pocket in HMM alternan formation

For all the complexes we solved, we have found isomaltooligosaccharides and oligoalternans bound in domain V. This domain shares a high percentage of identity with its counterpart exhibited by DSR-M. As for DSR-M we have identified two sugar binding pockets showing the structural traits close to those previously described for DSR-M. Unsurprisingly and similarly to what was observed with DSR-M, we found isomaltooligosaccharides in the pocket B of alternansucrase and signs of electron density in the pocket A and affinity gel electrophoresis revealed that ASRΔ2 binds dextran like DSR-M (REF). The positioning of isomaltooligosaccharide is different in DSR-M pocket V-A with the presence of Tyr238 that could prevent the orientation observed in ASR (Figure 9C). More strickingly, we also found oligoalternans bound in pocket B, and even if the affinity of domain V for alternan is lower compared to that of dextran, affinity gel confirmed the interaction between alternan and the domain V of the enzyme and demonstrates the binding promiscuity of this domain. In contrast, DSR-M was not retained by alternan, showing that despite structural similarities, the sugar binding pockets affinity could be subtly different from one enzyme to another. Lastly, in the ASRΔ2:I9 complex the I9 molecule seems to bind at the interface of the two domains V giving a striking example on how multiple glucansucrase chains can simultaneously bind to a single polymer chain and this is very much likely to take place in solution.

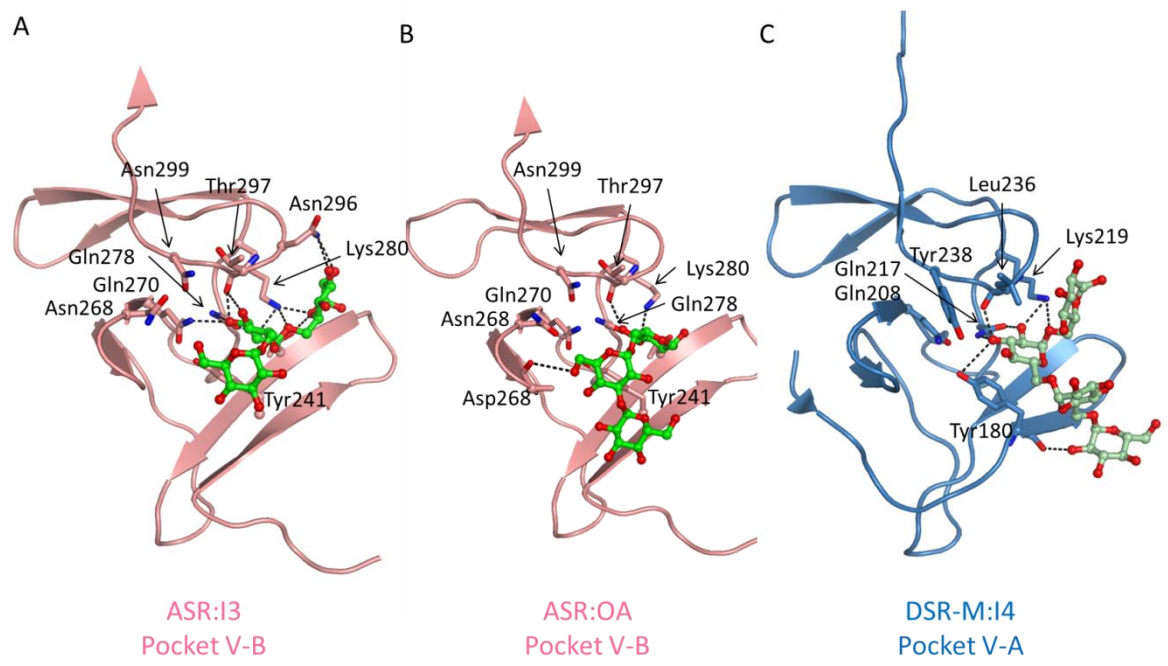


Figure 9: Comparison of (A) ASR:I3, (B) ASR:OA and (C) DSR-M:I4 complexes
Glucose residues are represented in green.

Analysis of the complexes revealed the same interaction involving a QxK motif and a conserved aromatic residue with the oligosaccharides. Changing the aromatic residues in each pocket induces a slight but significant effect on the HMM polymer yield, indicating that these pockets may interact with the polymer and promote its elongation by providing anchoring platforms for long chains. One thing worth to mention is that affinity gel indeed revealed binding with alternan. However, mutations in the pockets were much less detrimental to HMM alternan formation than the deletion of domain V. ASRΔ5 synthesized only 4.5 % of HMM polymer (Molina et al., 2019), versus 31.5 % for ASR-Δ2 and around 27% for the single or double mutants targeting the aromatic residue of the binding pocket V-A and V-B. The mutation of only the conserved aromatic residue of the pockets (Tyr 158 and Tyr 241) may not be sufficient to abolish all interactions, even in the double mutant, and to obtain similar effect to an entire deletion. Furthermore, deletion of the entire domain V may also modify the fold of the other domains compared to the full-length protein and this could also impact the production of HMM. Notably, the Tm of ASR-Δ5 decreased by 2°C. Altogether, our findings show that domain V plays a limited role in the formation of HMM polymer. Additional mutations targeting the conserved Gln and/or Lys of the QxK motif should help to conclude with more confidence on the role of the pockets in domain V.

The Sugar binding site A1, a signature of alternansucrase

Of the various ligands tested in our soaking experiments, I2 and I3 did bind to SBS-A1, nigerose did not as well I9 and I12 (Table 2). A close inspection of I3 complex revealed that, in the proposed configuration, a fourth α-1,6 linked glucosyl unit could not be added to the non-reducing end of I3 due to a steric clash with the protein surface. In contrast, the addition of an α-1,3 linked unit enables the steric clash to be avoided, which also supports the proposed binding mode of the DP6 OA in Figure 6C. In contrast, we have obtained a complex with an oligoalternan of DP9, indicating that the site is well-designed to specifically interact with oligomers containing α-1,6 and α-1,3 linkages above DP4. Two residues appeared to be particularly important: Tyr 717 and Gln700. Their replacement by Ala strongly reduced HMM polymer yield from 31.5 % for ASR-Δ2 to 18. % and 20% for the mutants, showing that there is a contribution of this site to alternan elongation. In agreement with this assumption, mutagenesis of Tyr717 together with the deletion of domain V led to a further decrease of the percentage of HMM polymer produced from 4.5% to 2.3%. Finally, the accommodation and glucosylation of short oligosaccharides (DP 3 to DP 6) are not at all impacted by the mutations operated in SBS-A1 confirming that SBS-A1, which is remote from the active center, comes on stage only when the formed oligosaccharides reach a sufficient length. Furthermore, we propose SBS-A1 binding site as a signature of alternansucrase specificity. Indeed, Gln700 and Tyr717 are conserved in all characterized (Côté and Robyt, 1982a; Wangpaiboon et al., 2018) and putative alternansucrase sequences, identified by BLASTp search (from *Ln. citreum* NRRL B-1501, NRRL B-1498, LBAE-C11, KM20, and EFEL 2700 strains). Furthermore, of 64 characterized GH70 glucansucrases, residue Tyr717 is only found in alternansucrase and is replaced by an arginine (22/64) or a proline residue (20/64) in the other glucansucrases (Figure 10). Finally, a question remains pending: are there connections between SBS-A1 site and the domain V of ASR? To address it, we constructed mutants bearing the Tyr717Ala mutation in SBS-A1 and in each of the pocket: Tyr158Ala+Tyr717Ala (pocket V-A) and Tyr241Ala+Tyr717Ala (pocket V-B). The HMM alternan yield further decreased from 31.5 % to 13.3 and 8.5 %. This was correlated with the slowdown of HMM alternan formation rate and provides evidence of a connection between SBS-A1 and the sugar binding pockets.

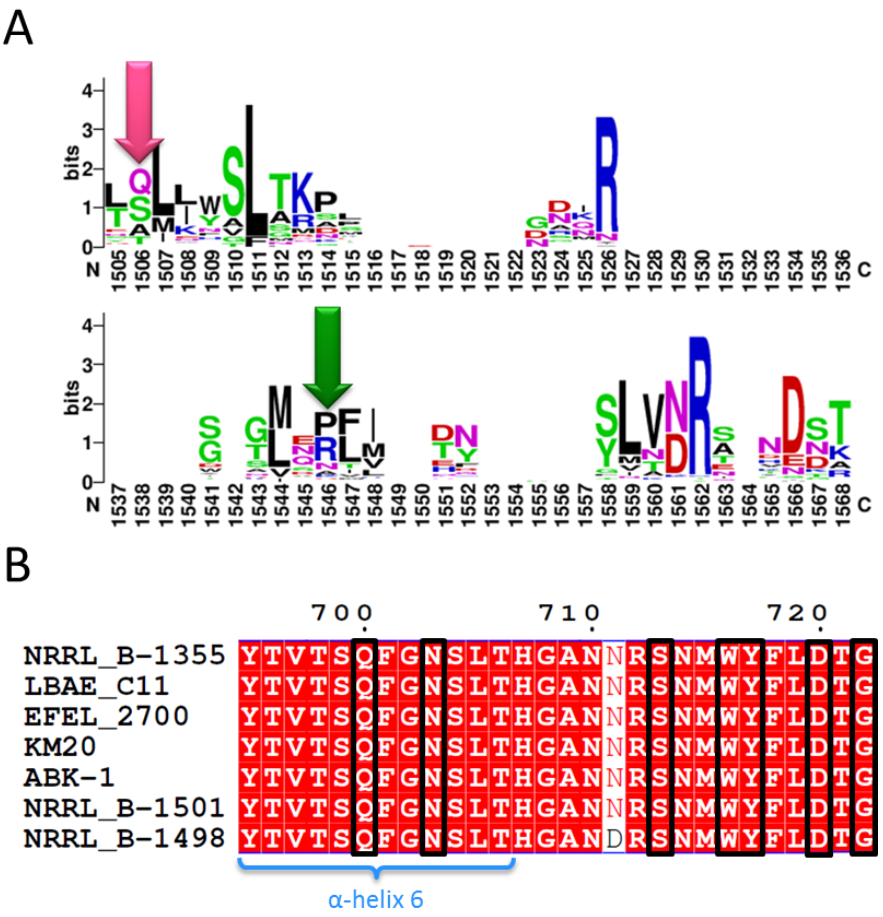


Figure 10: (A) Web logo of all GH70 characterized enzymes.
Pink arrow corresponds to Gln700 position and green arrow, to Tyr717 position.
(B) Alignment of the residues corresponding to SBS-A1 (in black boxes) in all characterized and putative alternansucrases.
Only the strain name is indicated. Species: *Leuconostoc citreum* or *mensenteroides*.

Conclusion

To sum up, we described herein the first 3D complexes of ASR with different sugars (isomaltose, isomaltotriose, isomaltononaose, panose and oligoalternan) bound either in domain A or V. The role of the domain V and the new site proposed as a signature of alternansucrase specificity has been clarified. We have generated mutants quasi exclusively specific for the formation of oligoalternans (ASRD5 Y717A produce less than one %). Engineering alternansucrase for the exclusive formation of HMM alternan remains highly challenging. One could think to engineer the domain V and the sugar pockets in order to increase the affinity for alternan and shift the enzyme mechanism towards more processivity. Having a clear vision of the interaction with a longer alternan chain would also be very useful. As co-crystallization and soaking experiments with long oligosaccharides (DP>9) are likely to be extremely difficult, using different techniques such as CryoEM could be of interest. Simulations could also be performed to predict the positioning of long chains connecting the active site to the domain V *via* the site SBS-A1.

Experimental procedures

Production and purification of ASRΔ2 and ASRΔ5

The *E. coli* BL21 DE3* strain was used for enzyme production. A preculture in LB medium supplemented with ampicillin 100 µg.mL^{-1} of transformed *E. coli* BL21 DE3* with plasmid pET53-*asrΔ2* or pET53-*asrΔ5* was used to inoculate a culture at an OD_{600nm} of 0.05 in ZYM-5052 auto-inducible medium (Studier, 2005) modified by supplementation with 100 µg.mL^{-1} ampicillin, 1% (w/v) α-lactose, and 1% (w/v) glycerol for enzyme production. After 26 hours of growing at 21°C, cells were harvested by centrifugation and resuspended in Binding buffer containing 20 mM phosphate buffer, 20 mM Imidazole (Merck Millipore), and 500 mM NaCl, pH 7.4 supplemented with EDTA-free anti-protease tablets (Roche). Cells were disrupted by sonication and debris was removed by a centrifugation step at 45,000 g for 30 minutes at 8°C. Purification was performed with the ÄKTA Xpress system (GE Healthcare). Two-step purification was performed in a cold chamber at 8°C using (i) a HisTrap HP 1mL column (GE Healthcare) for the affinity step and (ii) a Superose12 16/60 (GE Healthcare) for the size exclusion step, or a HiPrep desalting 26/10 column (GE Healthcare) for desalting. The size exclusion step was performed upstream of crystallization trials and Differential Scanning Fluorimetry assays, and protein was eluted in MES buffer pH 6.5 at 30 mM with 100 mM NaCl and 0.05 g.L^{-1} $CaCl_2$. The desalting step was performed for biochemical characterization, for which protein was eluted in 50 mM sodium acetate buffer pH 5.75. Purified fractions were pooled together and concentrated using AmiconUltra-15 with a cut-off of 50 KDa to 10-15 mg.mL^{-1}. Purification was checked by SDS-PAGE electrophoresis using NuPAGE 3-8% Tris-Acetate protein gels (Invitrogen), and protein concentration was assessed by spectroscopy at 280 nm using a NanoDrop instrument. The theoretical molecular weight and molar extinction coefficient of the enzyme were calculated using the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

Crystallization and Data collection

Crystals of ASRΔ2 were obtained using the conditions identified previously (Molina et al., 2019). Crystals were soaked in the reservoir solution complemented with 15% (v/v) ethylene glycol and a variable concentration of different oligo-saccharide (Table 1). Crystal was then directly cryo-cooled in liquid nitrogen. Data collection was made in the European Synchrotron Radiation Facility (Grenoble, France) on beamlines ID23-1 for Isomaltose and Panose complexes and in ALBA synchrotron (Barcelona, Spain) on beamline XALOC for Isomaltotriose, Isomatononaose and Oligoalternan complexes. Diffraction images were integrated using XDS (Kabsch, 2010) and converted to structure

factors using CCP4 programs. The structure was solved by molecular replacement using PHASER and the unliganded ASRΔ2 structure (PDB ID: 6HVG) as search model. To complete the model and build the oligosaccharides in the density, cycles of manual rebuilding using COOT were alternated to restrained refinement using REFMAC. The structures have been deposited in the protein data bank. Data collection and refinement statistics are shown in Table 1.

Table 1: Crystallographic statistics

Enzyme	**pET53-His-ASRΔ2-Strep**				
Ligand	Isomaltose	Isomaltotriose	Panose	Oligoalternan	Isomaltononaose
PDB ID	6SZI	6SYQ	6T16	6T18	6T1P
Soaking concentration and duration	100 mM 1 minute	100 mM 5 minutes	100 mM 5 minutes	100 g/L 5 minutes	100 mM 5 minutes
Data collection					
Wavelenght (Å)	0.9734	0.9793	0.9979	0.9793	0.9979
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Molecules per asymmetric unit	2	2	2	2	2
Cell constants a, b, c (Å)	100.74 134.69 236.04	100.72 135.65 238.94	101.00 134.30 235.37	100.87 135.65 239.42	100.57 135.21 237.20
Resolution (Å)	50.00-3.00	50.00-3.00	50.00-3.10	50.00-3.15	50.00-3.50
Measured reflections	440819	352514	275779	255328	187213
Unique reflections	65077	61391	58714	57606	39240
Data completeness %	99.9	92.4	99.7	99.9	94.9
Rmerge	0.11 (0.77)	0.07 (0.77)	0.10 (0.78)	0.09 (0.72)	0.14 (0.76)
< I/σ (I) >	6.1 (1.0)	8.5 (1.0)	6.7 (1.0)	7.5 (1.1)	4.9 (1.0)
$CC_{1/2}$	0.99 (0.85)	0.99 (0.82)	0.99 (0.80)	0.99 (0.71)	0.99 (0.71)
Wilson B-factor (Å²)	74.2	73.7	83.0	78.5	99.4
Refinement					
R_{work}/R_{free}	0.207/0.237	0.201/0.237	0.223/0.253	0.200/0.226	0.209/0.264
RMSD bonds (Å)	0.010	0.010	0.010	0.008	0.009
RMSD angles (°)	1.332	1.308	1.282	1.222	1.271
Ramachandran's:					
Favored / Allowed / Outliers	96/4/0	96/4/0	96/4/0	96/4/0	96/4/0
Number of atoms					
Protein	19318	19242	19251	19221	19271
Calcium	2	2	2	2	2
Carbohydrates	69	103	136	146	78
Average B-factor (Å²)	82.0	73.7	105.6	90.0	108.2
Clashscore (percentile)	3 (100^{th})	3 (100^{th})	4 (100^{th})	3 (100^{th})	4 (100^{th})

Site directed mutagenesis study

Mutants were constructed by inverse PCR using the pET53-*asr*-Δ2 or pET53-*asr*-Δ5 genes as template, Phusion® polymerase (NEB), and the primers described in Table S1. Following overnight *Dpn*I (NEB) digestion, the PCR product was transformed into competent *E. coli* DH5α and clones were selected on solid LB medium supplemented with ampicillin 100 $\mu g.mL^{-1}$. Plasmids were extracted with the QIAGEN spin miniprep kit and mutated *asr* genes were checked by sequencing (GATC Biotech). All mutants were produced and purified as described above.

Activity measurement

Activity was determined in triplicate at 30°C in a Thermomixer (Eppendorf) using the 3,5-dinitrosalicylic acid method (Miller, 1959). 50 mM sodium acetate buffer pH 5.75, 292 mM sucrose and 0.05 $mg.mL^{-1}$ of pure enzyme were used. One unit of activity is defined as the amount of enzyme that hydrolyzes 1 µmol of sucrose per minute.

Enzymatic reaction and product characterization

Polymer productions were performed using 1 $U.mL^{-1}$ of pure enzyme with 292 mM sucrose in 50 mM NaAc buffer pH 5.75 at 30°C over a period of 24 hours. The products were analyzed using High Pressure Size Exclusion Chromatography (HPSEC) with Shodex OH-Pak 805 and 802.5 columns in series in a 70°C oven with a flow rate of 0.250 $mL.min^{-1}$ connected to RI detector. The eluent was 50 mM sodium acetate, 0.45 M sodium nitrate and 1% (v/v) ethylene glycol. The polymer yield was calculated using the area of the peak corresponding to HMM glucan divided by the sum of the areas of all the peaks arising on the chromatogram. The same sample was analyzed in triplicate in HPSEC. The molar mass at peak apex was estimated using a calibration curve with fructose, sucrose and dextran standards of 39,000 $g.mol^{-1}$, 11,300 $g.mol^{-1}$, 6,000 $g.mol^{-1}$ and 1,500 $g.mol^{-1}$ at 10 $g.L^{-1}$.

The products were also analyzed by High Pressure Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) using a CarboPac TM PA100 guard column upstream of a CarboPac TM PA100 analytical column (2 mm x 250 mm) at a flow rate of 0.250 $mL.min^{-1}$. The eluents were A: 150 mM NaOH and B: 500 mM sodium acetate with 150 mM NaOH. Sugars were eluted with an increasing 0 to 60% gradient of eluent B for 30 minutes. Quantification was performed using standards of glucose and sucrose at 5, 10, 15 and 20 $mg.L^{-1}$. The hydrolysis percentage was calculated by dividing the final molar concentration of glucose by the initial molar concentration of sucrose.

Acceptor reactions were set up in the presence of maltose (sucrose:maltose mass ratio 2:1) using 1 $U.mL^{-1}$ of pure enzyme with 292 mM sucrose in 50 mM NaAc buffer pH 5.75 at 30°C over a period of

24 hours. The products were analyzed by High Pressure Anion Exchange Chromatography with the same conditions than described above. The structures corresponding to the nomenclature used are: OD4: α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc; OD5: α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc; OA4: α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc; OA5: α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc; OA6: α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc and α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→3)-α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc.

Acceptor reactions with a molar sucrose:glucose ratio of 2:1 were set up in the same conditions as above to produce oligoalternans. Oligoalternans were partially purified by size exclusion chromatography using two 1 m XK 26 columns (GE Healthcare) in series packed with Bio-Gel P6 and P2 resin(Biorad) and water as eluent. The fraction used for soaking experiments corresponds to a DP of approximately 9 as estimated by HPAEC-PAD (Figure S1).

NMR samples were prepared by dissolving 10 mg of the total products from sucrose in 0.5 mL D_2O. Deuterium oxide was used as the solvent, and sodium 2,2,3,3-tetradeuterio-3-trimethylsilylpropanoate ($TSPD_4$) was selected as the internal standard (δ^1H = 0 ppm, $\delta^{13}C$ = 0 ppm). ^1H and ^{13}C NMR spectra were recorded on a Bruker Avance 500-MHz spectrometer operating at 500.13 MHz for ^1H NMR and 125.75 MHz for ^{13}C using a 5-mm z-gradient TBI probe. The data were processed using TopSpin 3.0 software. 1D ^1H NMR spectra were acquired by using a zgpr pulse sequence (with water suppression). Spectra were performed at 298 K with no purification step, for all mutants.

Differential Scanning Fluorimetry was performed with 7 µM of pure enzyme in 50 mM sodium acetate buffer pH 5.75 supplemented with 0.5 $g.L^{-1}$ of calcium chloride and 10 X of SYPRO orange (Life Technologies). A ramp from 20 to 80°C was applied with 0.3°C increments at the rate of 0.3°C per second on a C100 Thermal Cycler.

Affinity gel electrophoresis

4 µg of purified enzyme were loaded in 6.5 % (w/v) acrylamide gels containing from 0 to 0.45% (w/v) of dextran 70 kDa (Sigma) or alternan produced by ASRΔ2 in the conditions described above. Alternan was purified by dialysis against water using a 14 kDa cut-off cellulose dialysis tubing (Sigma-Aldrich). Bovine Serum Albumine (BSA) in 1% (w/v) of NaCl and ladder All Blue Standard were used as negative control (BioRad). Migration was performed in mini PROTEAN system (BioRad) during 30 minutes at 65V followed by 2 hours at 95V in ice. Gels were stained with Colloidal Blue.

Multiple Sequence Alignment

Sequence alignment of putative sugar binding pockets was performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), was inspected and corrected manually using the structural superimposition of pockets V-A, V-B of ASR (PDB ID: 6HVG) and DSR-M (5NGY) and pocket V-L of GBD-CD2 (4TVD) to align the first aromatic residue. Then, the alignment was submitted to WebLogo3 (http://weblogo.threeplusone.com/) (Crooks et al., 2004).

Supporting information

Table S1: Primers used for mutant construction

Mutant	Forward primer (5' -> 3')	Reverse primer (5' -> 3')
Q700A	ACTTCAGCGTTTGGCAATTCTCTAACAC	GCCAAACGCTGAAGTAACTGTGTAATCC
N703A	AGTTTGGCGCTTCTCTAACACATGG	TAGAGAAGCGCCAAACTGTGAAGTAACTG
S713A	CAACAGGGCTAACATGTGGTATTTCTTAGATACTG	CCACATGTTAGCCCTGTTGTTGG
Y717A	ACATGTGGGCTTTCTTAGATACTGGCTATTATC	TAAGAAAGCCCACATGTTACTCCTGTTG
Y241A	GAAAAGTGCTGTCGCCGATAGTTC	CGGCGACAGCACTTTTCTTTAATTG
T249W	CTGGGCAATGGTACTATTTTGATGG	AAATAGTACCATTGCCCAGAACTATCGG

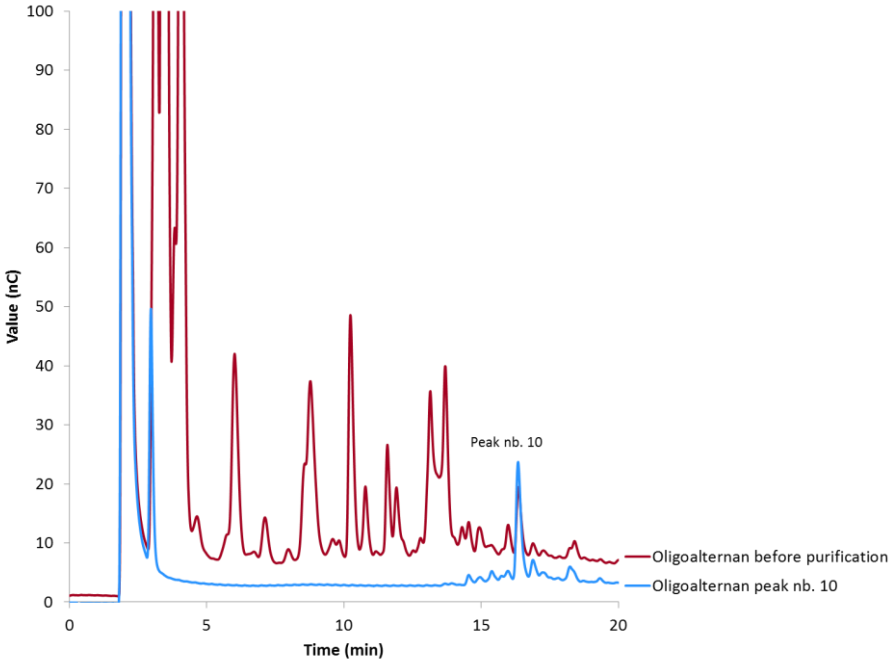


Figure S1: OA used for soaking experiments

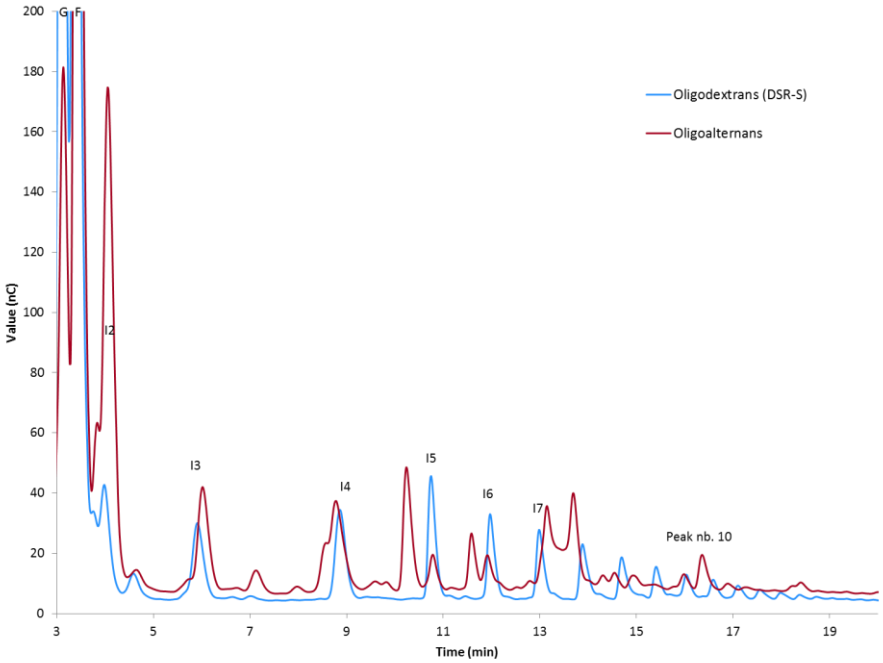


Figure S2: Superimposition of DSR-S and ASR glucose acceptor reaction products.

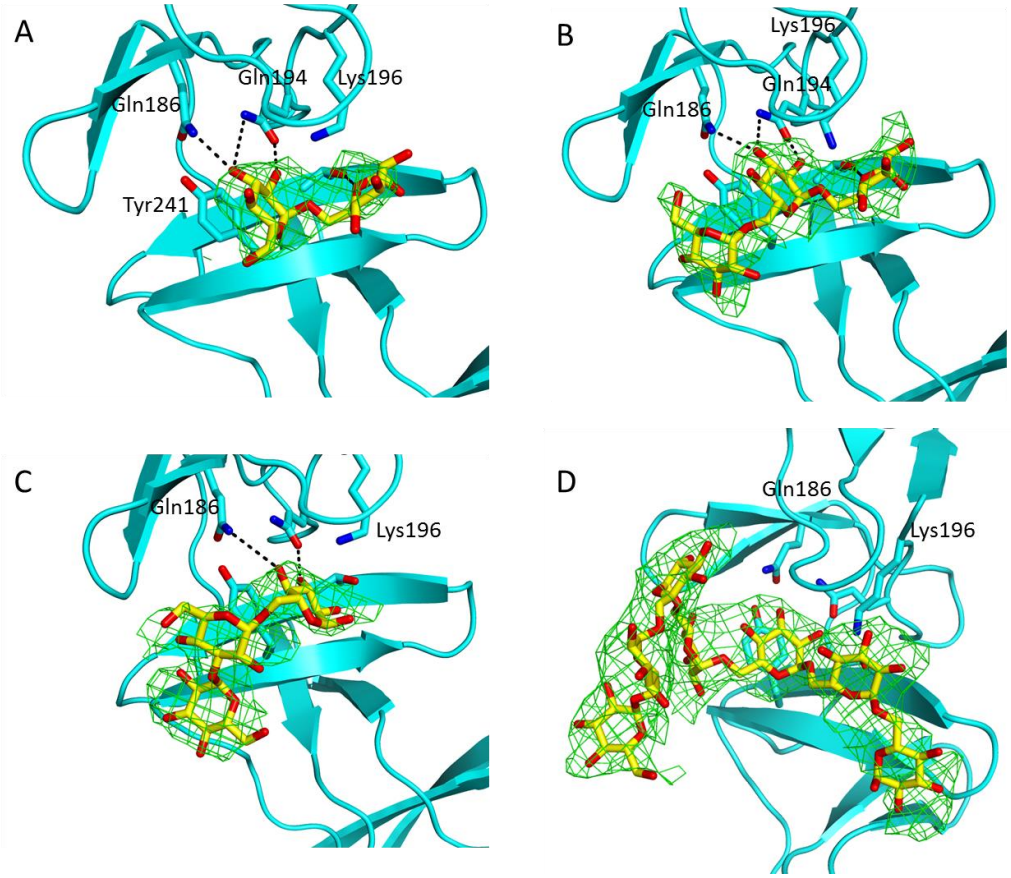


Figure S3: Model of (A) I2, (B) I3, (C) OA and (D) I9 binding in the domain V pocket V-B based on the I2, I3, OA and I9-bound crystal structures respectively.

The difference electron density map (*Fo – Fc)* around carbohydrates was calculated directly after the molecular replacement and contoured at 2.5σ (in green). Final structures of the carbohydrates are superimposed in yellow sticks.

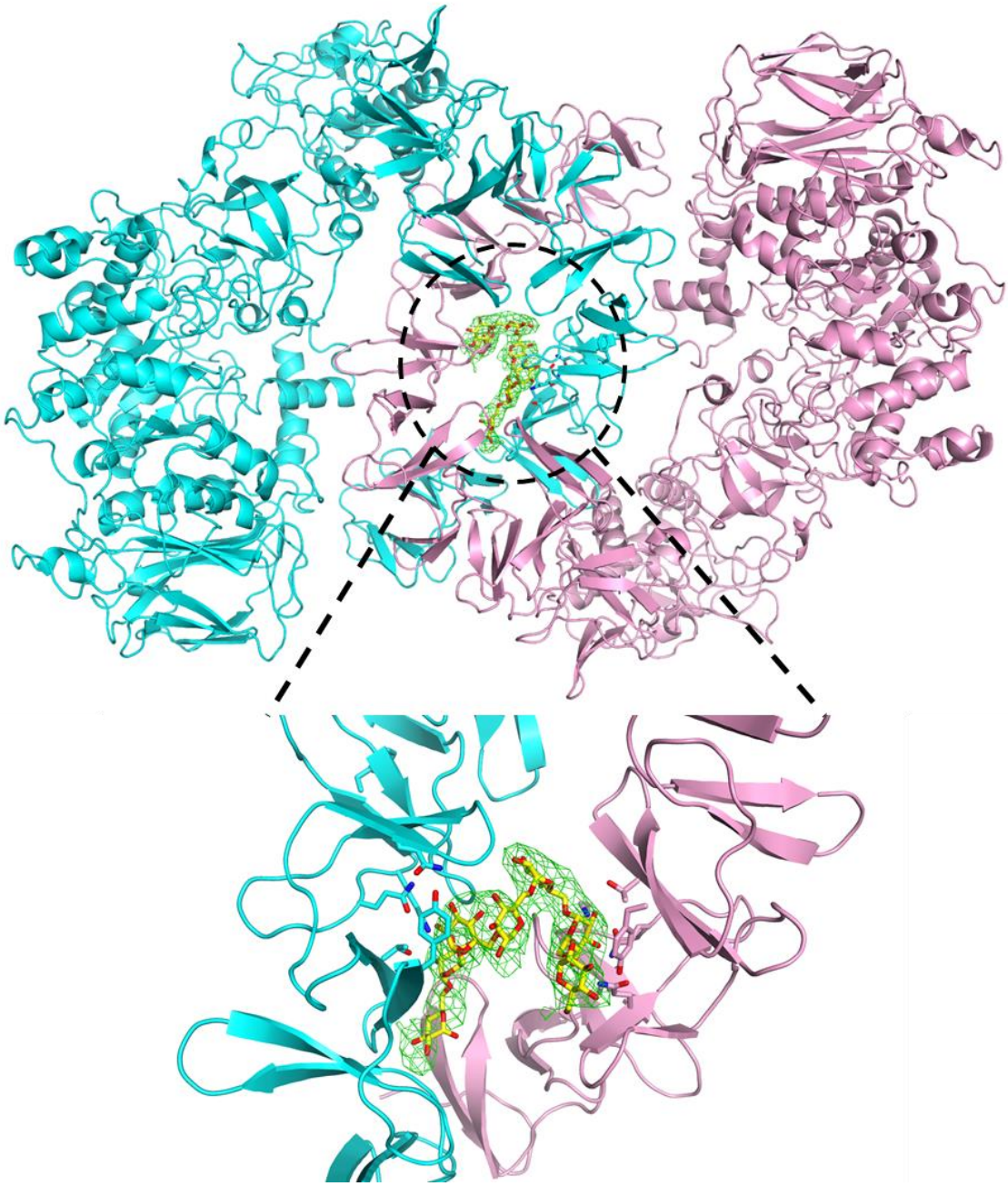


Figure S4: Proposed model of binding for the I9 in the pockets V-B of the two chains.

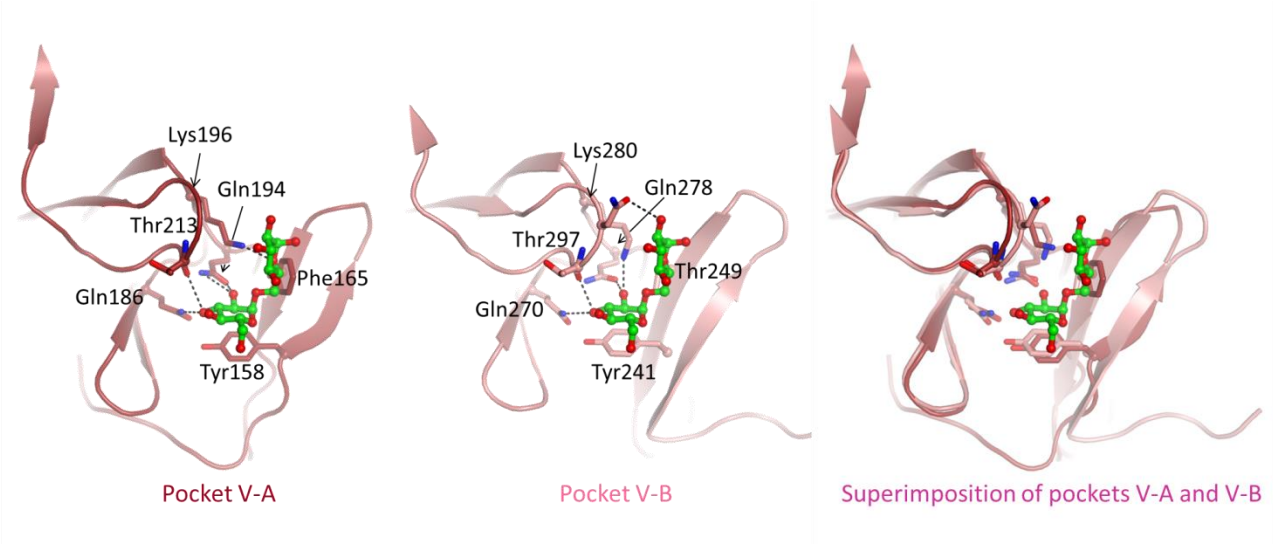


Figure S5: Structural comparison of the two sugar binding pockets

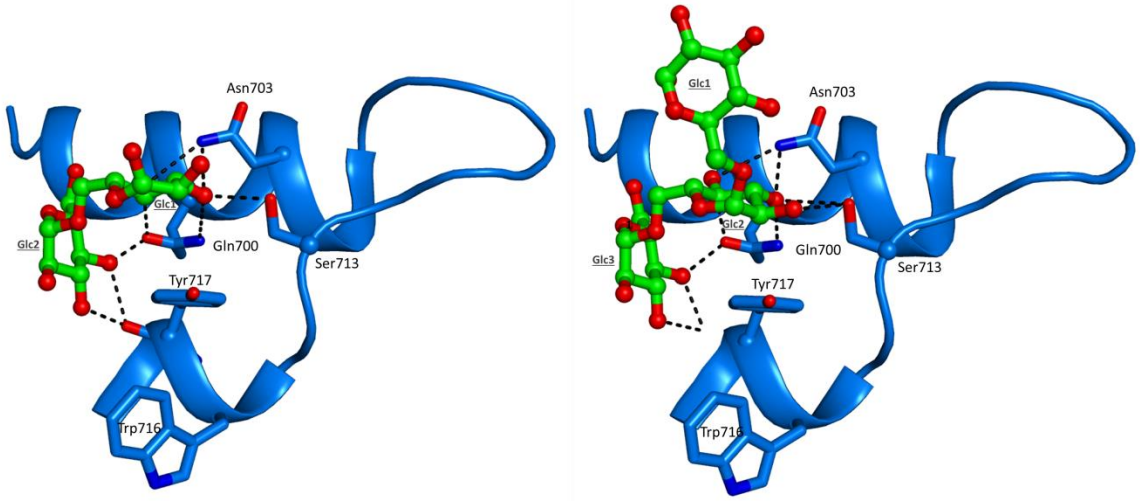


Figure S6: Model of I2 (left) binding versus I3 (right) binding in sugar binding site A1 based on the I2 and I3-bound crystal structure.

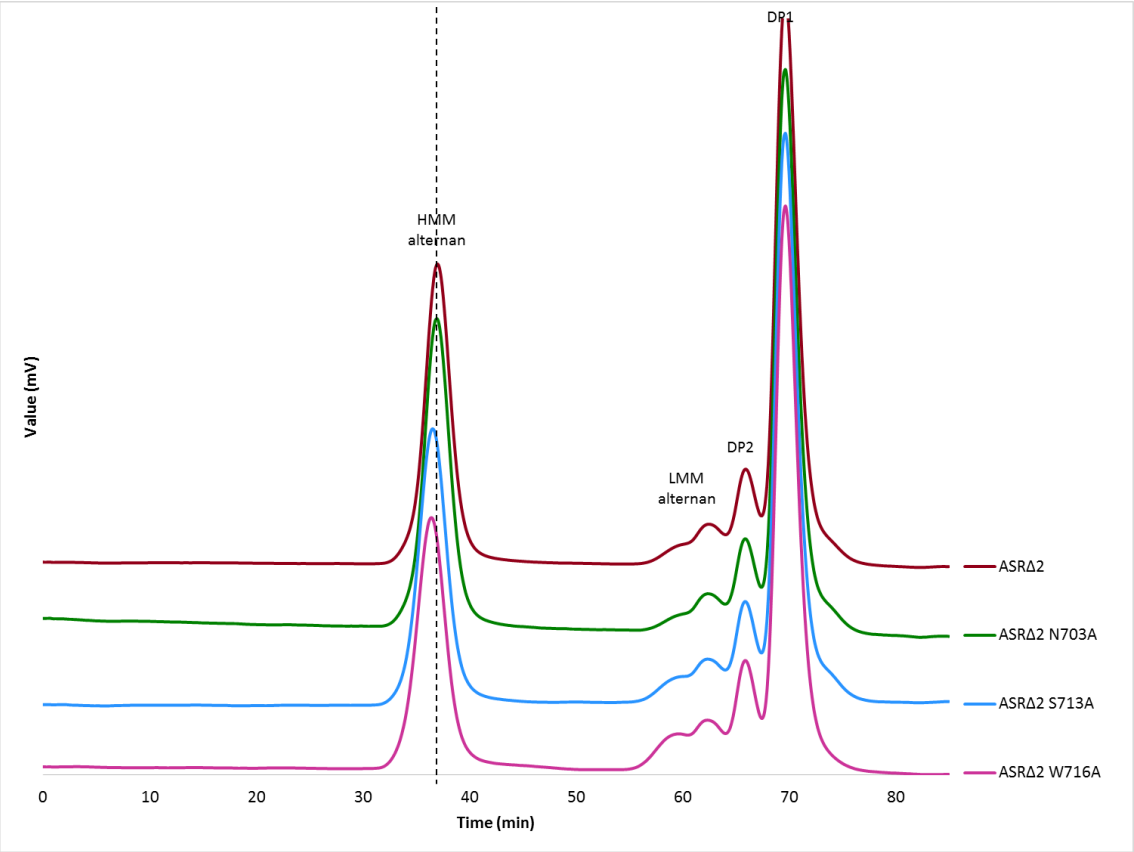


Figure S7: HPSEC chromatograms of SBS-A1 mutants with only slight effect.
Reaction from sucrose at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75.

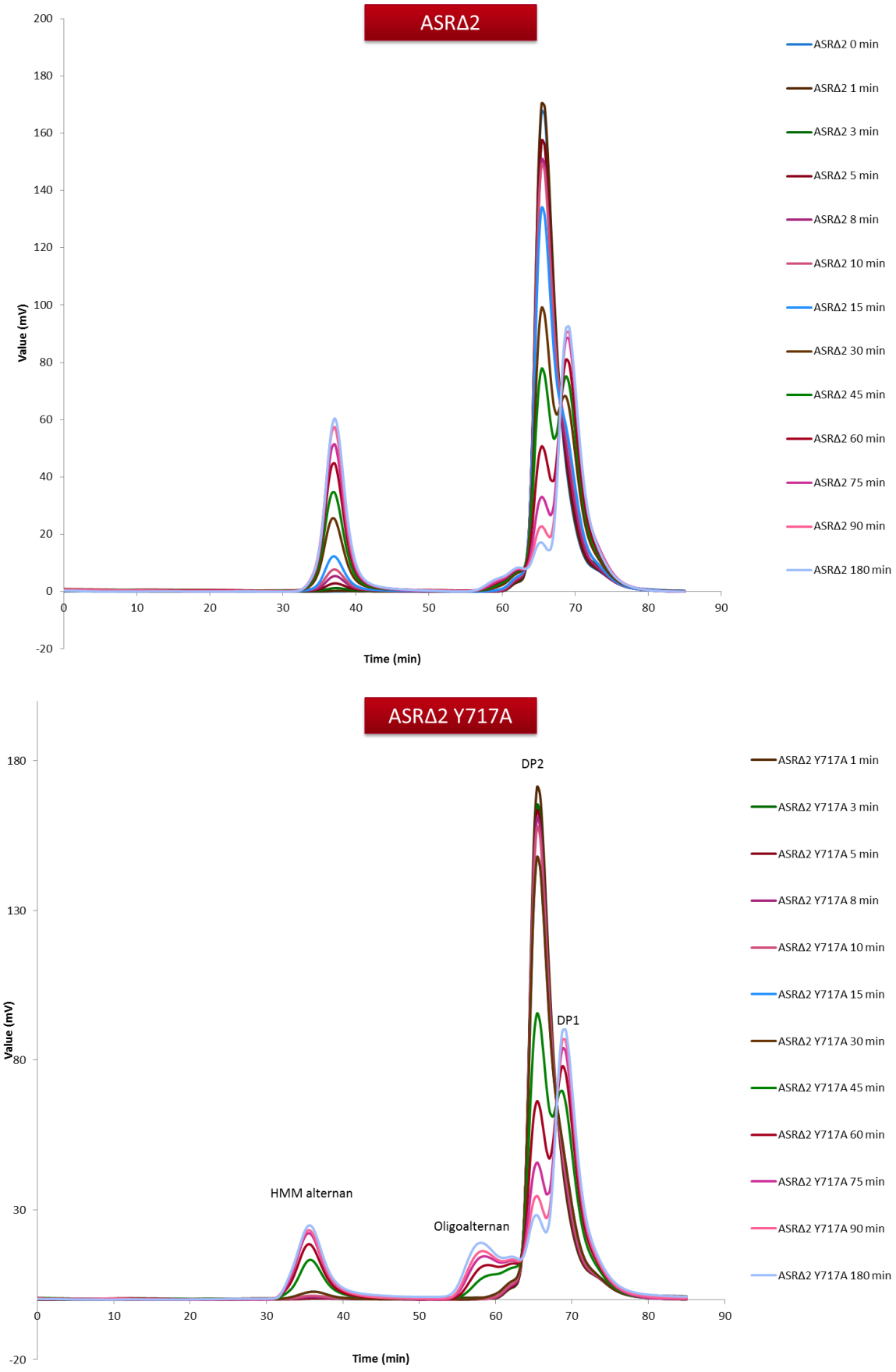


Figure S8: Reaction monitoring from sucrose

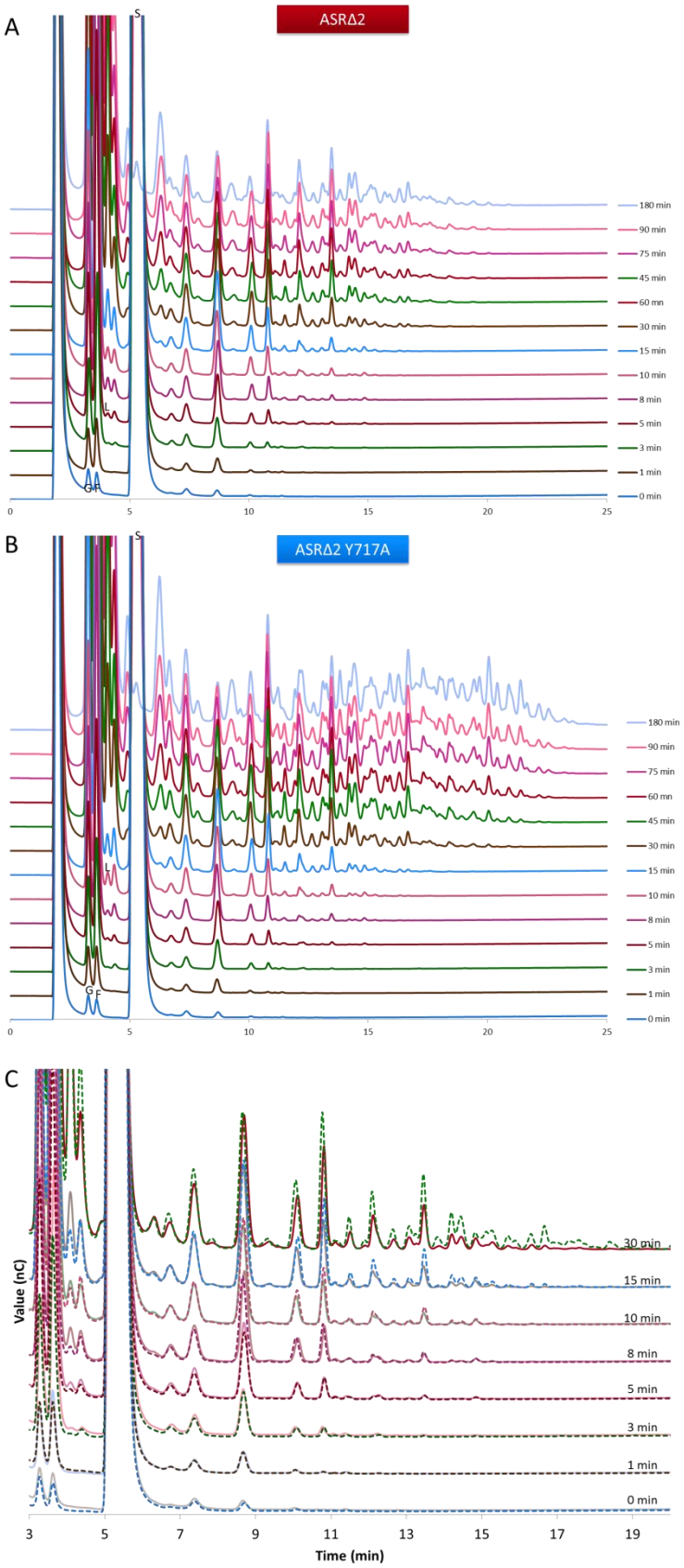


Figure S9: Oligoalternan formation monitoring for (A) ASRΔ2 and (B) the Tyr717Ala mutant. (C) Superimposition of the chromatograms (Dashed lines: Tyr717Ala mutant)

Reaction from sucrose at 30°C with 1 $U.mL^{-1}$ of pure enzyme and sodium acetate buffer 50 mM pH 5.75.

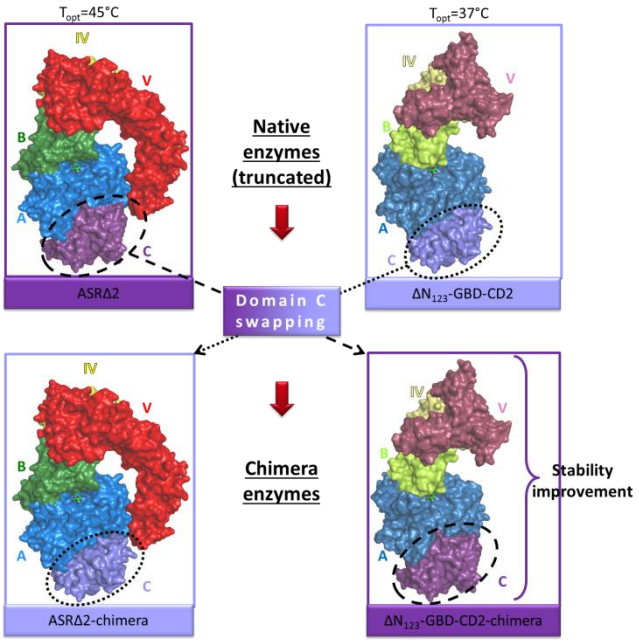
Chapter IV:

Understanding ASR stability and exploration of engineering approaches for ASR stabilization: from random to rational engineering strategies

Abstract

Alternansucrases and glucansucrases are very attractive enzymes for the synthesis of oligosaccharides, polysaccharides or glycoconjugates. However, their thermostability and/or resistance to organic solvent or harsh conditions are significant issues when it comes to the development of robust and cost-effective GH70-based industrial processes. Indeed, GH70 enzymes are mesophilic. No thermophilic equivalents have been described up to now in this family, making protein engineering approaches of prime interest to generate more stable catalysts. Surprisingly, very few studies have addressed the improvement of GH70 enzyme stability by structurally guided engineering or directed molecular evolution. This is likely due to the size of these proteins (between 120 and 200 kDa) and the fact that they are multi-domain enzymes comprising five sub-domains. In the following chapter, our goal was to get new insight in the structural determinants involved in ASR stability and also explore several approaches for ASR stabilization. As previously stated, this enzyme has already found commercial applications and several other potential ones can be listed. In addition, the protein is already known as one of the most stable glucansucrases.

First, and from our structural analysis, we tried to identify specific features that could be involved in the enzyme stability. In particular, the comparison of ASR to other glucansucrases provided structural information on its domain C, which was found to be more compact. We assumed that this atypical structure could be involved in ASR stability and decided to introduce this domain in another less stable GH70 enzyme of known 3D-structure, GBD-CD2.



In addition, we have initiated a work on ASR stabilization and set up screening assays to prepare ASR engineering using directed molecular evolution approaches and/or semi-rational and structurally-guided ones.

Introduction

As shown in our literature review, alternansucrase (ASR) like many glucansucrases is a promising enzyme for industrial processes (see III.4. Applications, page 43). This efficient transglucosylase uses a cheap and abundant substrate, sucrose, to produce high and/or low molar mass alternans, and oligosaccharides. It also shows a high potential for the glucosylation of acceptors including many di- and trisaccharides, α-glucans such as dextrans or amylose, polyols, flavonoids, terpens (André et al., 2018; Bertrand et al., 2006; Grimaud et al., 2018; Morel et al., 2017; Musa et al., 2014).

Improving the stability of alternansucrase would be of prime interest to further develop its industrial applications and also target new ones. Indeed, enzyme stability is a significant issue for the development of robust and cost-effective enzyme-based industrial processes. Using a stable enzyme often facilitates its purification and down-stream processing, and enables its usage on a longer period of time. In addition, working at high temperature often helps to solubilize molecules poorly soluble in water and attempt their glucosylation. In the same line, thermostable enzymes are often more resistant to organic solvents. This is not a general rule but if this could be verified for alternansucrase, glucosylation reactions of water-insoluble compounds could be conducted in organic media.

The ASR optimum temperature is of 45°C. By comparison, dextransucrase DSR-S from *L. mesenteroides* B-512F and GBD-CD2 branching sucrase display an optimum temperature of 37°C and 40°C respectively (Fabre, 2004; López-Munguía et al., 1993). In addition, ASR is one of the most stable enzyme in the GH70 family. The half-life time of the protein was of 75 hours at 30°C and 6 hours at 40°C (Joucla, 2003). By comparison, the half-life time of GBD-CD2 was of only 10 hours at 30°C and 15 minutes at 40°C (Brison et al., 2010).

In the following chapter, we have taken advantage of the 3D structure of ASR (disclosed in chapter II) to get further insight in some structural determinants suspected to be involved in the enzyme thermostability. In particular, the structural comparison of ASRΔ2 and ΔN_{123}-GBD-CD2 revealed striking differences in the 3D structure of their domain C. Notably, we observed the presence of longer loops in ASR domain C together with a higher level of interactions and longer β-sheets. These differences between ASR and GBD-CD2 domain C let us think that ASR domain C could be involved in ASR thermostability. To tackle this question, the domains C of ASR and GBD-CD2 were swapped and two chimera enzymes were constructed, (i) ASRΔ2 with the domain C of GBD-CD2 (ASRΔ2-chimera), and (ii) ΔN_{123}-GBD-CD2 with the domain C of ASR (ΔN_{123}-GBD-CD2-chimera). Here, we compare the

stability, specific activity and specificity of the two chimera enzymes relatively to the wild type enzymes.

In a second part, we have initiated a work to improve ASR stability. First, we analyzed the 3D structure of ASR to propose a structurally-guided strategy of stabilization based on (i) the introduction, in ASR, of ionic interactions predicted in other glucansucrases of known 3D structures using the web server RING (Piovesan et al., 2016) (http://protein.bio.unipd.it/ring/), and (ii) the introduction of disulfide bonds between the C-terminal and the N-terminal sides of each domain, using the calculations of the web server Disulfide by Design 2 (Craig and Dombkowski, 2013) (http://cptweb.cpt.wayne.edu/DbD2/). In parallel, a random approach of evolution was investigated to evolve variants with improved thermostability and higher resistance to solvent. We generated a first library of variants using random mutagenesis at a low mutation rate, from which around 300 hundred polymer-forming variants were isolated and used to undergo another cycle of mutagenesis and screening. This was repeated four times, to finally obtain a library of neutral variants which was screened for both thermostability and solvent tolerance.

I- Investigation on the role of the domain C

Structural comparison of domain C of GH70 enzymes

The structural comparison of the domain C of ASR with those of other GH70 sucrose active enzymes (*i.e.* GTF180, GTF-SI, GTFA, GBD-CD2 and DSR-M) highlighted significant differences (Figure 1). First, ASR presents three amino acid insertions in its domain C: ^{882}SSGKDLKDGE^{890}, ^{913}QDNS^{916} and ^{996}KQDGT^{1000}. Notably, the latter is part of a β-hairpin only found in the ASR structure (from Thr991 to Glu1005) (Figure 2A, B).

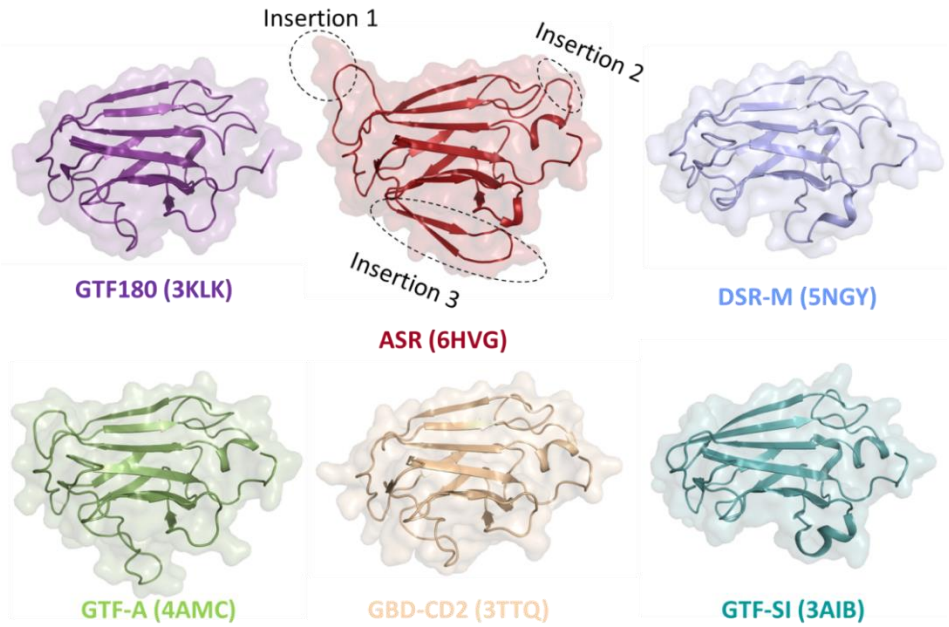
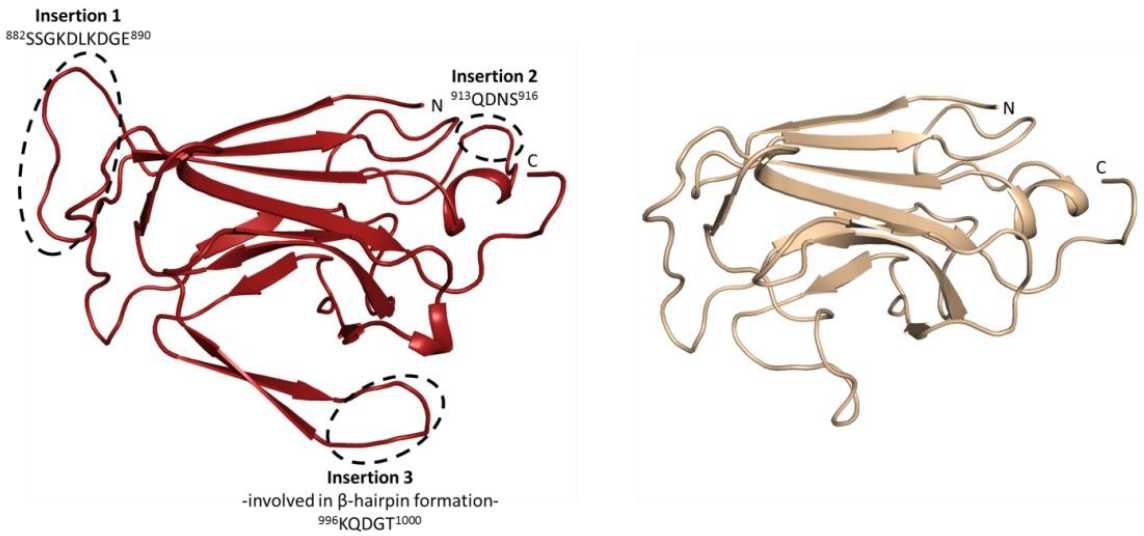


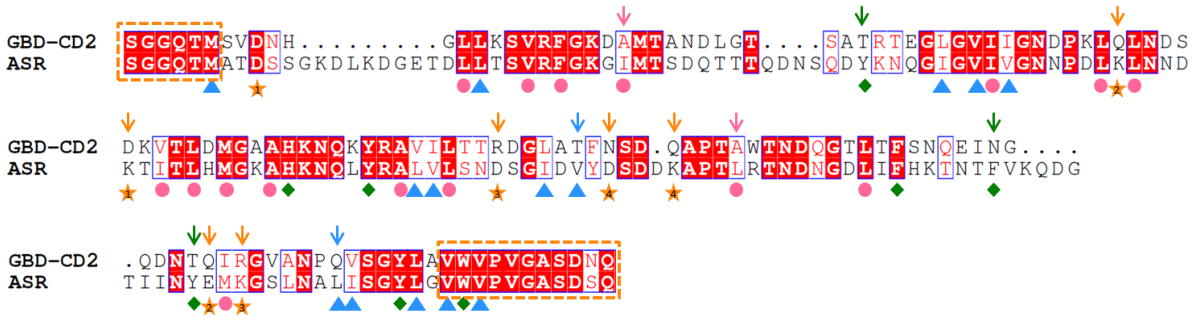
Figure 1: Comparison of the domain C of available glucansucrase and branching sucrase structures

Additionally, ASR β-sheets are of the same length or longer than the corresponding ones found in the other 3D-structures (Figure S1A). In addition, four ionic interactions in domain C can be predicted against two for DSR-M, GTF-SI and GTF180 and zero for GTFA and GBD-CD2. Similarly, ASR domain C harbors seven π-π stacking interactions whereas only two or three are predicted for the others. The alignment of ASR and GBD-CD2 domain C further revealed that ASR domain C contains 19 additional amino acids and 70 mutations, 22 of them being of high similarity score (Figure 2B). In addition, the number of residues Ile, Val, Leu and Phe showing the highest hydrophobicity according to Kyte & Doolite scale (Kyte and Doolittle, 1982) is higher in ASR than in GBD-CD2 domain C. Many of them participate in a hydrophobic interface with domain A or into local hydrophobic cores in domain C (Figure 2B, C).

A


Insertion 1
882SSGKDLKDGE890
Insertion 2
913QDNS916
N
C
Insertion 3
-involved in β-hairpin formation-
996KQDGT1000
N
C

B


GBD-CD2
ASR
GBD-CD2
ASR
GBD-CD2
ASR

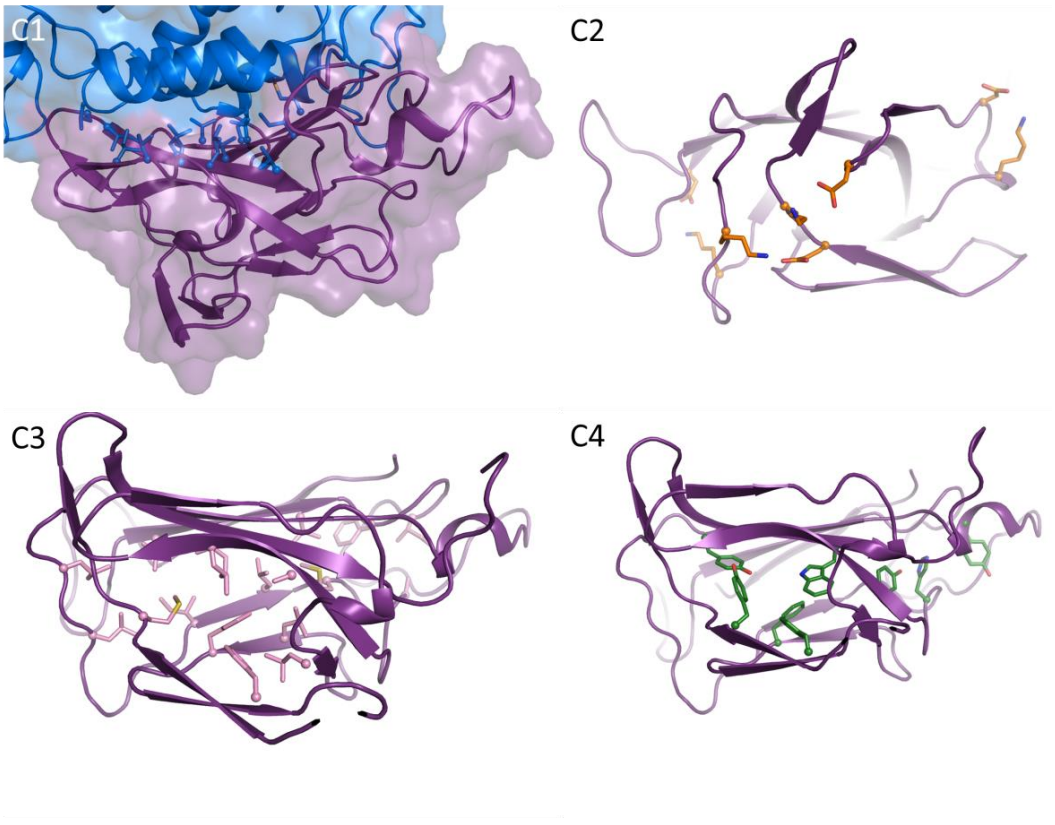

C1
C2
C3
C4

Figure 2: (A) View of domain C organization of ASR (left panel, 6HVG) and GBD-CD2 (right panel, 3TTQ). (B) Sequence alignment of the domain C of ASR and GBD-CD2 (ASR: 873-1030, GBD-CD2: 2424-2562). Arrows indicate the residue that differs in GBD-CD2 and that are thought to be involved in ASR domain C stability or folding. Blue triangle: hydrophobic interface with domain A; orange star: ionic interaction; pink circle: local hydrophobic core; green diamond: π-π stacking interaction. Orange dotted box: high similarity between the N-ter and C-ter of domains C. Alignment created with ENDscript 2 (Robert and Gouet, 2014). **(C1) View of ASR domain C hydrophobic residues (Ile, Val, Leu, Phe, Cys, Met, Ala) involved into the interface between domain A and C** (Met878, Leu895, Ile924, Val926, Val928, Leu958, Val959, Ile966, Val968, Leu1013, Ile1014, Leu1018, Val1020 and Val1022). **(C2) View of the residues involved into ionic interactions** (Asp881-Lys940, Lys935-Glu1005, Asp963-Lys1007 and Asp970-Lys974). **(C3) View of the residues involved into local hydrophobic cores** (Leu894, Phe900, Ile904, Ile927, Leu934, Leu936, Ile942, Leu944, Met946, Ala949, Ala957, Leu960, Leu978, Leu 986, Phe988, Phe994 and Met1006). **(C4) View of the residues involved into π-π stacking interactions** (Tyr919-His950, His950-Tyr955, Phe988-Phe994, Phe988-Trp1021, Phe994-Tyr1004, Phe994-Trp1021 and Tyr1004-Tyr1017).

All together, these specific structural features (*i.e.* hydrophobic cores, hydrophobic interface, ionic or π-π stacking interactions) are likely to be important for protein folding and thermal stability, and let us assume that the ASR domain C could be involved in the higher stability of this enzyme. To assess our hypothesis, we envisioned to introduce the domain C of ASR into a less stable GH70 enzyme. We turned our choice to GBD-CD2 as the recombinant enzyme was known to be less stable than ASR. Logically, two chimeras were constructed from ASRΔ2 and ΔN_{123}-GBD-CD2 in which the domains C were interchanged yielding: ΔN_{123}-GBD-CD2-chimera built of the domains A, B, IV, V of ΔN_{123}-GBD-CD2 and the domain C of ASR and ASRΔ2-chimera corresponding to the domains A, B, IV, V of ASRΔ2 with the domain C of GBD-CD2. Importantly, the design and the construction of these chimeric enzymes were facilitated by the high level of conservation of the amino acid sequences found at the N and C-terminal extremities of both domains C, limiting the risk of a complete destabilization of the overall fold due to domain C swapping (Figure 2B-orange dotted box, Figure 3).

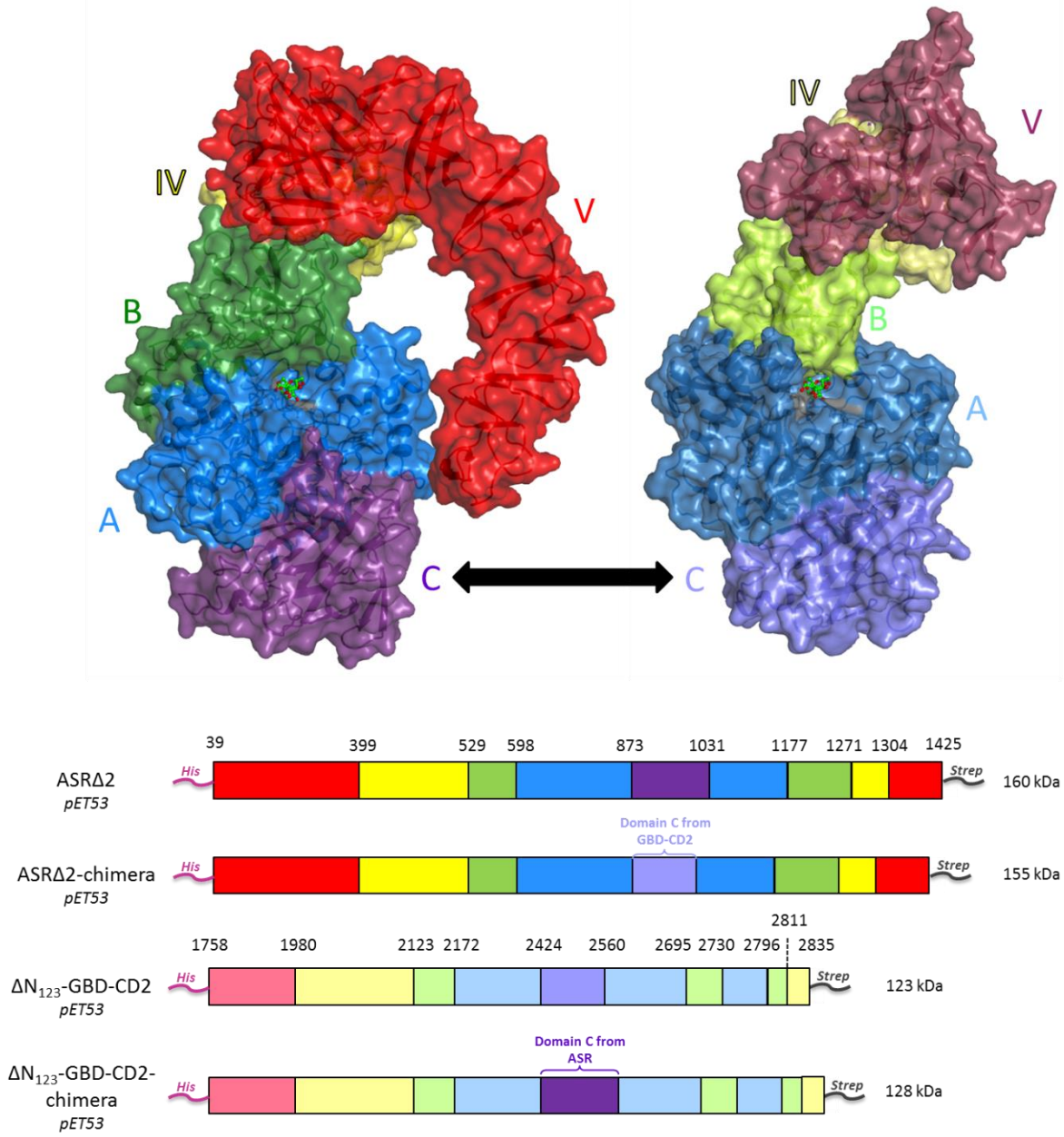


Figure 3: Schematic representation of the domain organization of ASRΔ2, ΔN_{123}-GBD-CD2 and their chimeras. The upper numbering corresponds to residues delimiting the different domains (red: domain V, yellow: domain IV, green: domain B, blue: domain A, purple: domain C) as described (Brison et al., 2012; Molina et al., 2019). His: His-tag, Strep: Strep-tag.

Production of the chimeric enzymes

Unfortunately, the protocols of production previously optimized for ASRΔ2 (Molina et al., 2019) and ΔN_{123}-GBD-CD2 wild type enzymes (Vuillemin et al., 2014) were not adapted for the production of both chimera. Most of the enzymes were recovered in the insoluble fraction as aggregates or inclusion bodies indicating that the exchange of domain C had impacted the chimera folding (Figure 4).

To fix that problem, we produced the wild type enzymes and their respective chimeras with a chaperone protein to facilitate protein folding. For that purpose, we used E. coli BL21 star cells previously transformed with the commercial pTf16 plasmid harboring the Tig chaperone (Takara). As shown in Figure 4, sufficient production levels of both wild type and chimeric enzymes were obtained using Tig chaperone to perform biochemical characterization (Figure 4). Unfortunately, due to a lack of time, ASRΔ2-chimera could not be purified and characterized. Consequently, we will only describe in the following part the preliminary results obtained for ΔN_{123}-GBD-CD2-chimera.

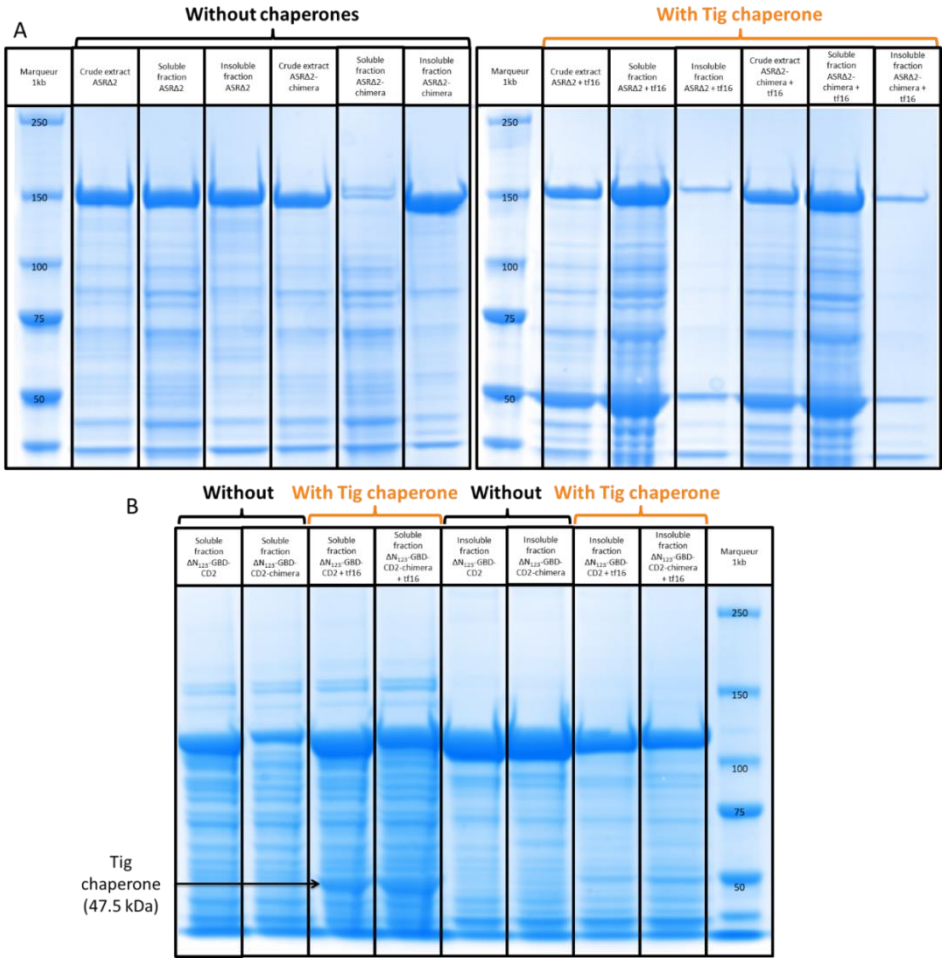


Figure 4: Protein gels with and without tf16 chaperone for (A) ASRΔ2 and ASRΔ2-chimera and (B) ΔN_{123}-GBD-CD2 and ΔN_{123}-GBD-CD2-chimera.

Effect of domain C swapping on GBD-CD2 specificity and thermal stability

The specific activity of ΔN_{123}-GBD-CD2-chimera was estimated at 11.6 $U.mg^{-1}$, versus 16.9 $U.mg^{-1}$ for the wild type enzyme. With 69% of residual activity compared to the wild type enzyme, ΔN_{123}-GBD-CD2-chimera was thus only slightly impacted by the domain swapping. Furthermore, ^{1}H NMR spectra of the reaction products obtained from sucrose and 70 kDa dextran with ΔN_{123}-GBD-CD2 and ΔN_{123}-GBD-CD2-chimera were perfectly stackable (Figure 5). The signals at 5.11 and 5.19 ppm are characteristic of the formation α-1,2 branching linkages in the linear T70 kDa dextran and their integration accounted for 36% α-1,2 linkage in both cases and 0.03% of α-1,3 linkages. These results indicate that domain C swapping did not affect the enzyme specificity, namely its ability to branch dextran molecules with glucosyl units linked through α-1,2 linkage.

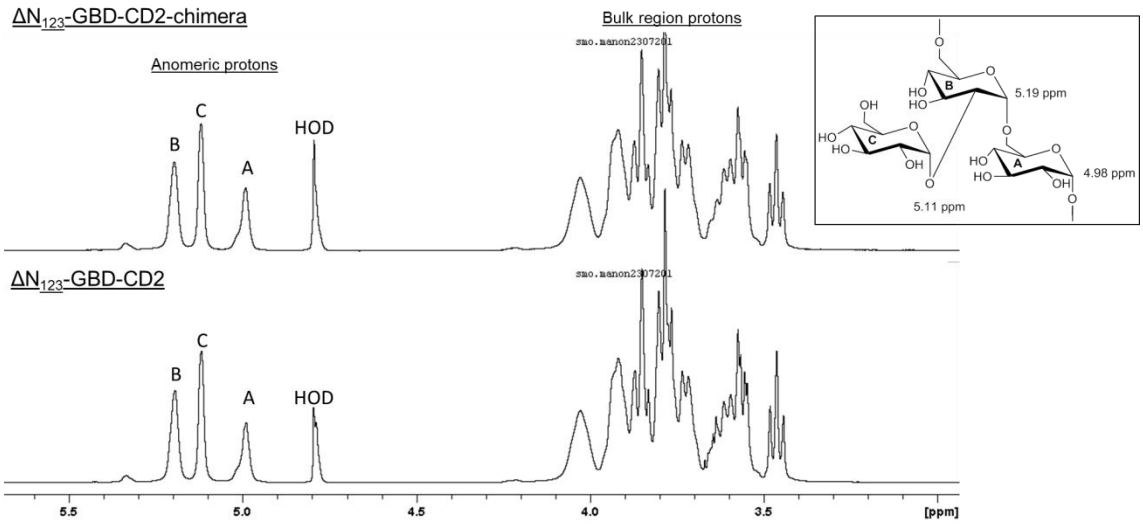


Figure 5: ^{1}H RMN spectra of ΔN_{123}-GBD-CD2 and its chimera.
Reaction from 292 mM sucrose, 309 mM of 70 kDa dextran in 50 mM NaAc buffer pH 5.75 at 30°C over a period of 24 hours.

The melting temperatures of ΔN_{123}-GBD-CD2 and ΔN_{123}-GBD-CD2-chimera were first compared (Figure 6). Both DSF curves revealed the presence of two peaks (A and B) corresponding to two melting temperatures (Tm). Peak B is not as clearly defined as peak A especially for the chimera; enzyme concentration should be increased to obtain a better resolution. The presence of different melting temperatures is often observed for multi-modular enzymes such as the GH70 family enzymes. The first T_m (peak A) and the second one (peak B) correspond to T_m values of 37.7 °C and 44.5 °C for the chimera versus 37.1 °C and 41.6 °C for the wild type enzyme. The increase of T_m especially for the second peak must be taken with caution due the bad peak resolution. The variation of 0.6 °C and 2.9 °C for the two Tms seems to indicate that the chimeric enzyme is more stable than the wild type ΔN_{123}-GBD-CD2.

Another method used to compare enzyme stability is the determination of the residual activity after a period of incubation at a given temperature. After 10 minutes incubation at 30°C, the percentage of residual activity was of 53% and 36% for the chimeric and wild type ΔN_{123}-GBD-CD2, respectively. This difference shows that the chimeric enzyme is more stable than its wild type homologue suggesting that the domain C swapping has improved the stability of ΔN_{123}-GBD-CD2.

Of course, these experiments have to be repeated, but they tend to show that the introduction of the domain C of ASR in GBD-CD2 stabilized the enzyme.

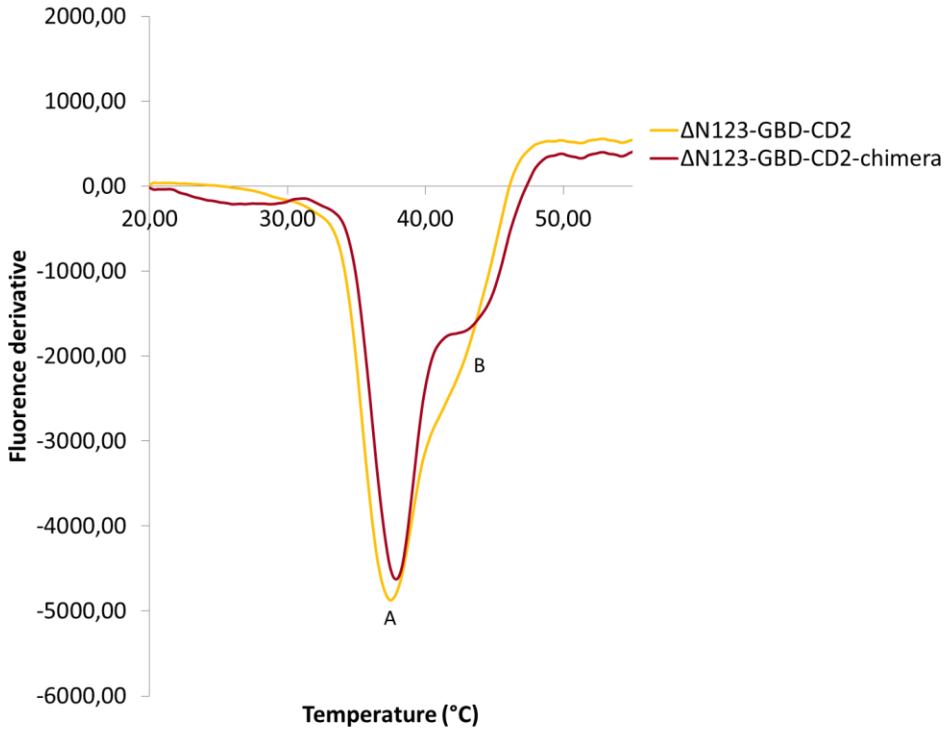


Figure 6: Derivative of the relative fluorescence units (RFU) during the denaturation of ΔN_{123}-GBD-CD2 and ΔN_{123}-GBD-CD2-chimera.

Discussion

In this study, we performed a domain swapping between two enzymes varying in thermal stability. The domain that was switched is the domain C. This domain is found at the basis of U-shape structure. It is more compact in ASR than in GBD-CD2, our two model enzymes, and was suspected to be involved in the highest stability of ASR compared to GBD-CD2. To investigate this issue, we have successfully constructed, produced and purified two chimeric enzymes. Our preliminary results indicate that the branching sucrase ΔN_{123}-GBD-CD2 has increased in stability after acquiring the domain C of ASRΔ2, which is in agreement with our initial hypothesis.

Regarding ASRΔ2, the comparison between the wild type and chimeric enzymes is in progress. It is very likely that the chimeric ASRΔ2 is less stable than its wild type homologue because it can hardly be produced in a soluble form without chaperones, which is usually an indicator of poor stability.

To note, soaking experiments using D-glucose and ΔN_{123}-GBD-CD2 enzyme resulted in the observation of glucose binding site in all domains (A, B, IV, V) (Brison et al., 2016) excepting the domain C. In the other glucansucrase structures (GTF180, GTF-SI, DSR-M or ASR), sugar ligands (maltose, sucrose, isomaltotetraose, glucose) have been observed in domains A, B, IV, V (Claverie et al., 2017; Ito et al., 2011; Vujičić-Žagar et al., 2010) but again, not in the domain C. This is consistent with the fact that this domain would neither be involved in enzyme specificity nor in polymer elongation. Furthermore, we showed that it does not contribute to ΔN_{123}-GBD-CD2 linkage specificity and branching activity. In contrast, our preliminary results tend to show that the domain C could play a role in stabilization. If confirmed, this finding could open the way to further engineering of the domain C of GH70 enzymes targeting stabilization improvements.

As previously underlined, our results have to be confirmed and repeated. Several experiments can be envisioned. Using the production conditions of the chimera set up herein, higher amounts of pure enzymes could be easily obtained to secure the comparison of the chimeric and native enzymes. In particular, T_m assessment will have to be repeated in conditions for which sufficient fluorescence response, at least 4-folds higher than the basic values, will be obtained after the thermal treatment using DSF. In addition, half-life time of the four enzymes will also have to be rigorously determined in similar reaction conditions using the same enzyme concentration for all the assays. Biochemical characterization should also be completed with the determination of the enzyme kinetic parameters to compare their catalytic efficiency (k_{cat}/K_m). Finally, the specificity of the chimeric ASR will also have to be checked to ensure that domain C swapping has not impacted the enzyme specificity and its ability to elongate polymer.

II- Stabilization of ASR using protein engineering

Several strategies can be used to stabilize enzymes and enhance their resistance to co-solvent including enzyme immobilization, chemical modification, the use of different additives and finally protein engineering (Dey et al., 2016). In our work, only **enzyme engineering** was considered for ASR stabilization. Before describing our preliminary results, we will first briefly survey the various methods available to engineer protein thermostability or co-solvent resistance. Interestingly, as several studies described correlations between solvent tolerance and thermostability (Koudelakova et al., 2013; Reetz et al., 2010), we will not distinguish the two properties.

Methods used to engineer enzyme thermostability and co-solvent tolerance

Methods of enzyme engineering can be classified in two categories: the random approaches such as directed molecular evolution, which does not require any structural information, and the rational approaches in which the sequence and structural knowledge is used to target amino acid changes in the protein. The two approaches are more and more combined to design smarter and smaller-sized variant libraries.

- **Random approach**

The relations between the structure of an enzyme and its specificity, catalytic efficiency, thermostability or solvent resistance are not well understood yet. In addition, enzymes have evolved to optimally accomplish their function in their natural environment and in conditions, which may be rather far, particularly in terms of catalytic efficiency or stability, from those required for the development of industrial process. To circumvent this issue, Frances Arnold (Nobel laureate in chemistry, 2018) and Willem Stemmer introduced in the nineties the concept of **directed molecular evolution**, inspired from the **Darwinian process** of evolution (Arnold, 1996; Moore and Arnold, 1996; Stemmer, 1994; You and Arnold, 1996; Zeymer and Hilvert, 2018; Zhang et al., 1997). Enzymes are submitted to successive rounds of *in vitro* **mutations** and/or recombinations and phenotypic **selection** or **screening** to identify variants showing desired properties. Numerous methods have been described, which can be used to generate genetic diversity (for review one can refer to Packer and Liu (Packer and Liu, 2015)). The most employed method is still the error prone PCR (**epPCR**) using Taq polymerase with excess of Mg^{2+} or Mn^{2+} or low fidelity polymerases such as GeneMorph II Random Mutagenesis Kit (Agilent). The generated libraries are usually of large size ranging from 10^4 to 10^9 mutants. However, due to the rare occurrence of beneficial mutations, accurate, **fast** and

high-throughput screening or selection assays have to be developed to isolate the enzyme showing the properties of interest.

Directed evolution has often been successful to generate enzymes with improved thermostability and/or solvent resistance. The protease subtilisin E was engineered to function in polar organic solvent. A mutant with ten mutations was obtained with a **256-fold increase** in its hydrolytic activity on a peptide substrate in presence of 60% dimethylformamide (Chen and Arnold, 1993). The psychrophilic proterase substilisin S41 was also evolved using directed evolution in order to stabilize the protein. A mutant with seven mutations was found to have a half-live time at 60°C **500 times higher** than the one of the wild type. Its temperature optimum also had a 10°C increase but its activity at low temperature was still retained (Miyazaki et al., 2000).

In particular, these methods have been applied to the recombinant amylosucrase from *Neisseria polysaccharea*, a glucansucrase from GH13 family, which produces an amylose-like polymer from sucrose. epPCR was used to generate a library of around 60,000 variants in which 7,000 active variants were selected using sucrose as the sole carbon source. Then, the variants were screened for increased stability by 20 minutes heat shock at 50°C and three positive hits were isolated. One of them, variant R20C/A451T, exhibits a **10-fold increased** half-life time at 50°C (Emond et al., 2008).

Another random-based method of evolution inspired from the **neutral theory** of evolution (Kimura, 1989) was also more recently applied to engineer enzymes. This iterative method consists in introducing mutations through the generation of diversity and retains variants showing the natural phenotype under **non-adaptive conditions**. After **several rounds** of mutations and selection or screening, a **small library** (hundreds of variants) is obtained from which enzymes diverging from their natural phenotype and showing new substrate specificity or increased stability can be identified. The approach has been applied to different enzymes (Bershtein et al., 2008; Bloom et al., 2007; Gupta and Tawfik, 2008; Smith et al., 2011) and more recently to the amylosucrase from *Neisseria polysaccharea* (Daudé et al., 2019). From a library of around 400 variants, several mutants showing an increased ability to use an alternative glucosyl donor (*p*-nitrophenyl-α-D-glucopyranoside, *p*NP-Glc) or glucosylate a range of acceptors were identified. Notably, one mutant (7946E10: Q5R/D231N/I330M/G348S) with a **4°C increased** Tm was also isolated. Altogether, these results illustrate how effective the neutral drift can be for finding variants with a range of new or improved properties from small-sized libraries.

- **Rational and semi/rational approach**

Improving enzyme thermal stability in a rational way is very challenging as it needs first to understand the **structural determinants** of stability. Scientists have collected meaningful information from the sequence and structure comparison of mesophilic, thermophilic and psychrophilic enzymes belonging to the same family or class. From these studies, different structural features favoring thermostability can be listed, which are summarized in Table 1. However, It remains difficult to deduce general trends across all the families (Kumar et al., 2000). In addition, large variations of thermal stability between enzymes can be due to only one or few points mutations, what renders these positions quite hard to locate (Eijsink et al., 2004).

Computational tools provide an effective support to the generation of structurally-guided libraries of variants (Steiner and Schwab, 2012). For example, the **consensus sequence** within a protein family can be rapidly sorted out by multiple sequence alignment and used to introduce conserved amino acids in sequences that diverge from the consensus. Indeed, the consensus residues are thought to participate in protein fitness more than the others and thus could be important players for ensuring protein thermostability. Using this approach, the thermostability of different enzymes including a phytase (**T_m + 30°C** compared to the wild type), a penicillin G acylase or a glucose dehydrogenase (Vazquez-Figueroa et al., 2008) for example was improved and beneficial mutations could be introduced in less stable enzymes of the same family to positively enhance their thermostability. Reconstruction of **ancestral sequences** is another alternative that can be attempted. Indeed, ancestral enzymes are often considered as more stable and less promiscuous than the more recent ones. They can be inferred from nods in the phylogenetic trees allowing the ancient proteins to be reconstructed for increasing thermal stability (Wheeler et al., 2016). FastML is a useful web server (http://fastml.tau.ac.il/) for ancestral sequence reconstruction. FastML codes a Multiple Sequence Alignment (MSA) into a binary indels matrix (presence/absence of characters) and then reconstructs each internal nodes of the phylogenetic tree (Ashkenazy et al., 2012). This server was successfully used to construct ancestors of the human phosphate binding protein (HPBP) and serum paraoxonases (PON) exhibiting a much **higher solubility** or a **30°C higher T_m**, respectively, compared to the wild type enzymes (Gonzalez et al., 2014; Trudeau et al., 2016). Other web servers for ancestral sequence reconstruction such as ProtASR (Arenas et al., 2017) or PROSS (Protein Repair One Shot-Stop web server) (http://pross.weizmann.ac.il/. PROSS) can be used. Notably, PROSS needs only a dozen of homologous sequences and structures to provide several ancestor sequences, expected to be more stable (Goldenzweig et al., 2016).

The **B-Factor Iterative Test** (**B-FIT**) also enabled the stabilization of LipA (181 amino acids) by a **5-fold** improvement of the half-life time at 55°C (Reetz et al., 2006). In this approach, B-factor profiles, providing the oscillation amplitudes of atoms around their equilibrium position from crystallographic data, were used to target mutagenesis positions and served as a guide to evolve the enzyme using using Iterative Saturation Mutagenesis (ISM) (Reetz et al., 2006). This strategy was applied to larger proteins such as the *Rhizomucor miehei* lipase (269 residues) and phosphatidylinositol-synthesizing *Streptomyces* phospholipase D (509 residues) with less success. A **1.2-fold** increase of the lipase thermostability was observed for the lipase after a heat shock at 70°C for five hours, and the half-life time of the phospholipase increased by only **8.7 min** at 65°C (Damnjanović et al., 2012; Zhang et al., 2012). In other studies, B-Factor analyses were successfully applied to improve the thermostability of an amine transaminase and a lipase (Huang et al., 2017; Zhang et al., 2016). For the latter enzyme, the method relied on **Active Center Stabilization** (ACS) consisting in the selection of high flexible residues within 10Å of the catalytic residues as target for saturation mutagenesis was applied (Zhang et al., 2016).

Table 1: Determinants suggested to be involved in thermostability according to: 1= (Jaenicke and Böhm, 1998); 2=(Prakash and Jaiswal, 2010); 3= (Dey et al., 2016); 4= (Hakulinen et al., 2003); 5= (Kumar et al., 2000); 6= (Guérin et al., 2012); 7= (Eijsink et al., 2004)

Thermostability determinants	References
Compactness	1
Loops shortening	1, 2, 3
α-helix stabilization (helix capping)	1,2, 4, 7
Presence of disulfide bridge	2, 3, 4
Presence of salt bridges	1, 2, 3, 4, 5, 7
More hydrophobic residues and hydrophobic interactions	1, 2, 5, 6, 7
Hydrogen bonds	1, 2, 4, 5
More charged residues	2, 4, 5, 6
More proline	2, 3, 6, 7
Electrostatic interactions	7

Algorithms have also been developed to find interaction network in proteins. RING (Residue Interaction Network Generator) web server (http://protein.bio.unipd.it/ring/) is an example. The algorithm uses a PDB file as input to determine the **interaction network** and list all the residues involved in hydrogen bonds, Van der Waals interactions, Disulfide bond, salt bridge, π-π stacking and π-cation (Piovesan et al., 2016). **Disulfide bridge** can also be designed using Design 2 (http://cptweb.cpt.wayne.edu/DbD2/). This algorithm needs a PDB file as input to give a list of mutations predicted to be compatible with the formation of disulfide bonds through the introduction

of cysteine residues (Craig and Dombkowski, 2013). Among the structure-guided approaches, the algorithm SCHEMA was also proved to be rather efficient to permute domains or segment fragments from one protein to another and generate chimeras of cellobiohydrolases, cellulases and P450s with improved thermostability (Heinzelman et al., 2009, 2010; Li et al., 2007).

The growing number of available structures and the advent in computational protein design also enables different methods to be combined. This is typically the strategy proposed by **FRESCO**, Framework for Rapid Enzyme Stabilization by COmputational libraries. Various algorithms are used to generate stabilizing mutations, which are screened *in silico* using molecular modelling and molecular dynamics simulations before being experimentally verified, confirmed and combined. The thermostability (ΔTm, app +22-35°C) and cosolvent (DMSO, DMF, methanol) tolerance of an epoxide hydrolase, two dehalogenases and a peptide amidase were improved using such a computational design and screening strategy (FRESCO, framework for rapid enzyme stabilization by computational library design) (Arabnejad et al., 2017; Floor et al., 2014; Wijma et al., 2014; Wu et al., 2016).

II.1. Exploration of rational approaches to improve ASR stability

To the best of our knowledge, only one study aimed at improving GH70 family enzyme stability using protein engineering. It concerns the dextransucrase from *L. mesenteroides* 0326. A model of the structure was generated and six residues, proline or lysine, located on enzyme surface in domains A, B, C or IV were selected for mutagenesis and replaced by amino acids enabling the introduction of hydrogen bonds (serine or threonine). Following this approach, the P473S/P856S mutant was obtained. It displayed a 2-fold increase in catalytic efficiency, the same optimum temperature of 25°C and a 8-fold increase of half-life time at 35°C (6.6 minutes for the wild type to 48.8 minutes for the mutant) (Li et al., 2018).

The salt-bilization approach

The level of amino acid interactions is described as a key determinant of stability (Dey et al., 2016). To compare the interactions found in ASR and in the other available structures of GH70 enzymes, we first performed a RING analysis (Piovesan et al., 2016). The number of predicted interactions on an average of 830 residues (Table 2) selected by structural alignment is reported in Table 3.

Table 2: Residues considered for each structure for the interaction network analysis

Enzyme	PDB ID	Chain	N-term	C-term	Length
GTF-SI	3AIB	C	244	1079	836
GTF180	3KLK	A	796	1627	832
GTFA	4AMC	A	795	1632	838
ASR	6HVG	A	402	1294	893
DSR-M	5NGY	A	424	1305	882
GBD-CD2	3TTQ	A	1986	2825	840

Table 3: Interaction network predicted by RING. Energy from (Piovesan et al., 2016)

Interaction Energy (Kj/mol)	**ASR**	**GTF-SI**	**GTF180**	**GBD-CD2**	**DSR-M**	**GTFA**
IONIC 20.0	**21**	20	15	11	11	7
HBOND 17.0-40.0-115.0 (based on distance)	710	696	**739**	704	681	650
PIPISTACK 9.4	**49**	45	47	41	46	45
PICATION 9.6	3	**4**	3	1	3	**4**
VDW 6	927	909	**989**	988	853	930

The ASR was found to have the highest number of ionic and π-π stacking interactions. The enzyme is ranked second for the H-bond interactions right behind GTF-180. These results must be taken with caution as the analysis is based on the static structures and also because we have noticed that the predictions sometimes varied from one enzyme to another although amino acids were strictly conserved. However, the comparison of interactions is consistent with the stability of ASR compared to the other GH70 members.

Furthermore, RING analysis revealed ionic interactions in GTF-SI, GTF180, GBD-CD2, DSR-M or GTFA that were absent in ASR. From this observation, we devised a **"Salt-bilization" strategy** consisting in introducing these interactions into the ASR template and examined their effect on enzyme stability (Figure 7). The suggested mutations were first checked visually to discard those introducing structure disruption or targeting positions important for catalysis, specificity, calcium binding site or enzyme packing (residues involved in local hydrophobic cores). Based on this analysis, a total of 68 possible mutations could have been implemented in ASR. However, we considered that this high number of mutations could be detrimental for enzyme folding and stability, and preferred to introduce mutations step by step. To do so, we distinguished three types of mutations to design 3 synthetic sequences (Table 4, Figure 8). The first one contains only single mutations that were predicted to result in ionic interactions with an amino acid partner already present in ASR. The second one contains a set of double mutations creating ionic interactions between the N- and C-term parts of the various fragments constituting the domains A, B, IV, V. Finally the third sequence regroups all the predicted mutations targeting only the domain A. The synthetic genes corresponding to these three sequences were ordered from TWIST. In parallel, several single mutants representative of the most conserved ionic interactions were constructed manually (in domain A: Asn750Asp, His1049Arg and Pro1073Lys; in domain B: Asn1209Asp; the Asn624Asp-Glu1157Lys junction).

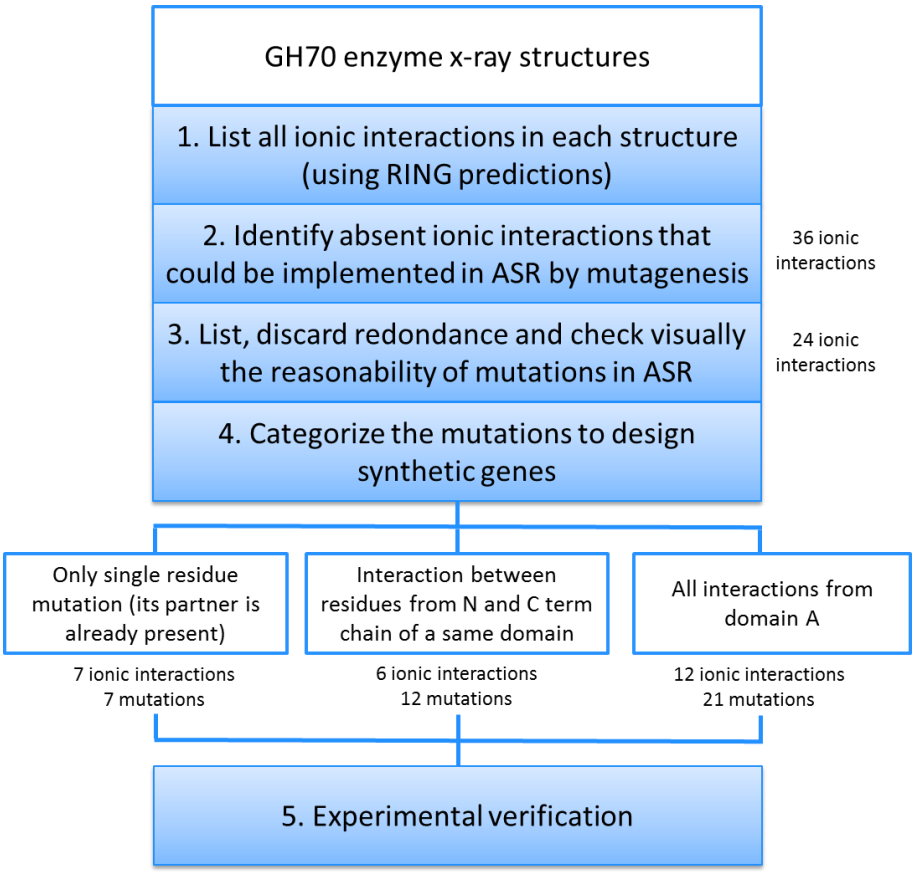


Figure 7: Framework of the "salt-bilization" strategy.

Table 4: Mutations of the three synthetic genes ordered trying to increase ionic interaction number

Gene	Strategy	Nb. of possible new ionic interactions	Nb. of mutations	Mutations (bold) -- represent the ionic interaction
1	All single mutations	7	7	K832--**N750D** D427--**D520H** E447--**S415K** D1044--**H1049R** D449--**T443K** D1151--**P1073K** R1270--**N1209D**
2	N-C-ter junction in domains A and B	6	12	**G540D--E1191K** **A550K--Q1195D** **N624D--E1157K** **D659E--G1045K** **F793D--M1069K** **S862E--L1087K**
3	All domain A	12	21	**N624D--E1157K** **K653R--A689D** **D659E--G1045K** **A663K--K667D** **N684H--N688D** K832--**N750D** **F793D--M1069K** **S862E--L1087K** D1044--**H1049R** **N1057R--S1095D** D1151--**P1073K** **N1086D--N1162K**

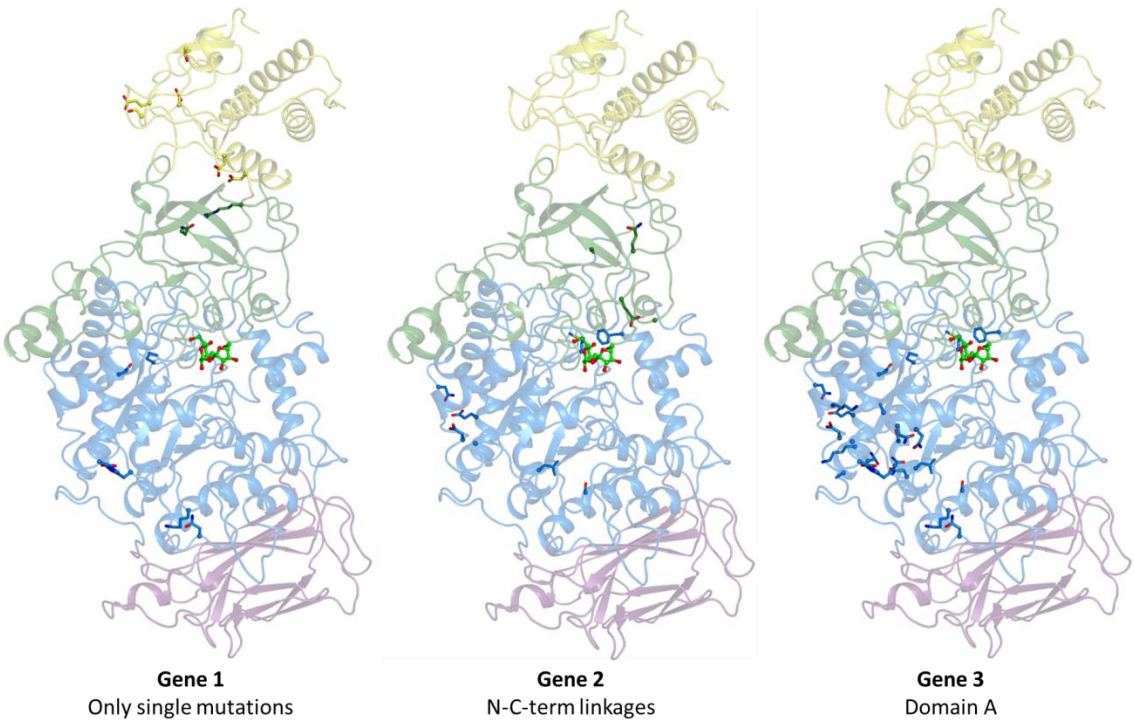


Figure 8: Location of mutations for salt-bilization.
Residues predicted to be in an ionic interaction are indicated in sticks. Domain V not shown.

Unfortunately, we did not have enough time to produce, purify and fully characterize all these mutants, only those listed in Table 5 were characterized. As shown in Table 5, the T_m values did not reveal any stability improvement. Their specific activities were also determined at 30°C and vary between 69% and 102% of residual activity compared to the wild type enzyme (Figure 9). The synthetic gene 2 and 3 still need to be expressed and their product characterized.

Table 5: Melting temperatures of "salt-bilization" strategy mutants.

Mutant	T_m1 (°C)	SD	ΔT_m	T_m2	SD	ΔT_m
ASRΔ2	55,2	0	0	39,9	0	0
N750D (domain A)	55,1	0,2	-0,1	39,9	0	0
H1049R (domain A)	55,2	0	0	39	0	-0,9
P1073K (domain A)	55,2	0	0	39	0	-0,9
N1209D (domain B)	51	0	-4,2	39,6	0	-0,3
N624D-E1157K (domain A, junction N-C-sides)	55,2	0	0	39,3	0	-0,6
Synthetic gene 1 (all single mutations)	53,9	0,2	-1,3	38,1	0	-1,8

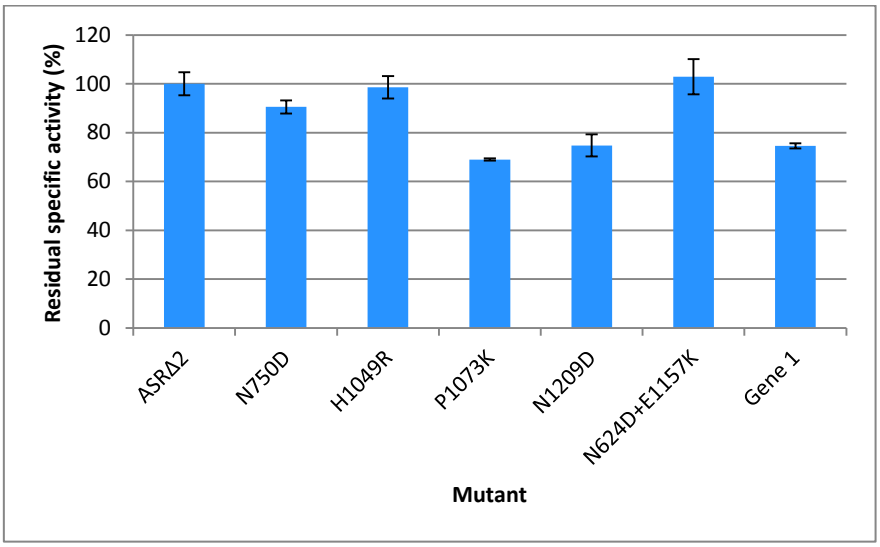


Figure 9: "Salt-bilization" strategy mutants residual specific activity (%) compared to the wild type ASRΔ2
Reaction at 30°C with 292 mM sucrose and 50mM sodium acetate buffer measured in triplicates.

Introduction of disulfide bridge

Another strategy, based on the introduction of disulfide bridges (167 kJ/mol) was envisioned. The objective was to create a disulfide bridge between the N-terminal side and the C-terminal side of each domain, to reinforce the enzyme packing and the links between the two arms of the U-shape fold. Disulfide bridges were designed using DbD 2 and firstly selected on the basis of the lowest predicted energy. Indeed, 90% of 1418 known disulfide bonds have an energy lower than 1.94 kcal/mol (Craig and Dombkowski, 2013). The N access, representing the solvent accessible surface, was also used as a selection criterion. The positions with low N access were favored to protect the disulfide bridge from water. Finally, the B-factors were also examined to favor the residues with highest B-factors for which the mutation is assumed to generate a better stabilization effect. The mutations were checked visually using COOT and constructed by inverse PCR (Table 6, Figure 10). The mutants will be produced and characterized soon.

Table 6: Mutations for disulfide bridge introduction

Domain	Mutations	Energy kcal/mol	Σ B-factor
Domain V	A330C, A1410C	2,81	139,29
Domain IV	A408C, N1301C	1,75	119,7
Domain B	Q545C, N1266C	0,93	95,56
Domain A	P601C, D1249C	1,32	103,65
Domain C	G874C, F900C	1,75	98,89

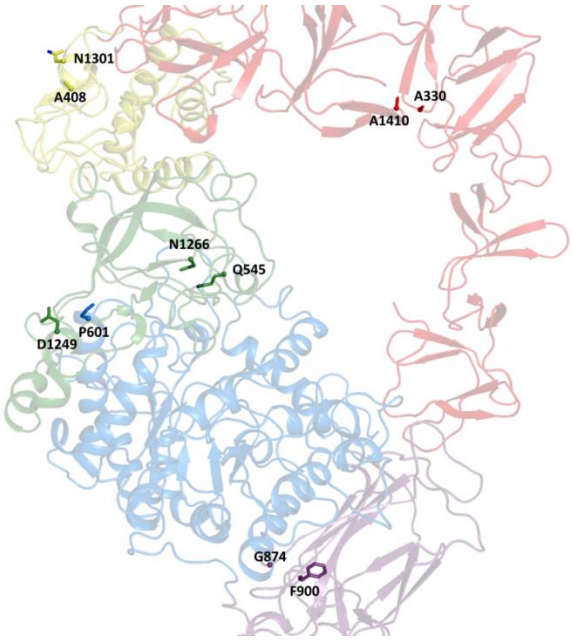


Figure 10: Localization of mutant position for disulfide bond addition

II.2. Random engineering of ASR

Implementation of a screening assay

The first step was the development of a screening assay to allow the isolation of variants keeping the polymerase activity. When grown on solid medium supplemented with sucrose, the colonies of *E. coli* producing the wild type ASR are surrounded by a bubble of polymer ASR. The natural lysis of some recombinant cells may be sufficient to release the active ASR, which produces alternan from sucrose (Figure 11).

This observation allowed us to implement a screening assay to isolate active variants comprising the following steps (Figure 12):

(i) transformation and growing of competent *E. coli*,
(ii) induction of glucansucrase production using IPTG,
(iii) addition of sucrose and enzyme buffer to the medium to allow alternan formation.

We determined the best parameters to be used for the three steps, in terms of duration, temperature and IPTG, DMSO or sucrose concentration (Figure 11). To note, the maximum temperature to be used was 40°C because above, the colonies were not able to grow in liquid LB supplemented with ampicillin. Thus, DNA could not be recovered. *E. coli* was first transformed and grown on a membrane placed on solid LB medium at 37°C to allow the formation of the colonies. Then, the membrane was transferred on solid LB medium supplemented with 1 mM IPTG at 21°C for 24h to induce ASR production. Finally, the membrane was transferred on a solid medium supplemented with sucrose, sodium acetate buffer and DMSO 20 % at 40°C (Figure 12). Using this DMSO concentration and this temperature allowed the observation of the formation of very small bubbles of polymers distinguishable from the colonies producing only an inactive ASR (Glu673Gln mutant).

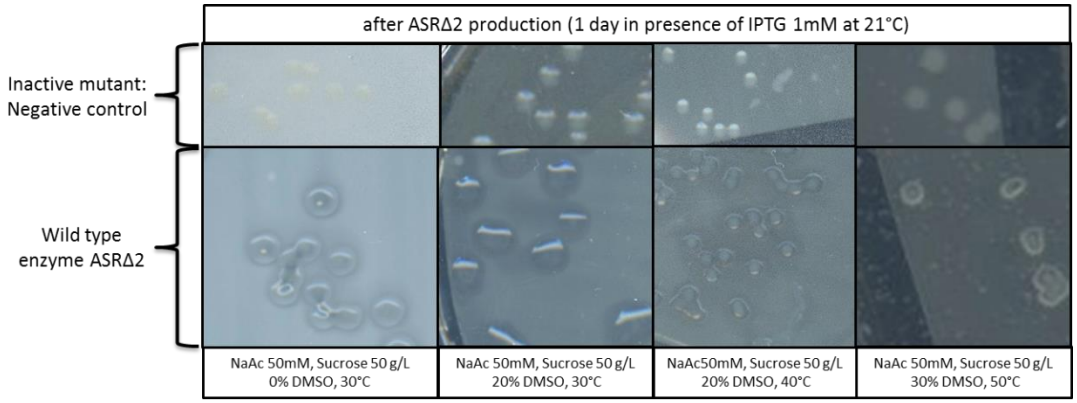


Figure 11: Bubbles observed on solid medium in presence of 50 g/L sucrose, 50 mM sodium acetate buffer pH 5.75 at various temperature and DMSO concentration.

The first two steps of the screen were prior to these experiments as detailed below.

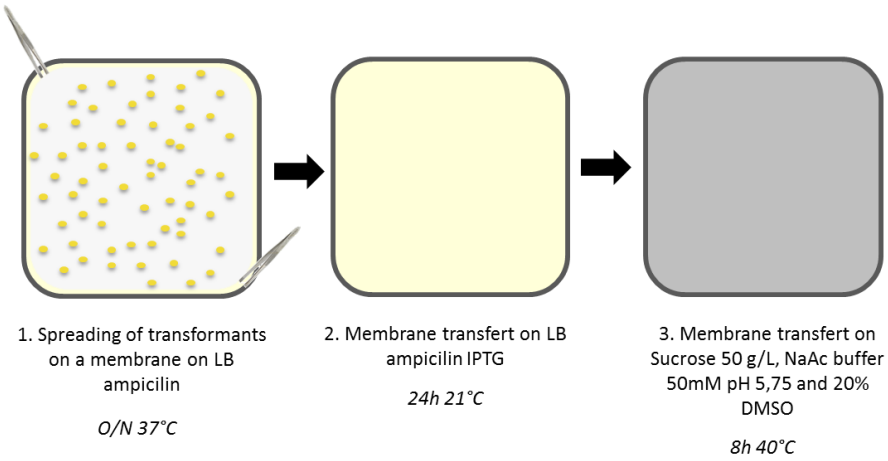


Figure 12: Scheme of the three-steps screening.

Construction of the variant libraries

We decided to set up an evolution process inspired from the neutral drift approach (Gupta and Tawfik, 2008). We first constructed a library of 7,530 variants using epPCR and a low mutation rate (between 0 and 4.5 mutation/kb). We decided to focus the introduction of genetic diversity only on the catalytic domain (A, B and C domains; corresponding to 2.2 kb) as mutation away from the catalytic center may not be sufficient to protect the catalytic core from heat-induced effects for large enzymes (Zhang et al., 2016). The library was screened using the assay described above and we retained ≈ 300 polymer-producing mutants, which were pooled and used as template for the next round of mutagenesis/screening. This procedure was repeated four times.

The bank sizes, the average mutation rates, the maximum and minimum number of mutations per mutant were calculated from around 15 mutants taken randomly that were sequenced (Table 7). Overall, the results are in agreement with those expected, and a low mutation rate was observed, except that the A and T were more mutated than the G and C (Table 8).

Table 7: Data on the random bank generated (before screening for polymer-forming variants).
Around 15 mutants sequenced after transformation of the epPCR product

N° round	Round 1	Round 2	Round 3	Round 4
Bank size	7,530	3,450	7,950	13,350
Average number of mutations per plasmid	7	10.3	9.4	10.6
Average number of mutations per kb	3.1	4.6	4.2	4.8
Average number of non-silent mutations per plasmid	4.6	7.9	7	7.4
Average number of non-silent mutations per kb	2.1	3.6	3.1	3.3
Average number of silent mutations per gene	2.4	2.4	2.4	3.3
Average number of silent mutations per kb	1.1	1.1	1.1	1.5
Highest number of mutations found among mutants	11	21	22	17
Lowest number of mutations found among mutants	6	0	2	7

Table 8: Comparison with informations from the GeneMorph II kit (Agilent).

N° Tour	Round 1	Round 2	Round 3	Round 4	Mutazyme II	Mutazyme I	Taq DNA polymerase
Transitions (A<->G, T<->C) (%)	46.4	43.6	52.4	50.4	43	54	41.2
Transversions (A<->T, A<->C, T<->G, G<->C) (%)	53.6	56.3	47.6	49.6	51.4	44.1	54.1
Ts/Tv ratio	0.86	0.77	1.1	1	0.9	1.2	0.8
AT to N (%)	73.9	66.3	69.7	73.9	50.7	25.6	75.9
GC to N (%)	26.1	33.7	30.3	26.1	43.8	72.5	19.6

The mutated plasmids were transformed in *E. coli* DH5α, recovered and pooled to be transformed in *E. coli* BL21 DE3* for screening. The bubbles were clearly visible after 3 hours of incubation (Figure 13). A total of 28,892 colonies were screened and 1,380 mutants were selected for their ability to produce polymer in the screening conditions (Table 9).

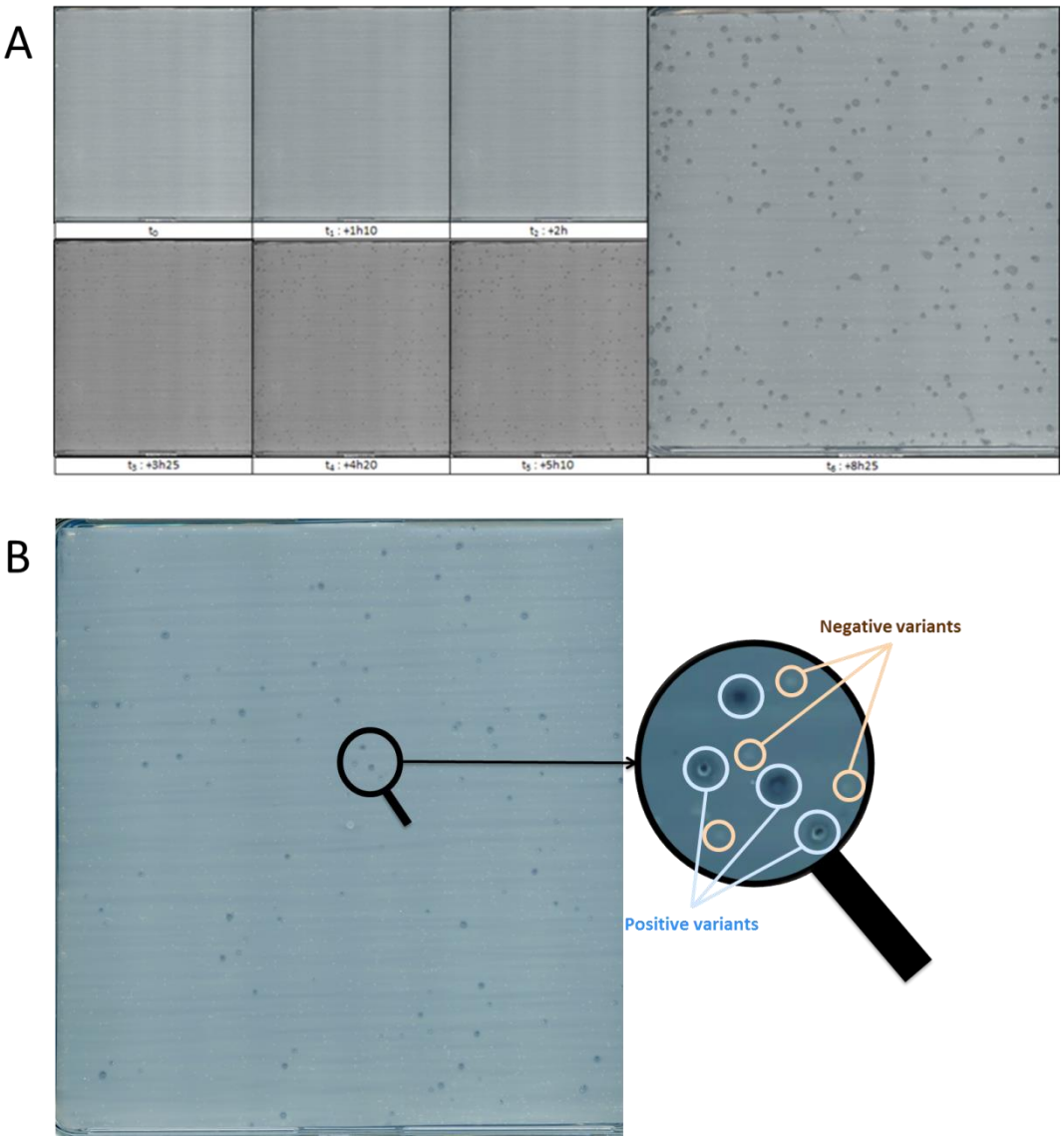


Figure 13: (A) Following bubble formation of round 3 screening.
(B) Zoom on round 4 screening.

Table 9: Data on the colonies screened

N° round	Round 1	Round 2	Round 3	Round 4
Total number of colony screened (cfu)	4,063	8,537	8,636	7,656
Positive colony (%)	25	17.5 (2) 12 (2 bis)	9.6	5.3
Number of picked colonies Total number of picked colonies	188 188	386 574	400 974	406 1,380

Among the positive hits, around 12 were randomly chosen to be sequenced and two negative hits were also sequenced for rounds 2, 3 and 4 (Table 10). Notably, wild type genes were recovered in each round. However, their number decreases from seven (round 1) to four (round 4) out of 12 mutants tested. This result comes from the fact that the screen did not exclude wild type and that we used a low mutagenesis rate between 0 and 4.5 mutation per kb, which can sometimes result in no

mutation of the template. In addition, we can notice that the percentage of positive colonies is inversely correlated with the mutation rate (Figure 14). Moreover, the average mutation number is higher in the in negative mutants than in the positive mutants (11.25 against 0.58 for the round 2). This indicates that a high number of mutations is on average deleterious to activity and/or polymer formation ability.

Table 10: Sequencing results of 12 randomly taken positive hits (after screening)

N° round	Round 1	Round 2	Round 3	Round 4
Percentage of wild type (include silent mutations) (%)	64 (7/11)	75 (9/12)	16.7 (2/12)	33 (4/12)
Average number of mutations per gene	1.25	0.58	2.42	3.42
Average number of mutations per kb	0.56	0.26	1.09	1.54
Average number of non-silent mutations per gene	0.58	0.5	1.5	2.25
Average number of non-silent mutations per kb	0.26	0.22	0.67	1.01
Average number of silent mutations per gene	0.67	0.08	0.92	1.17
Average number of silent mutations per kb	0.3	0.04	0.4	0.52
Average mutation number per gene except wild type genes	2.1	2.3	3.1	5.6
Maximum mutation number	3	3	5	8
Minimum mutation number	0	0	0	0
Average number of mutation per gene for negative mutants	n.d.	11.25	8	9.5

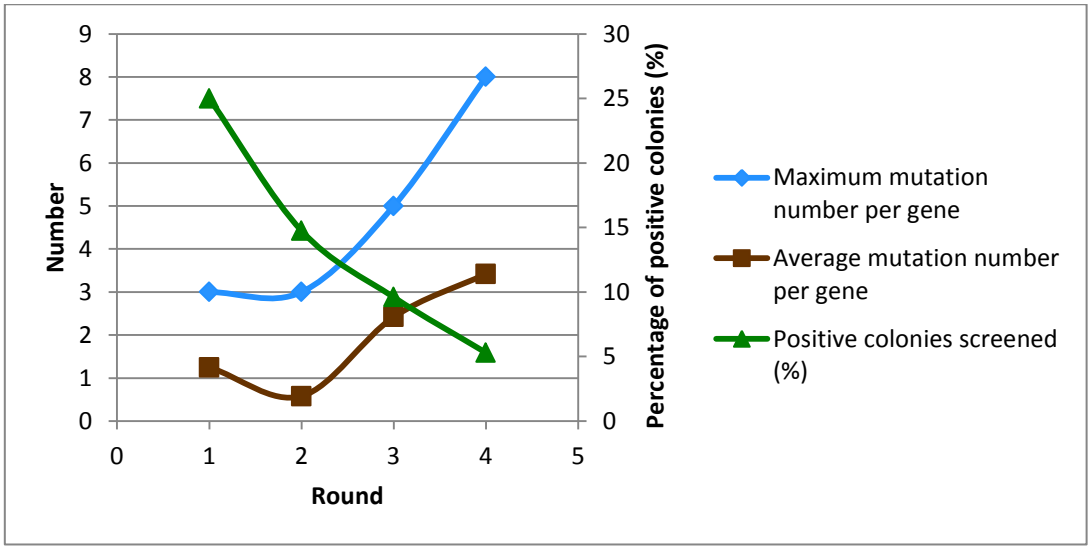


Figure 14: Evolution of the percentage of positive colonies screened (green triangles), the maximum (blue diamond) and the average (brown square) mutation number per gene.

Screening of the libraries

Thus, we decided to produce the 1,380 isolated polymer-forming variants in liquid medium and then characterize their properties. The recombinant cells were grown in deep well plates, lyzed and centrifuged. The culture supernatant was recovered for activity measurement in the following conditions: 30 °C (reference), 60°C (high temperature), 50°C and 30% of DMSO (both high temperature and DMSO concentration) and 30°C 50% DMSO (high DMSO concentration). All 1,380 variants were screened. From data analysis, 33 positive hits and 6 negative hits were chosen (see Table S2 for sequencing results). However, the screen was found to be not reproducible because a second end-point activity measurement of the hits in duplicate showed a very bad repeatability between both the results of the screening and the duplicates. The variability obtained by activity determination of 91 wild type enzymes was 37%, 24%, 26% and 56% at 30°C, 60°C, 50°C 40% DMSO and 30°C 50% DMSO respectively. Thus our liquid screening method was not adapted. This probably results from heterogeneous production level in deep well correlated to the very variable OD determined after cell growth.

Hits characterization

The hits were sequenced (Table S2), produced in deep well plate and purified by HisMultiTrap microplate. All the mutants were active on sucrose, with a specific activity between 12% and 115% of relative activity compared to the wild type (Figure 15). Specific activity was determined only one time for each mutant due to low amount of pure enzyme. The melting temperature was determined using DSF (Figure 16). As for the wild type enzyme, two peaks were distinguished on the DSF curve. The wild type melting temperature was 55.05°C (for the higher peak) and five variants (8, 15, 1, 24 and 9) revealed slightly improved activities, the best T_m improvement corresponding to 1.3°C increase compared to the wild type enzyme. In the presence of DMSO, the best mutants N°24 and 8 showed T_m values of 53.85°C and 54.75°C, respectively *versus* 53.32°C for the wild type. Again, the improvement was modest (1.4°C). Notably, the improved mutants identified were almost the same in both conditions but in a different order: 9-24-1-15-8 without DMSO and 8-24-22-9-15-26 with DMSO.

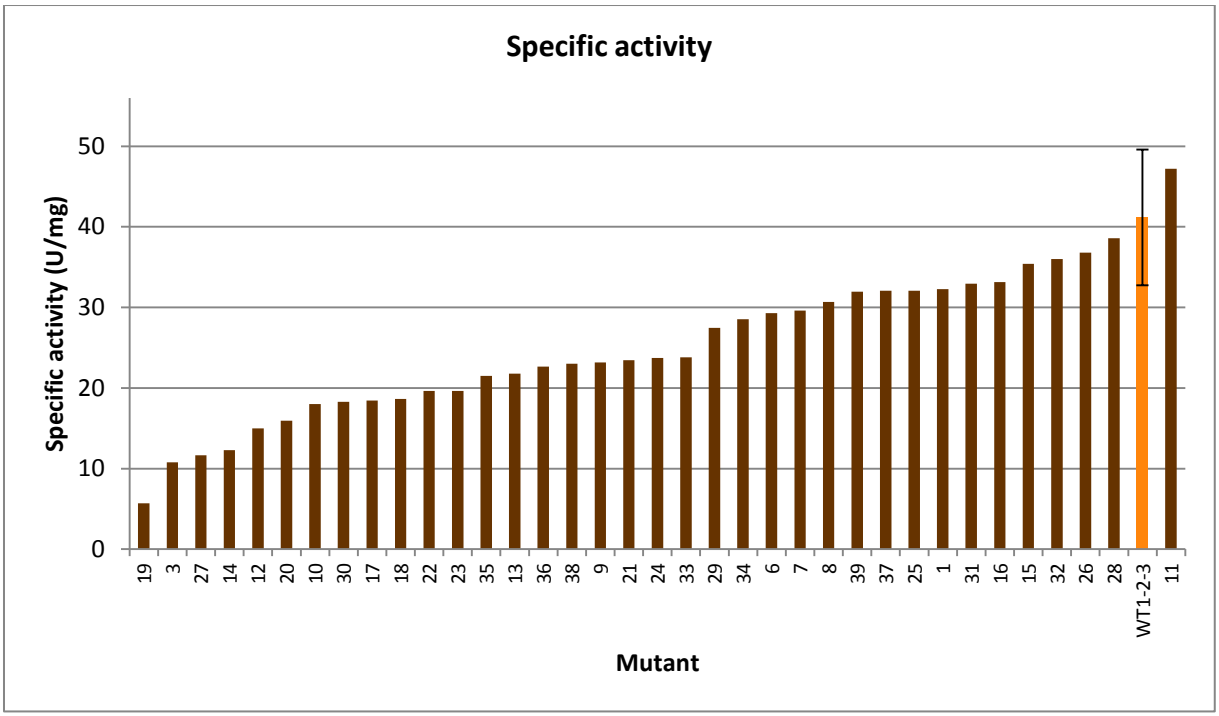


Figure 15: Hits specific activity (U/mg) at 30°C with 292 mM sucrose and 50mM sodium acetate buffer.
Triplicate for the wild type; only one measurement for the variants.

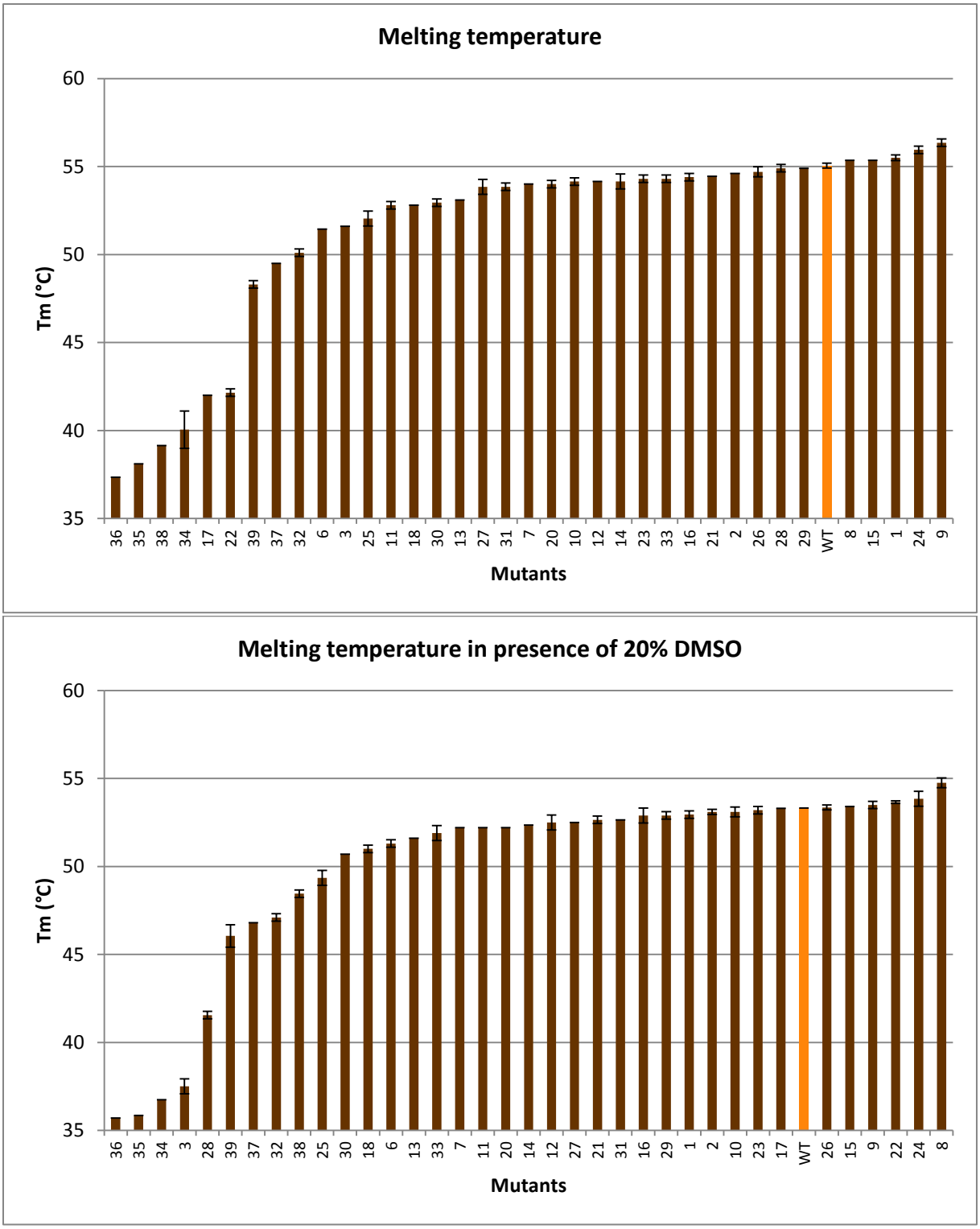


Figure 16: Melting temperature determination of the hits without DMSO and with 20% (v/v) DMSO

The low amount of pure enzyme obtained using HisMultiTrap purification did not allow further characterization. The best hits identified by DSF should be produced in flask and their half-life time as well as T_m will be determined in order to confirm or infirm the improvement.

Conclusion

Our rational strategy of “salt-bilization” was not successful. However, we still need to pursue the experiments in order to finish mutant characterization, in particular the ones with disulfide bonds. To improve salt-bilization strategy, it could be coupled to computational calculations to allow the evaluation of the effect that will trigger the mutation of the residues to form ionic interactions; maybe the mutation will be more detrimental to the stability and thus unwanted effect. This could be evaluated by Gibbs free energy calculations.

Our random strategy based on a neutral drift approach was not successful, maybe more rounds of mutagenesis should have been done or maybe our libraries were too small. We also need to optimize the second screening as the variability was found to be too high in these conditions for the wild type enzyme only. We could have purified the variants using a heat shock to remove most host proteins, because enzyme stability is very dependent on protein concentration and we observed production growth variation in deep well plates. Instead of an end-point activity measurement at high temperature, the activity would have been determined before and after a heat shock at 30°C and higher temperature.

Additionally, in our case, the presence of two melting temperatures observed during DSF analysis is still unclarified. Maybe, the presence of more interactions near the catalytic center of the enzyme will not be visible by DSF because of enzyme size whereas these interactions may stabilize the enzyme during the reaction. We need to perform more than DSF to evaluate accurately the stability of such large enzyme, doing half-live times and temperature optimum determination. Further, the increase in thermostability described using mutagenesis is generally between 2 and 15°C improvement of the melting temperature. This does not seem a large improvement but a little change in melting temperature can trigger more than a 10-fold longer life-time (Wijma et al., 2013). Thus, DSF was maybe not a good screening method in our case. Nevertheless, the hits obtained deserve further characterization to evaluate their stability comparatively to the wild type. Moreover, the bank created could be also screened for new properties *e.g.* change in substrate specificity or product analysis or serve as basis for other round of mutagenesis.

Large enzyme are difficult to stabilize (Eijsink et al., 2004) as a mutation away from the catalytic center may not be sufficient to protect the catalytic core from heat-induced effects (Zhang et al., 2016). Our work is consistent with these observations; more efforts need to be done to succeed glucansucrase stabilization.

Experimental procedures

Random bank construction and solid screening

Random mutants were constructed using GeneMorph II Random Mutagenesis kit (Agilent), pET53-*asr*-Δ2 gene as a template, forward primer 5'-GATAGTTTTGTGAAGACGCAAGCTAATTGG and reverse primer 5'-GCCCAGTCTTTCAAAACATAATAAGCACC. After PCR amplification and overnight digestion at 37°C with *Dpn*I, mutated genes were recombined with the vector previously linearized with forward primer 5'-GGTGCTTATTATGTTTTGAAAGACTGGGC and reverse primer 5'-CCAATTAGCTTGCGTCTTCACAAAACTATC with NEBuilder DNA Assembly kit using a ratio vector:insert of 1:2 and ligation product was transformed in home-made TSS (10 $g.L^{-1}$ tryptone, 5 $g.L^{-1}$ Yeast extract, 10 $g.L^{-1}$ NaCl, 100 $g.L^{-1}$ PEG (M=3350 g/mol), 5% (v/v) DMSO, 50 mM $MgCl_2$) competent *E. coli* DH5α cells. 100 µL of the transformation product was spread on agar plate to evaluate bank size. Mutated plasmids were extracted after O/N culture at 37°C in LB medium supplemented with ampicillin 100 $mg.mL^{-1}$ using NucleoBond Xtra Midiprep kit (Macherey-Nagel). Mutated plasmids were then transformed in home-made TSS competent *E. coli* BL21 DE3* cells and spread on Durapore PVDF 0.22 µm membrane (Sigma) positioned on Q-tray containing LB agar medium supplemented with ampicillin 100 $mg.mL^{-1}$. After O/N incubation at 37°C, the membrane was transferred on Q-tray containing LB agar medium supplemented with ampicillin 100 $mg.mL^{-1}$ and IPTG 1 mM at 21°C. After 24 hours incubation, membrane was transferred on Q-tray containing 20 % (v/v) DMSO, 50 mM sodium acetate buffer pH 5.75 and 50 $g.L^{-1}$ sucrose. After 8 hours of incubation at 40°C, between 200 and 400 positive variants *i.e* colonies surrounded with a bubble of polymer were manually resuspended in 96-wells microplate in LB medium supplemented with ampicillin 100 $mg.mL^{-1}$ and 8% glycerol. Microplates were incubated at 37°C 600 RPM O/N and then conserved at -80°C. A print of each microplate was taken on Q-tray or OmniTray containing LB agar medium supplemented with ampicillin 100 $mg.mL^{-1}$. After O/N incubation at 30°C, colonies were scrapped to seed a culture of LB supplemented with ampicillin 100 $mg.mL^{-1}$ which was incubated 4 hours at 37°C. Positive mutated plasmids were recovered by Nucleobond Xtra Midi kit (Macherey Nagel) and this mix was used as template for another round of error-prone PCR. All these steps (epPCR, *Dpn*I digestion, recombination, transformation in *E. coli* DH5α, Midiprep, transformation in *E. coli* BL21 DE3*, screening, positive hit selection and organization in microplate) were carried out in six days.

Production in 96 deep well plates and liquid screening

The microplates containing the positives variants were thaw, and used to seed new microplates containing LB supplemented with ampicillin 100 $mg.mL^{-1}$. After overnight incubation, 96 deep well microplates containing 1 mL of optimized ZYM medium were seeded with 10 µL of the preculture. After 26 hours of culture at 21°C, the plates were centrifuged 15 minutes at 2,000 g 10°C and the supernatant was discarded. 300 µL of Lysis buffer (0.5 $g.L^{-1}$ lysozyme, DNAseI, 50mM sodium acetate buffer pH 5.75) was added in each well to resuspend the pellet. Cells were then frozen at -80°C, at least for one day.

Five deep wells can be screened in one day. Deep well plates were thaw and vortexed to lyze the cells. The crude extract was then centrifuged 20 minutes at 2,250 g. The supernatant was transferred in another deep well using TECAN® liquid handler, diluted ten times and kept on ice. 20 µL of diluted supernatant was dispensed in microplates. Then, 80µL of a 1.25X concentrated mix was dispensed in each well. The concentrated mix contained either (i) 125 $g.L^{-1}$ sucrose and 62.5 mM sodium acetate buffer pH 5.75 for the two conditions at 60°C and 30°C without DMSO; (ii) 125 $g.L^{-1}$ sucrose, 62.5 mM sodium acetate buffer pH 5.75 and 50% DMSO (w/v) for the condition at 50°C with 40% DMSO or (iii) 125 $g.L^{-1}$ sucrose, 62.5 mM sodium acetate buffer pH 5.75 and 62.5% DMSO (w/v) for the condition at 30°C with 50% DMSO. The microplates containing the enzymatic extract with the reaction mix were incubated at different temperatures during 20 minutes under 600 rpm stirring. Enzymatic reactions were stopped by 100 µL DNS addition and the OD at 540 nm was measured to evaluate the end-point activity (Miller, 1959).

Mutagenesis study

Mutants were constructed by inverse PCR using the pET53-*asr*-Δ2 gene as a template, Phusion® polymerase (NEB), and the primers described in Table S1. Following overnight *Dpn*I (NEB) digestion, the PCR product was transformed into competent *E. coli* DH5α and clones were selected on solid LB medium supplemented with ampicillin 100 $µg.mL^{-1}$. Plasmids were extracted with the QIAGEN spin miniprep kit and mutated *asr* genes were checked by sequencing (GATC Biotech). Mutants were produced and purified as described above.

Synthetic fragments were order at TWIST® and recombined with NEBuilder DNA Assembly.

Chimera design and construction

Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) was used to align ASR and GBD-CD2 sequence and Genome Compiler (http://www.genomecompiler.com/) was used to construct and manipulate plasmid cards.

Domain C of GBD-CD2 was amplified by PCR using pET53-ΔN_{123}-gbd-cd2 as template, Phusion® polymerase (NEB), forward primer AAAGCTAGAATAAAATATGTTAGCGGTGGTCAAACTATGAGTG and reverse primer TGTACGAGCATCTTGTGAATCTGATGCACCCACAGGCAC. The underlined sequence corresponds to the overlap with asr-Δ2 gene. Similarly, domain C of ASR was amplified by PCR using pET53-asr-Δ2 as template, forward primer ACGGCTCGTAAAAGCTATGTCTCTGGTGGGCAAACAATG and reverse primer TGTACGGGCATCTTGATTGTCACTAGCTCCAACTGGCACC. Plasmids of pET53-asr-Δ2 and pET53-ΔN_{123}-gbd-cd2 were amplified without the domain C with the following primers: GATTCACAAGATGCTCGTACAGTG and AACATATTTTATTCTAGCTTTGAGCAACGC for pET53-asr-Δ2 and GACAATCAAGATGCCCGTACAG and GACATAGCTTTTACGAGCCGTC for pET53-ΔN_{123}-gbd-cd2. PCR products were purified with GenElute PCR Clean-up kit (Sigma) and DNA was quantified using a NanoDrop instrument. Assemblage of asr-Δ2 plasmid with ΔN_{123}-gbd-cd2 domain C or ΔN_{123}-gbd-cd2 plasmid with asr-Δ2 domain C was made using NEBuilder HiFi DNA Assembly kit with a molar ratio vector:insert of 1:2. Ligation product was transformed in home-made *E. coli* DH5α competent cells and sequences of the constructions were checked by sequencing (GATC).

Production and purification of Bank hits, rational mutants and wild type enzyme

The *E. coli* BL21 DE3* strain was used for enzyme production. A preculture of transformed *E. coli* BL21 DE3* in LB medium supplemented with ampicillin 100 µg.mL^{-1} was used to inoculate a culture at an OD_{600nm} of 0.05 in ZYM-5052 auto-inducible medium (Studier, 2005) modified by supplementation with 100 µg.mL^{-1} ampicillin, 1% (w/v) α-lactose, and 1% (w/v) glycerol for pET53 enzyme production. After 26 hours of growing at 21°C, cells were harvested by centrifugation and resuspended in Binding buffer containing 20 mM phosphate buffer, 20 mM Imidazole (Merck Millipore), and 500 mM NaCl, pH 7.4 supplemented with EDTA-free anti-protease tablets (Roche). Cells were disrupted by sonication and debris was removed by a centrifugation step at 45,000 g for 30 minutes at 8°C. Purification was performed with the ÄKTA Xpress system (GE Healthcare). Two-step purification was performed in a cold chamber at 8°C using (i) a HisTrap HP 1mL column (GE Healthcare) for the affinity step and (ii) a Superose12 16/60 (GE Healthcare) for the size exclusion step, or a HiPrep desalting 26/10 column (GE Healthcare) for desalting. The size exclusion step was performed upstream of crystallization trials and Differential Scanning Fluorimetry assays, and protein

was eluted in MES buffer pH 6.5 at 30 mM with 100 mM NaCl and 0.05 $g.L^{-1}$ $CaCl_2$. The desalting step was performed for biochemical characterization, for which protein was eluted in 50 mM sodium acetate buffer pH 5.75. Purified fractions were pooled together and concentrated using AmiconUltra-15 with a cut-off of 50 KDa to 10-15 $mg.mL^{-1}$. Purification was checked by SDS-PAGE electrophoresis using NuPAGE 3-8% Tris-Acetate protein gels (Invitrogen), and protein concentration was assessed by spectroscopy at 280 nm using a NanoDrop instrument. The theoretical molecular weight and molar extinction coefficient of the enzyme were calculated using the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

200 randomly taken mutants of the fourth round of mutagenesis were purified using HisMultiTrap HP (GE Healthcare) and Zeba spin desalting plates (Thermofischer).

Production and purification of chimera and wild type enzymes

Chimeras and their relative wild type were produced using chaperones. Thus, the *E. coli* BL21 DE3* strain transformed with a plasmid harboring the Tf16 chaperone gene (Takara) was used for enzyme production. A preculture of transformed *E. coli* BL21 DE3* in LB medium supplemented with ampicillin 100 $\mu g.mL^{-1}$ and chloramphenicol 20 $\mu g.mL^{-1}$ was used to inoculate a culture at an OD_{600nm} of 0.05 in ZYM-5052 auto-inducible medium (Studier, 2005) modified by (i) supplementation with 100 $\mu g.mL^{-1}$ ampicillin and 20 $\mu g.mL^{-1}$ chloramphenicol, 4 $mg.mL^{-1}$ arabinose, 1% (w/v) α-lactose, and 1% (w/v) glycerol for ASRΔ2 and ASRΔ2 chimera production; or (ii) supplementation with 100 $\mu g.mL^{-1}$ ampicillin and 20 $\mu g.mL^{-1}$ chloramphenicol, 4 $mg.mL^{-1}$ arabinose, 0.75%(w/v) α-lactose, 1.5%(w/v) glycerol and 0.05% glucose for ΔN_{123}-GBD-CD2 and ΔN_{123}-GBD-CD2 chimera production (conditions optimized previously (Vuillemin et al., 2014)). The protocol is then the same than described above for ASRΔ2. For ΔN_{123}-GBD-CD2, after 24 hours at 23°C, cells were harvested by centrifugation and resuspended in 50 mM sodium acetate buffer pH 5.75. Final OD_{600nm} were 80 and 30 for ASRΔ2 and ΔN_{123}-GBD-CD2 respectively. Cells were disrupted by sonication and debris was removed by a centrifugation step at 45,000 g for 30 minutes at 8°C. Purification was performed with ÄKTA Xpress system (GE Healthcare). 20 mM imidazole and 500 mM NaCl were added to the soluble fraction prior to the purification of ΔN_{123}-GBD-CD2 and its chimera. Two-step purification was performed in a cold chamber at 8°C using (i) a HisTrap HP 1mL column (GE Healthcare) for the affinity step and (ii) a Superose12 16/60 (GE Healthcare) for the size exclusion step, and protein was eluted in MES buffer pH 6.5 at 30 mM with 100 mM NaCl and 0.05 $g.L^{-1}$ $CaCl_2$. Purified fractions were pooled together and concentrated using AmiconUltra-15 with a cut-off of 50 KDa to 10-15 $mg.mL^{-1}$. Purification was checked by SDS-PAGE electrophoresis using NuPAGE 3-8% Tris-Acetate protein gels (Invitrogen), and protein concentration was assessed by spectroscopy at 280 nm using a NanoDrop instrument. The

theoretical molecular weight and molar extinction coefficient of the enzyme were calculated using the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

Activity measurement

Activity was determined in triplicate at 30°C in a Thermomixer (Eppendorf) using the 3,5-dinitrosalicylic acid method (Miller, 1959). 50 mM sodium acetate buffer pH 5.75, 292 mM sucrose and 0.05 $mg.mL^{-1}$ (ASRΔ2, its chimera and its variants) or at 0.002 $mg.mL^{-1}$ (ΔN_{123}-GBD-CD2 and its chimera) of pure enzyme were used. One unit of activity is defined as the amount of enzyme that hydrolyzes 1 µmol of sucrose per minute.

Enzymatic reactions and product characterization

ASRΔ2 and ASRΔ2 chimera polymer productions were performed using 1 $U.mL^{-1}$ of pure enzyme with 292 mM sucrose in 50 mM NaAc buffer pH 5.75 at 30°C over a period of 24 hours. ΔN_{123}-GBD-CD2 and ΔN_{123}-GBD-CD2 chimera branching reactions were performed using 1 $U.mL^{-1}$ of pure enzyme with 292 mM sucrose and 309 mM of 70 kDa dextran in 50 mM NaAc buffer pH 5.75 at 30°C over a period of 24 hours. These polymers were purified by dialysis using 14 kDa cut-off cellulose dialysis tubing (Sigma-Aldrich) in water and analysed by NMR. NMR samples were prepared by dissolving 10 mg of purified product in 0.5 mL D_2O. Deuterium oxide was used as the solvent, and sodium 2,2,3,3-tetradeuterio-3-trimethylsilylpropanoate ($TSPD_4$) was selected as the internal standard (δ^1H = 0 ppm, $\delta^{13}C$ = 0 ppm). ^1H spectra were recorded on a Bruker Avance 500-MHz spectrometer operating at 500.13 MHz. The data were processed using TopSpin 3.0 software. 1D ^1H NMR spectra were acquired by using a zgpr pulse sequence (with water suppression). Spectra were performed at 298 K.

Enzyme melting temperature determination

Differential Scanning Fluorimetry experiments were performed with 7 µM of pure enzyme in 50 mM sodium acetate buffer pH 5.75 supplemented with 0.05 $g.L^{-1}$ of calcium chloride and 10 X of SYPRO orange (Life Technologies). A ramp from 20 to 80°C was applied with 0.3°C increments at the rate of 0.3°C per second on a C100 Thermal Cycler.

Structural analysis of the enzymes

To analyze the network of interactions in the structure, the RING (Piovesan et al., 2016) web server was used (Residue Interaction Network Generator). It identifies and lists all the non-covalent interactions in the PDB loaded (hydrogen bond, Van der Waals, Ionic, π-π stacking, π-cation).

Supporting information

A

β-sheet	Longest
1	GTF-SI
2	GTF-SI
3	All except DSR-M
4	GTFA
5	ASR
6	ASR
7	All except DSR-M and GTF-SI
8	GTFA
9	Only in ASR
10	Only in ASR
11	ASR, GBD-CD2
12	ASR

B

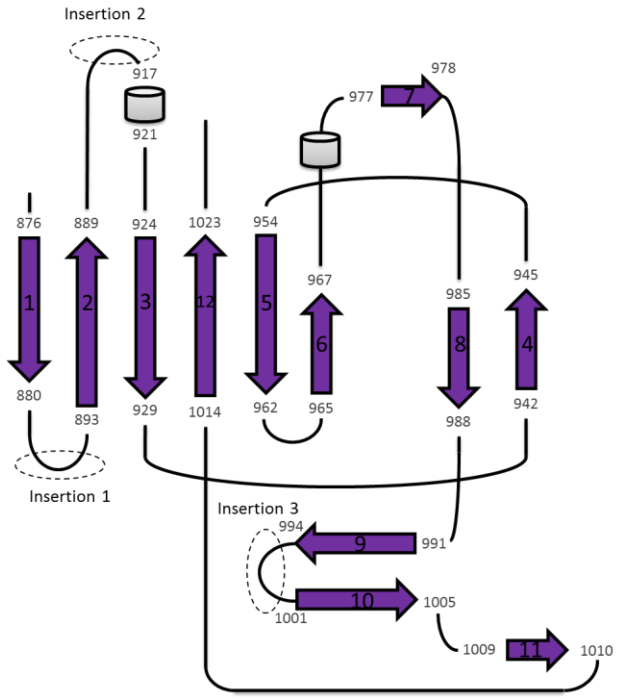


Figure S1: (A) Comparison of β-sheet size between all available sucrose-using GH70 family structure Comparison of the domain C of available GH70 family structures. For β-sheet nomenclature see B. (B) Schematic representation of the structural elements of ASR domain. Cylinder: α-helix (grey), arrow: β-sheet (purple).

Table S1: Primers used for "salt-bilization" and disulfide bridge mutant constructions

Mutation	Forward primer 5'-> 3'	Reverse primer 5' -> 3'
N750D	CAGAATTGCTGATTCAGGTGATACAAAAG	GTATCACCTGAATCAGCAATTCTGTTAACC
H1049R	GTTCTGTATTCCGTTCAAATGCTGCATTAG	CAGCATTTGAACGGAATACAGAACCATC
P1073K	CCGACTTCTAAGGAGCAAAGTACAAATG	TACTTTGCTCCTTAGAAGTCGGCATC
N1209D	CTCTATGTTGCTGATACTAAAAGTAGTGGTGTGG	CCACTACTTTTAGTATCAGCAACATAGAGTAAGTTGAC
N624D	TACAGGTAATGATGACAATGCGAATTTTG	TTCGCATTGTCATCATTACCTGTAATACTACC
E1157K	GTAATGCAATAAAGGCATTACACAAAAACG	TTGTGTAATGCCTTTATTGCATTACGTAAATC
A330C	GTAAAGATGTTTGCGGAAATACATTTTATTTTGACG	ATGTATTTCCGCAAACATCTTTACTAAATGCATTC
A1410C	TAACGATGATTGTAATAATTTTTATCACTTAAATG	GATAAAAATTATTACAATCATCGTTAATGTATTTATC
A408C	CTCCTCACAATTGCGCAAAGTCTTATG	AGACTTTGCGCAATTGTGAGGAGTATTTTC
N1301C	TACAGTTGCAGTGTAAAGATGCGCAAAC	TGCGCATCTTTACACTGCAACTGTAGTG
Q545C	AGTGGCTTTGTGGGGGATTCCTAG	GAATCCCCCACAAAGCCACTG
N1266C	GAATGGGACCTGTATTTTACATCGAGGTG	CCTCGATGTAAAATACAGGTCCCATTCATATATTTAGC
L865C	TCAGCGTTGTGCAAAGCTAGAATAAAATATG	TTCTAGCTTTGCACAACGCTGAAATGAC
L1091C	TATTTAAAGAATGTGGTATTACTAGTTTTGAGTTAGCACC	TAGTAATACCACATTCTTTAAATAAGTTAGCCTTTGTTGC
G874C	ATATGTTTCTTGTGGGCAAACAATGG	TTGTTTGCCCACAAGAAACATATTTTATTC
F900C	CAAGTGTTCGATGTGGTAAAGGAATTATGAC	TTCCTTTACCACATCGAACACTTGTTAAC

Table S2: Neutral drift bank hits sequencing

Variant	Round	Non-silent mutations	Hit
1	1	K717E	positive
		Y778F	
		N965H	
2	1	A825T	positive
		M867T	
		G927R	
3	1	Q662H	positive
		V890I	
		T939S	
		M1075V	
		Y1223C	
4	1	V602F	positive
		Q647K	
		R674G	
		S1003L	
		D1162N	
5	1	S581T	positive
6	3	T544I	positive
		D848N	
		V982A	
		V984M	
7	3	N1097I	positive
8	3	S675C	positive
		N748H	
		A911T	
		D1162N	
9	3	P701L	positive
		L758V	
		L1094I	
		K1182Q	
10	4	Q662L	positive
		P821A	
		V921I	
11	4	T660I	positive
		D780N	
		A1120V	
12	4	D1135G	positive
13	4	A625T	positive
		N626K	
		S661T	
		K902T	
		A1170G	
14	4	S785N	positive
		D851A	
		K1123E	
		G1125D	
		V1204I	
15	4	K717T	positive
		Q879R	
		L922V	
16	4	I558V	positive
		N626Y	
		K1222R	
17	4	D1162Y	positive
		P1196T	
18	4	M789L	positive
		N820T	
		D895N	
		V957G	
		G1142V	
19	4	I538T	positive
		G584C	
		F1058L	
20	4	N877K	positive
		I963V	
		G1024C	
21	4	K1200I	positive
22	4	K1051I	positive
23	4	G744C	positive
		Q750R	
		A1120V	
24	4	D960N	positive
		P1035H	
		A1117V	
25	4	F755Y	positive
26	4	N626Y	positive
		Q643L	
		K832R	
		D960N	
27	4	Q662L	positive
		P821A	
		V921I	
28	4	I963T	positive
		A1030T	
29	4	A616T	positive
		A786T	
30	3	I564V	positive
		D734E	
		T792S	
31	4	D697V	positive
		H951L	
		K969T	
32	4	N672Y	positive
		N820K	
		N877I	
		L920I	
		Y1072C	
		A1170V	
33	4	K969E	positive
		D1006E	
34	4	V599F	negative
		K608I	
		K698N	
		P735T	
		E763D	
		I928L	
35	1	D524G	negative
		P720R	
36	1	D524G	negative
		P720R	
37	3	F753V	negative
38	4	R519Q	negative
		L906M	
		D1162E	
		I1210N	
39	4	D546E, N1171K	negative

Conclusion

&

Prospects

All along this manuscript, our objective was to deepen the understanding of the **mechanism** of the alternansucrase (ASR) from *Leuconostoc citreum* NRRL B-1355 in terms of specificity, stability and ability to produce high or low molar mass alternans.

The ASR clearly stands apart among glucansucrases, due to the presence of **unusual sequence** stretches which place it on an independent branch of the phylogenetic tree of GH70 sucrose active enzymes (Figure 1) and are likely involved in its unique **specificity and stability**. Indeed, its sequence is longer (2,057 residues) compared to the average size of the other glucansucrases, in part because of the presence of seven APY repeats in the C-terminal domain V of the enzyme. These APY repeats do not affect the biochemical properties of the enzyme and their function in GH70 family enzyme are still unraveled. In addition, ASR is the only glucansucrase to possess the signature "**YDA**" following the TSS in the conserved motif IV, suggested to be an indicator of specificity. ASR is the only glucansucrase performing the synthesis of **alternated α-1,6 and α-1,3 linkages** and one of the **most stable** enzymes of the family GH70. Surprisingly, Blastp analysis revealed that there is no intermediates between the putative ASR sequences sharing very high identity (more than 97%) and the other glucansucrases: the next hit drops down to 61% of identity and correspond to DSR-M. Using only the ASR catalytic domain as query, again, the next hits following the putative ASR sequences showing more than 96% identity drop down to less than 57% identity with the catalytic domain of putative enzymes from *Lactobacillus* or *Oenococcus*. These observations clearly indicate that the ASR is a **very special enzyme** in the GH70 enzyme family.

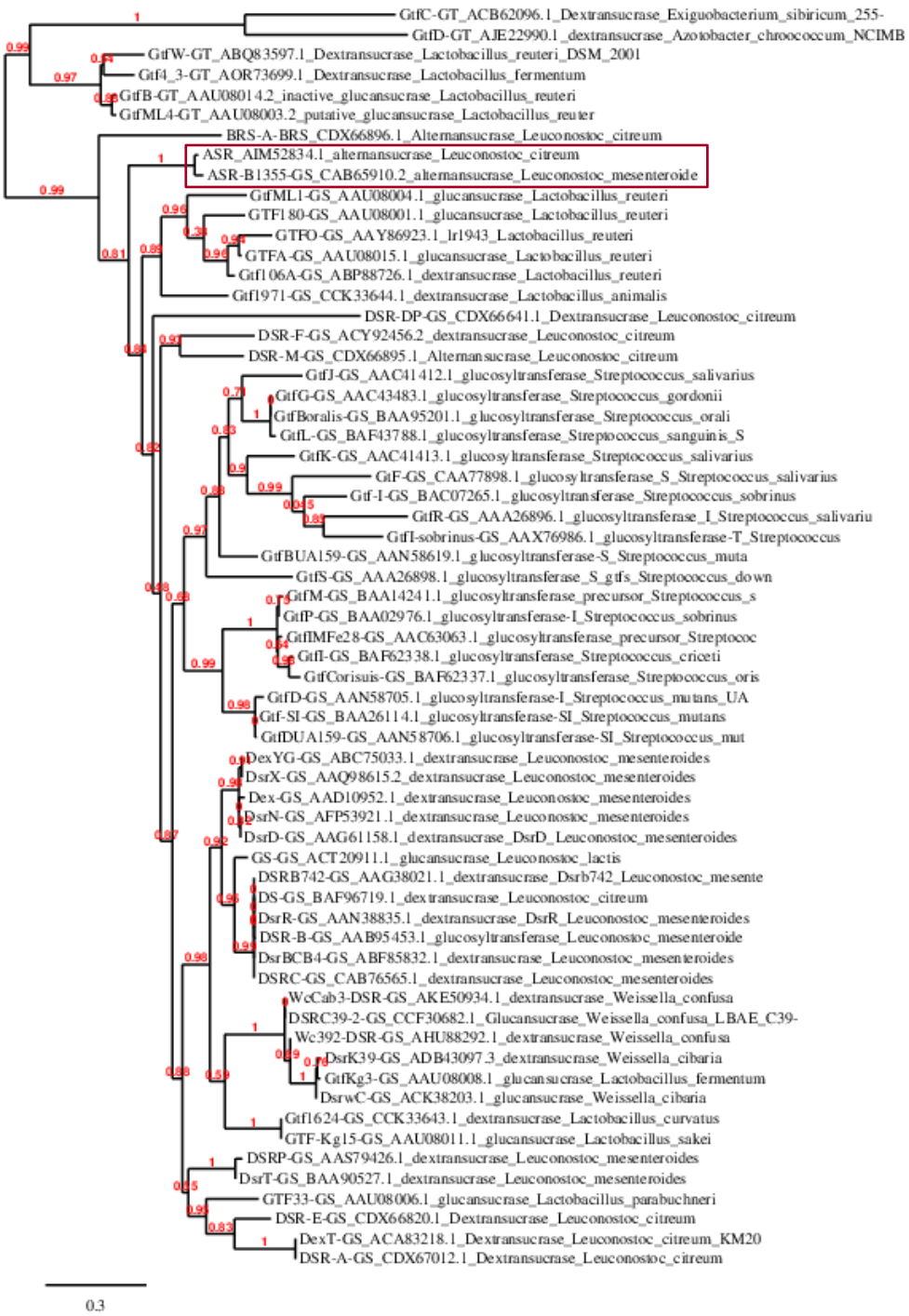


Figure 1: Phylogenetic tree of the 64 characterized sequences of GH70 in the CAZy database (December, 2018). Created using phylogeny.fr (Dereeper et al., 2008)

ASR is an **old** glucansucrase, the first proof of this intringuing activity dating from **1954**. The enzyme was placed at the core of two PhD thesis at LISBP. The first one was defended by Martha Arguello Morales in 2000 and aimed to clone, sequence and express the *asr* gene in *E. coli*. It was the first time that ASR was recombinantly produced. The second PhD thesis was defended by Gilles Joucla in 2006, and focused on the optimization of ASR production by construction and characterization of truncated variants. This work led to the **ASR-C-APY-del** variant that displayed an increased stability and the same specificity as the full-length enzyme. This truncated version was produced in much higher amounts in *E. coli*. Additionally, mutants were constructed to investigate the role of the particular sequence stretches identified in ASR motifs III and IV.

Despite these two studies and further works of Professor Côté and his collaborators, who characterized many ASR acceptor reaction products, the mechanism of alternated linkage synthesis remained ununderstood at the molecular level. In the same line, the molecular determinants at the origin of ASR stability were unknown. We absolutely needed to resolve the **three-dimensional structure** of ASR to address these issues and to get in hand useful data to rationally engineer the enzyme towards modulated specificity or improved stability. Following is a compilation of the main findings of our study (overview in Figure 2) with the corresponding perspectives (written in deep blue) that could be envisioned.

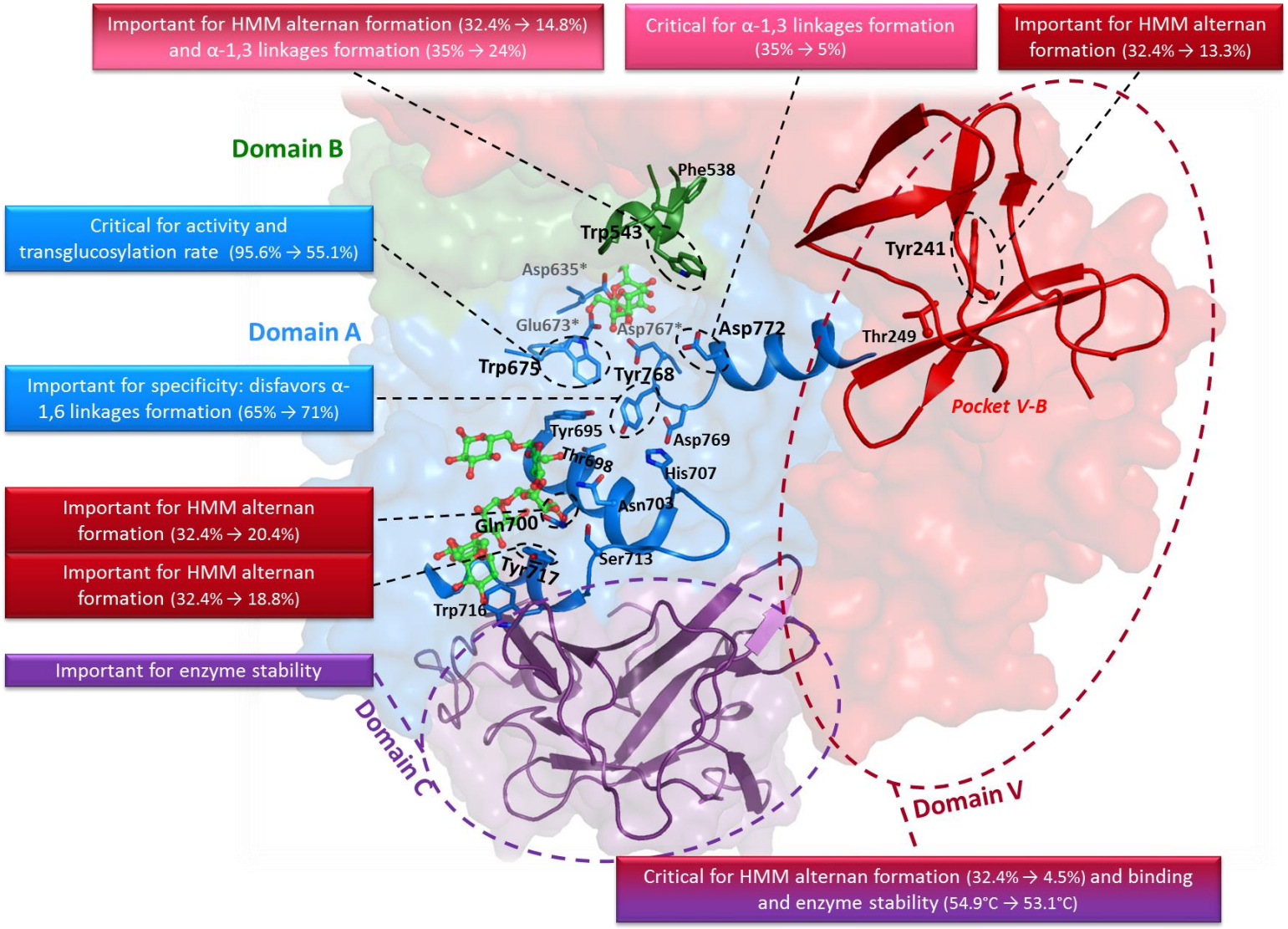


Figure 2: Overview of the 16 targeting residues (in sticks) in our work (chapter II and III) and domain C (purple cartoon, chapter IV) swapped with GBD-CD2.
Grey*: catalytic residues, not targeted. Sucrose manually docked from 3HZ3.

Unraveling the mechanism of alternated linkages

To resolve the 3D structure of the enzyme, crystallization campaigns were initiated using the truncated form ASR C-APY-del that was constructed and characterized by Gilles Joucla. Conditions were identified in which ASR C-APY-del **crystals** were obtained in around one month but their maximum diffraction achieved only 3,7 Å. To improve crystal diffraction, four N-terminal **truncated mutants** (ASRΔ2, ASRΔ3, ASRΔ4 and ASRΔ5) were constructed and produced. Good diffraction data were finally obtained for ASRΔ2. First, the 3D structure analysis showed that ASR adopts like the other glucansucrases a **U**-shaped fold made by **five** distinct **domains**. Like for DSR-M, the domain V fold of ASR, leaning toward the active site, confers a **horseshoe-shape** to the protein. Overall, ASR distinguishes from the other glucansucrases by **amino acid insertions** in both domains A and C resulting in longer **loops** on the enzyme surface. Comparison of the catalytic site of available glucansucrase structures with the new one of ASR motivated the rational construction of **18 mutants** that were biochemically characterized. The obtained results combined with molecular docking and maltose acceptor reaction monitoring enabled us to formulate a plausible scenario explaining the alternance mechanism.

We suggest that the linkage type is determined by the position of the acceptor in either the +2 or the +2' subsites located on either side of the +1 subsite. The +2 subsite is defined by Trp675, a very conserved residue among glucansucrases, and orients acceptor binding exclusively toward α-1,6 linkage synthesis. In fact, ASR can form consecutive α-1,6 linkages, but this reaction is not favored. Indeed, the other subsite comprises Asp772 and Trp543, two residues defining the +2' and +3' subsites respectively, which are critical for α-1,3 linkage formation right after an α-1,6 linkage. We submit that the interplay between these two acceptor sites controls linkage alternance.

Investigating the formation of high molar mass alternan

ASR crystals were soaked in the presence of different ligands. We obtained several complexes with panose, isomaltooligosaccharides and oligoalternans at a decent resolution.

In particular, the analysis of these complexes enabled us to identify a **sugar binding site** (SBS-A1), unique to ASR, in the catalytic **domain A**. This site comprises several residues binding the oligosaccharides through a network of interactions mediated by Asn703, Ser713, Trp716, Gln700 and Tyr717. Single mutant construction and characterization pinpointed the importance of Gln700 and **Tyr717** as their mutations severely reduced **HMM alternan formation** without affecting at all the synthesis of oligoalternans formed by acceptor reaction. The stacking with Tyr717 provides a major

contribution to sugar binding. Notably, this residue is **unique** to ASR, located near the interface between domains A and C and is rather far from the catalytic center at 24.9Å from Asp767 (TSS).

Furthermore, SBS-A1 accommodates oligoalternans of DP higher than 4 but not oligodextrans of the same size. We suggested that this intermediary site could drive a correct positioning between domain A and domain V for a specific elongation of alternan.

However, if SBS-A1 is important, the most striking contribution to polymer elongation is conferred by the **domain V**. Indeed, the truncation of domain V leads to an enzyme that is almost unable to produce **HMM polymer**. In addition, this enzyme does not **bind** either alternan or dextran polymer. This indicates that polymer binding in domain V is important for HMM alternan synthesis. In accordance with this observation, the analyses of the sugar complexes revealed the presence of isomaltooligosaccharides in a sugar binding pocket named V-B (showing structural homology with other sugar binding pockets previously described in the domain V of different GH70 sucrose-active enzymes). Rendering this pocket unfunctional through the mutation of a Tyr residue identified as important for sugar stacking clearly reduced HMM polymer formation. Another sugar binding pocket, named V-A was identified and is also suggested to participate in polymer binding and to play a role in polymer elongation.

Notably, these investigations enabled us to produce a stable and efficient variant -deleted of domain V and mutated in the SBS-A1 subsite- that is highly specialized for the production oligoalternans.

To go further in the comprehension of the enzyme mechanism, several additional experiments can be envisioned:

- **Soaking** experiments as well as **co-crystallization** should be pursued to try to obtain complex with a oligoalternans or oligodextrans of various size in the catalytic center active core. This could permit to better map the **subsites** of the protein. In particular, co-crystallization of ASRΔ5 isomaltooligosaccharides and oligoalternans of increasing size could be attempted.
- It would also be interesting to determine whether **Asp772** is also important for other putative alternansucrases and investigate the effect of Asp772 mutation to Ala in an ASRΔ5 template to evaluate domain V role for dextran elongation as Asp772Ala mutant mainly produces α-1,6 linked polymer.
- In addition, we did not investigate the mechanism of **branching** formation. We also have some interrogations on the **regularity** of the α-1,3 and α-1-6 linkage **alternance** in the polymer. To address these issues, more precise information on the **polymer structure** is required and would

necessitate the combination of NMR with methylation, partial acetolysis as well as enzymatic degradation for example. To this end, alternanase and debranching enzymes could be used but this would necessitate an in depth characterization of the specificity of these enzymes.

- Regarding the role of domain V, the construction of a single mutant in pocket A and a double mutant targeting both **pocket V-A** and **V-B** is planned to investigate their **individual** and/or **synergistic** roles in polymer binding, linkage specificity and polymer size.
- The putative **connection** between the catalytic center, the SBS-A1 subsite and the domain V would also deserve to be experimentally studied and or simulated using **molecular modelling techniques**. This is not easy to do without additional data. In particular, we would need to obtain complex with ligands enough long in order to help us to find a **possible path** connecting the various domains. Such complexes may not be obtainable using classical crystallization techniques: would Cryo-EM or XFEL techniques be of interest? This is an important question. We could adopt an alternative strategy consisting in replacing individually the aromatic residues at the surface of the protein with an alanine and examine the effect on the product profile.

What confers stability to ASR? How could we further improve it?

Our structural study of ASR allowed us to identify several possible structural determinants of its stability.

Firstly, the presence of several **amino acid insertions** is suggested to protect the enzyme from solvent penetration. Secondly, based on RING analysis (Piovesan et al., 2016), ASR is predicted to display more **ionic** and **π-π stacking** interactions than the following GH70 enzymes: GTF180, GTF-SI, GBD-CD2, GTFA, DSR-M. In particular, the **domain C** of ASR contains three amino acid insertions, longer β-sheets, more ionic interactions, more hydrophobic residues and π-π stacking interactions than the domain C of the other GH70 enzymes and notably that of the α-1,2 branching sucrase GBD-CD2. We performed domain C swapping between ASR and GBD-CD2 and the resulting chimera of ΔN_{123}-GBD-CD2 with ASR domain C gained in stability with a 1.5 fold improvement of its residual activity after 10 minutes incubation at 30°C. The other chimera (ASR with GBD-CD2 domain C) remains to be characterized.

To note, no specific functions have been assigned to domain C. From our preliminary work, we submit that it could contribute to ASR stability by reinforcing its pedestal position at the bottom of the U-shaped fold.

Experiments have to be pursued to confirm our preliminary results. In addition, we could **swap** the domain C of ASR with those of **other** GH70 enzymes, similarly to what was done with the branching sucrase. We could also try to introduce a simpler and a shorter linker connecting the **two arms** of the **U**-shape and examine the effect on the enzyme stability.

To further improve the stability of ASR, we propose to extend the number of **salt bridges** in the protein, what we called the "**salt-bilization**" strategy. From the analysis of the interaction network in various GH70 family structures, we thus attempted to introduce, in ASR, ionic interactions found in other glucansucrase structures. Introduction of **disulfide bridges** was also initiated based on the proposal provided by disulfide by Design 2.0 web server (Craig and Dombkowski, 2013). The aim was to link the **N-** and **C**-terminal **branch** of each domain to stabilize the overall scaffold. This strategy is still under progress and the characterization of these rationally designed mutants has still to be performed. If successful, we could envisioned to further extend the **connection** between the **two arms** forming the domains A, B, IV and V. We could support the construction of such a protein, using **molecular modelling**. Moreover, **genome mining** and **bioinformatics** analysis could be regularly performed to identify putative glucansucrases in highly **thermophilic** organisms that could be used to help the identification of stability determinants.

In parallel to this semi-rational work, we have also laid the fundations for **directed molecular evolution** of ASR. A screening assay was developed on solid medium to enable variant screening ASR in 20% DMSO at 40°C. A first library of around 4,000 variants (generated by epPCR) was screened in these conditions, and around 300 polymer-forming colonies were retained. The same cycle was repeated four times to finally keep 1,380 variants, which were further screened at higher temperatures with or without DMSO. Unfortunately, we did not obtain significant improvement but also figured out that our secondary screening assays were not sufficiently reproducible.

Our results highlight the fact that it is particularly difficult to stabilize such big and multi-domain proteins. However, alternative assays could be envisioned, which could be tested and applied to larger libraries of variants after validation:

- A more robust secondary screening assays based on **heat shock** and residual activity measurement in liquid medium could be used as secondary screen
- The variant libraries could be expressed in **organisms more thermostable** than *E. coli* to facilitate the discrimination between the wild type and improved mutants on solid medium

Finally, our work enabled us to extend our toolbox of engineered enzymes for tailor made α-glucans. For example, two enzymes, ASRΔ2-Asp772Ala and ASRΔ5-Tyr717Ala, could be really useful for prebiotic synthesis. They are almost as active and stable as the wild type enzyme and can produce high amounts of oligodextrans or oligoalternans, respectively (Table 1). To augment the indigestibility of oligoalternans and modulate their potential as prebiotics, we could imagine a process in which these oligosaccharides could be glucosylated using an **α-1,2** branching sucrase. In addition, ASR as well as the mutant that we have generated, could be tested to **glucosylate molecules** such as flavone or hydroxylated terpens, for example with the view of modifying their properties.

Table 1: Constructions with potential applications

Construction	Activity		Stability	Specificity		Product	
	Residual activity compared to ASRΔ2 (%, DNS method)	Hydrolysis (%, HPAEC-PAD)	ΔTm compared to ASRΔ2 (°C, DSF)	α-1,6 (%, NMR)	α-1,3 (%, NMR)	Residual ability to produce HMM glucan (%, HPSEC)	Average oligosaccharide size (g/mol, HPSEC)
ASRΔ2 (reference)	100 (29.9 U/mg)	4.4	0	65	35	100 (32.4%)	700 (≈DP5)
ASRΔ2 D772A-W543A	64	5.4	0	94	6	10	1000 (≈DP6-7)
ASRΔ5	80	5.3	-2	70	30	14	1900 (≈DP12)
ASRΔ5 Y717A	64	*n.d.*	-2.4	70	30	7	2200 (≈DP14)

In summary, our work has enabled the accumulation of very useful data on ASR. It also showed that even if there are common traits among GH70 glucansucrases or branching sucrases, the length, the amino acid composition and the position of the various loops defining the active site are **unique** to **each enzyme specificity**. That is why it is important to conduct structure/function relationship studies in this family and extend the number of resolved structures. **Trying to better understand the dynamics of these enzymes is also probably essential to better guide rational or semi-rational glucansucrase engineering and optimized the direct production of tailor made polymers from sucrose to *in fine*, open new fields of applications.**

References

A

Amari, M., Valérie, G., Robert, H., Morel, S., Moulis, C., Gabriel, B., Remaud-Siméon, M., Fontagné-Faucher, C., 2015. Overview of the glucansucrase equipment of *Leuconostoc citreum* LBAE-E16 and LBAE-C11, two strains isolated from sourdough. FEMS Microbiol. Lett. 362, 1–8. https://doi.org/10.1093/femsle/fnu024

André, I., Grelier, S., Guieysse, D., Lafraya, A., Monsan, P., Moulis, C., Peruch, F., Remaud-Siméon, M., Vuillemin, M., 2018. Enzymatic Production of Glycosylated Synthons. US20180258456A1.

Aoki, H., Shiroza, T., Hayakawa, M., Sato, S., Kuramitsu, H.K., 1986. Cloning of a *Streptococcus mutans* glucosyltransferase gene coding for insoluble glucan synthesis. Infect. Immun. 53, 587–594.

Arabnejad, H., Dal Lago, M., Jekel, P.A., Floor, R.J., Thunnissen, A.-M.W.H., Terwisscha van Scheltinga, A.C., Wijma, H.J., Janssen, D.B., 2017. A robust cosolvent-compatible halohydrin dehalogenase by computational library design. Protein Eng. Des. Sel. 30, 175–189. https://doi.org/10.1093/protein/gzw068

Arenas, M., Weber, C.C., Liberles, D.A., Bastolla, U., 2017. ProtASR: an evolutionary framework for ancestral protein reconstruction with selection on folding stability. Syst. Biol. 66, 1054–1064. https://doi.org/10.1093/sysbio/syw121

Argimón, S., Alekseyenko, A.V., DeSalle, R., Caufield, P.W., 2013. Phylogenetic analysis of glucosyltransferases and implications for the coevolution of mutans *Streptococci* with their mammalian hosts. PLoS ONE 8, e56305. https://doi.org/10.1371/journal.pone.0056305

Argüello Morales, M.A., Remaud-Siméon, M., Willemot, R.-M., Vignon, M.R., Monsan, P., 2001. Novel oligosaccharides synthesized from sucrose donor and cellobiose acceptor by alternansucrase. Carbohydr. Res. 331, 403–411. https://doi.org/10.1016/S0008-6215(01)00038-6

Argüello-Morales, M.A., Remaud Siméon, M., Pizzut-Serin, S., Sarcabal, P., Willemot, R.-M., Monsan, P., 2000a. *Leuconostoc mesenteroides* NRRL B-1355 dsrC gene for dextransucrase. GenBank Access. No AJ250172.

Argüello-Morales, M.A., Remaud-Siméon, M., Pizzut-Serin, S., Sarçabal, P., Willemot, R.-M., Monsan, P., 2000b. Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. FEMS Microbiol. Lett. 182, 81–85. https://doi.org/10.1111/j.1574-6968.2000.tb08878.x

Arnold, F.H., 1996. Directed evolution: Creating biocatalysts for the future. Chem. Eng. Sci. 51, 5091–5102. https://doi.org/10.1016/S0009-2509(96)00288-6

Asensio, J.L., Ardá, A., Cañada, F.J., Jiménez-Barbero, J., 2013. Carbohydrate–aromatic interactions. Acc. Chem. Res. 46, 946–954. https://doi.org/10.1021/ar300024d

Ashkenazy, H., Penn, O., Doron-Faigenboim, A., Cohen, O., Cannarozzi, G., Zomer, O., Pupko, T., 2012. FastML: a web server for probabilistic reconstruction of ancestral sequences. Nucleic Acids Res. 40, W580-584. https://doi.org/10.1093/nar/gks498

B

Badel, S., Bernardi, T., Michaud, P., 2011. New perspectives for *Lactobacilli* exopolysaccharides. Biotechnol. Adv. 29, 54–66. https://doi.org/10.1016/j.biotechadv.2010.08.011

Bai, Y., Gangoiti, J., Dijkstra, B.W., Dijkhuizen, L., Pijning, T., 2017. Crystal structure of 4,6-α-glucanotransferase supports diet-driven evolution of GH70 enzymes from α-amylases in oral bacteria. Struct. Lond. Engl. 1993 25, 231–242. https://doi.org/10.1016/j.str.2016.11.023

Bershtein, S., Goldin, K., Tawfik, D.S., 2008. Intense neutral drifts yield robust and evolvable consensus proteins. J. Mol. Biol. 379, 1029–1044. https://doi.org/10.1016/j.jmb.2008.04.024

Bertrand, A., Morel, S., Lefoulon, F., Rolland, Y., Monsan, P., Remaud-Siméon, M., 2006. *Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. Carbohydr. Res. 341, 855–863. https://doi.org/10.1016/j.carres.2006.02.008

Biswas, S., Biswas, I., 2012. Complete genome sequence of *Streptococcus mutans* GS-5, a serotype c strain. J. Bacteriol. 194, 4787–4788. https://doi.org/10.1128/JB.01106-12

Bloom, J.D., Romero, P.A., Lu, Z., Arnold, F.H., 2007. Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. Biol. Direct 2, 17. https://doi.org/10.1186/1745-6150-2-17

Bounaix, M.-S., Gabriel, V., Robert, H., Morel, S., Remaud-Siméon, M., Gabriel, B., Fontagné-Faucher, C., 2010. Characterization of glucan-producing *Leuconostoc* strains isolated from sourdough. Int. J. Food Microbiol., The 16th CBL (Club des Bactéries Lactiques) Symposium, May 2009, Toulouse, France 144, 1–9. https://doi.org/10.1016/j.ijfoodmicro.2010.05.026

Bozonnet, S., Dols-Laffargue, M., Fabre, E., Pizzut-Serin, S., Remaud-Siméon, M., Monsan, P., Willemot, R.-M., 2002. Molecular characterization of DSR-E, an alpha-1,2 linkage-synthesizing dextransucrase with two catalytic domains. J. Bacteriol. 184, 5753–5761.

Brison, Y., Fabre, E., Moulis, C., Portais, J.-C., Monsan, P., Remaud-Siméon, M., 2010. Synthesis of dextrans with controlled amounts of α-1,2 linkages using the transglucosidase GBD–CD2. Appl. Microbiol. Biotechnol. 86, 545–554. https://doi.org/10.1007/s00253-009-2241-z

Brison, Y., Malbert, Y., Czaplicki, G., Mourey, L., Remaud-Siméon, M., Tranier, S., 2016. Structural Insights into the Carbohydrate Binding Ability of an α-(1→2) Branching Sucrase from Glycoside Hydrolase Family 70. J. Biol. Chem. 291, 7527–7540. https://doi.org/10.1074/jbc.M115.688796

Brison, Y., Pijning, T., Malbert, Y., Fabre, É., Mourey, L., Morel, S., Potocki-Véronèse, G., Monsan, P., Tranier, S., Remaud-Siméon, M., Dijkstra, B.W., 2012. Functional and structural characterization of α-(1->2) branching sucrase derived from DSR-E glucansucrase. J. Biol. Chem. 287, 7915–7924. https://doi.org/10.1074/jbc.M111.305078

Buchan, D.W.A., Minneci, F., Nugent, T.C.O., Bryson, K., Jones, D.T., 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. 41, W349-357. https://doi.org/10.1093/nar/gkt381

C

Carlson, T.L., Woo, A., 2013. Use of low-glycemic sweeteners in food and beverage compositions. US8512739B2.

Carlson, T.L., Woo, A., Zheng, G.-H., 2006. Methods of making syrups. WO2006088884A1.

Champion, E., André, I., Mulard, L.A., Monsan, P., Remaud-Siméon, M., Morel, S., 2009. Synthesis of L-rhamnose and N-acetyl-D-glucosamine derivatives entering in the composition of bacterial polysaccharides by use of glucansucrases. J. Carbohydr. Chem. 28, 142–160. https://doi.org/10.1080/07328300902755796

Cheetham, N.W.H., Slodki, M.E., Walker, G.J., 1991. Structure of the linear, low molecular weight dextran synthesized by a D-glucosyltransferase (GTF-S3) of *Streptococcus sobrinus*. Carbohydr. Polym. 16, 341–353. https://doi.org/10.1016/0144-8617(91)90053-F

Chen, K., Arnold, F.H., 1993. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proc. Natl. Acad. Sci. U. S. A. 90, 5618–5622.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21. https://doi.org/10.1107/S0907444909042073

Claverie, M., Cioci, G., Guionnet, M., Schörghuber, J., Lichtenecker, R., Moulis, C., Remaud-Siméon, M., Lippens, G., 2019a. Futile Encounter Engineering of the DSR-M Dextransucrase Modifies the Resulting Polymer Length. Biochemistry (Mosc.) 58, 2853–2859. https://doi.org/10.1021/acs.biochem.9b00373

Claverie, M., Cioci, G., Vuillemin, M., Bondy, P., Remaud-Siméon, M., Moulis, C., 2019b. Identification of key molecular determinants in the domain V of *Oenococcus kitaharae* DSM 17330 dextransucrase involved in high molar mass dextran synthesis. submitted.

Claverie, M., Cioci, G., Vuillemin, M., Monties, N., Roblin, P., Lippens, G., Remaud-Siméon, M., Moulis, C., 2017. Investigations on the Determinants Responsible for Low Molar Mass Dextran

Formation by DSR-M Dextransucrase. ACS Catal. 7, 7106–7119. https://doi.org/10.1021/acscatal.7b02182

Côté, G.L., 2009. Acceptor products of alternansucrase with gentiobiose. Production of novel oligosaccharides for food and feed and elimination of bitterness. Carbohydr. Res. 344, 187–190. https://doi.org/10.1016/j.carres.2008.10.017

Côté, G.L., 2002. I. Polysaccharides from Prokaryotes, in: Vandamme, E.J., De Baets, S., Steinbüchel, A. (Eds.), Biopolymers: Polysaccharides. pp. 232–350.

Côté, G.L., 1992. Low-viscosity α-d-glucan fractions derived from sucrose which are resistant to enzymatic digestion. Carbohydr. Polym. 19, 249–252. https://doi.org/10.1016/0144-8617(92)90077-4

Côté, G.L., Dunlap, C.A., 2003. Alternansucrase acceptor reactions with methyl hexopyranosides. Carbohydr. Res. 338, 1961–1967. https://doi.org/10.1016/S0008-6215(03)00324-0

Côté, G.L., Dunlap, C.A., Vermillion, K.E., 2009. Glucosylation of raffinose via alternansucrase acceptor reactions. Carbohydr. Res. 344, 1951–1959. https://doi.org/10.1016/j.carres.2009.06.023

Côté, G.L., Dunlap, C.A., Vermillion, K.E., Skory, C.D., 2017. Production of isomelezitose from sucrose by engineered glucansucrases. Amylase 1, 82–93. https://doi.org/10.1515/amylase-2017-0008

Côté, G.L., Holt, S.M., 2007. Prebiotic oligosaccharides via alternansucrase acceptor reactions. US7182954B1.

Côté, G.L., Holt, S.M., Miller-Fosmore, C., 2003. Prebiotic Oligosaccharides via Alternansucrase Acceptor Reactions, in: Oligosaccharides in Food and Agriculture, ACS Symposium Series. American Chemical Society, Washington, DC, pp. 76–89. https://doi.org/10.1021/bk-2003-0849.ch007

Côté, G.L., Robyt, J.F., 1983. The formation of α-D-(1→3) branch linkages by an exocellular glucansucrase from *Leuconostoc mesenteroides* NRRL B-742. Carbohydr. Res. 119, 141–156. https://doi.org/10.1016/0008-6215(83)84053-1

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

Côté, G.L., Robyt, J.F., 1982b. Acceptor reactions of alternansucrase from *Leuconostoc mesenteroides* NRRL B-1355. Carbohydr. Res. 111, 127–142. https://doi.org/10.1016/0008-6215(82)85013-1

Côté, G.L., Sheng, S., 2006. Penta-, hexa-, and heptasaccharide acceptor products of alternansucrase. Carbohydr. Res. 341, 2066–2072. https://doi.org/10.1016/j.carres.2006.04.044

Côté, G.L., Sheng, S., Dunlap, C.A., 2008. Alternansucrase acceptor products. Biocatal. Biotransformation 26, 161–168. https://doi.org/10.1080/10242420701789023

Côté, G.L., Skory, C.D., 2017. Isomelezitose formation by glucansucrases. Carbohydr. Res. 439, 57–60. https://doi.org/10.1016/j.carres.2017.01.004

Côté, G.L., Skory, C.D., 2014. Effects of mutations at threonine-654 on the insoluble glucan synthesized by *Leuconostoc mesenteroides* NRRL B-1118 glucansucrase. Appl. Microbiol. Biotechnol. 98, 6651–6658. https://doi.org/10.1007/s00253-014-5622-x

Craig, D.B., Dombkowski, A.A., 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics 14, 346. https://doi.org/10.1186/1471-2105-14-346

Crooks, G.E., Hon, G., Chandonia, J.-M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190. https://doi.org/10.1101/gr.849004

D

Damnjanović, J., Takahashi, R., Suzuki, A., Nakano, H., Iwasaki, Y., 2012. Improving thermostability of phosphatidylinositol-synthesizing *Streptomyces* phospholipase D. Protein Eng. Des. Sel. 25, 415–424. https://doi.org/10.1093/protein/gzs038

Daudé, D., Vergès, A., Cambon, E., Emond, S., Tranier, S., André, I., Remaud Siméon, M., 2019. Neutral genetic drift-based engineering of a sucrose-utilizing enzyme toward glycodiversification. ACS Catal.

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

Dennes, T.J., Perticone, A.M., Paullin, J.L., 2015. Cationic poly alpha-1,3-glucan ethers. WO2015095358A1.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.-M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465-469. https://doi.org/10.1093/nar/gkn180

Dertli, E., Colquhoun, I.J., Côté, G.L., Le Gall, G., Narbad, A., 2018. Structural analysis of the α-d-glucan produced by the sourdough isolate *Lactobacillus brevis* E25. Food Chem. 242, 45–52. https://doi.org/10.1016/j.foodchem.2017.09.017

Devulapalle, K.S., Goodman, S.D., Gao, Q., Hemsley, A., Mooser, G., 1997. Knowledge-based model of a glucosyltransferase from the oral bacterial group of mutans *Streptococci*. Protein Sci. Publ. Protein Soc. 6, 2489–2493.

Dey, T.B., Kumar, A., Banerjee, R., Chandna, P., Kuhad, R.C., 2016. Improvement of microbial α-amylase stability: strategic approaches. Process Biochem. https://doi.org/10.1016/j.procbio.2016.06.021

Dimopoulou, M., Vuillemin, M., Campbell-Sills, H., Lucas, P.M., Ballestra, P., Miot-Sertier, C., Favier, M., Coulon, J., Moine, V., Doco, T., Roques, M., Williams, P., Petrel, M., Gontier, E., Moulis, C., Remaud-Siméon, M., Dols-Lafargue, M., 2014. Exopolysaccharide (EPS) Synthesis by *Oenococcus oeni*: From Genes to Phenotypes. PloS One 9, e98898. https://doi.org/10.1371/journal.pone.0098898

Djouzi, Z., Andrieux, C., Pelenc, V., Somarriba, S., Popot, F., Paul, F., Monsan, P., Szylit, O., 1995. Degradation and fermentation of α-gluco-oligosaccharides by bacterial strains from human colon: in vitro and in vivo studies in gnotobiotic rats. J. Appl. Bacteriol. 79, 117–127. https://doi.org/10.1111/j.1365-2672.1995.tb00924.x

Dobruchowska, J.M., Meng, X., Leemhuis, H., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2013. Gluco-oligomers initially formed by the reuteransucrase enzyme of *Lactobacillus reuteri* 121 incubated with sucrose and malto-oligosaccharides. Glycobiology 23, 1084–1096. https://doi.org/10.1093/glycob/cwt048

E

Eijsink, V.G.H., Bjørk, A., Gåseidnes, S., Sirevåg, R., Synstad, B., Burg, B. van den, Vriend, G., 2004. Rational engineering of enzyme stability. J. Biotechnol., Highlights from the ECB11: Building Bridges between Biosciences and Bioengineering 113, 105–120. https://doi.org/10.1016/j.jbiotec.2004.03.026

Emond, S., André, I., Jaziri, K., Potocki-Véronèse, G., Mondon, P., Bouayadi, K., Kharrat, H., Monsan, P., Remaud-Siméon, M., 2008. Combinatorial engineering to enhance thermostability of amylosucrase. Protein Sci. Publ. Protein Soc. 17, 967–976. https://doi.org/10.1110/ps.083492608

F

Fabre, E., 2004. Caractérisation de la dextrane-saccharase DSR-E de *Leuconostoc mesenteroides* NRRL B-1299 et applications à la synthèse de composés prébiotiques (Thèse de doctorat). Institut national des sciences appliquées de Toulouse, Toulouse.

Fabre, E., Bozonnet, S., Arcache, A., Willemot, R.-M., Vignon, M., Monsan, P., Remaud-Siméon, M., 2005. Role of the two catalytic domains of DSR-E dextransucrase and their involvement in the formation of highly α-1,2 branched dextran. J. Bacteriol. 187, 296–303. https://doi.org/10.1128/JB.187.1.296-303.2005

Fabre, E., Joucla, G., Moulis, C., Emond, S., Richard, G., Potocki-Veronese, G., Monsan, P., Remaud-Siméon, M., 2006. Glucansucrases of GH family 70: What are the determinants of their specifities? Biocatal. Biotransformation 24, 137–145. https://doi.org/10.1080/10242420600556713

Ferretti, J.J., Gilpin, M.L., Russell, R.R., 1987. Nucleotide sequence of a glucosyltransferase gene from *Streptococcus sobrinus* MFe28. J. Bacteriol. 169, 4271–4278.

Finkenstadt, V.L., Côté, G.L., Willett, J.L., 2011. Corrosion protection of low-carbon steel using exopolysaccharide coatings from *Leuconostoc mesenteroides*. Biotechnol. Lett. 33, 1093–1100. https://doi.org/10.1007/s10529-011-0539-2

Flemming, H.-C., Wingender, J., 2010. The biofilm matrix. Nat. Rev. Microbiol. 8, 623–633. https://doi.org/10.1038/nrmicro2415

Floor, R.J., Wijma, H.J., Colpa, D.I., Ramos-Silva, A., Jekel, P.A., Szymański, W., Feringa, B.L., Marrink, S.J., Janssen, D.B., 2014. Computational Library Design for Increasing Haloalkane Dehalogenase Stability. ChemBioChem 15, 1660–1672. https://doi.org/10.1002/cbic.201402128

Freitas, F., Alves, V.D., Reis, M.A., 2011. Advances in bacterial exopolysaccharides: from production to biotechnological applications. Trends Biotechnol. 29, 388–398.

Funane, K., Ookura, T., Kobayashi, M., 1998. Glucan binding regions of dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. Biosci. Biotechnol. Biochem. 62, 123–127.

G

Gangoiti, J., Lamothe, L., van Leeuwen, S.S., Vafiadi, C., Dijkhuizen, L., 2017a. Characterization of the *Paenibacillus beijingensis* DSM 24997 GtfD and its glucan polymer products representing a new glycoside hydrolase 70 subfamily of 4,6-α-glucanotransferase enzymes. PLoS ONE 12. https://doi.org/10.1371/journal.pone.0172622

Gangoiti, J., Leeuwen, S.S. van, Gerwig, G.J., Duboux, S., Vafiadi, C., Pijning, T., Dijkhuizen, L., 2017b. 4,3-α-Glucanotransferase, a novel reaction specificity in glycoside hydrolase family 70 and clan GH-H. Sci. Rep. 7, 39761. https://doi.org/10.1038/srep39761

Gangoiti, J., Pijning, T., Dijkhuizen, L., 2015. The *Exiguobacterium sibiricum* 255-15 GtfC Enzyme Represents a Novel Glycoside Hydrolase 70 Subfamily of 4,6-α-Glucanotransferase Enzymes. Appl. Environ. Microbiol. 82, 756–766. https://doi.org/10.1128/AEM.03420-15

Gangoiti, J., van Leeuwen, S.S., Vafiadi, C., Dijkhuizen, L., 2016. The Gram-negative bacterium *Azotobacter chroococcum* NCIMB 8003 employs a new glycoside hydrolase family 70 4,6-α-glucanotransferase enzyme (GtfD) to synthesize a reuteran like polymer from maltodextrins and starch. Biochim. Biophys. Acta 1860, 1224–1236. https://doi.org/10.1016/j.bbagen.2016.02.005

Gibson, G.R., Hutkins, R., Sanders, M.E., Prescott, S.L., Reimer, R.A., Salminen, S.J., Scott, K., Stanton, C., Swanson, K.S., Cani, P.D., Verbeke, K., Reid, G., 2017. Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. Nat. Rev. Gastroenterol. Hepatol. 14, 491–502. https://doi.org/10.1038/nrgastro.2017.75

Giffard, P.M., Allen, D.M., Milward, C.P., Simpson, C.L., Jacques, N.A., 1993. Sequence of the gtfK gene of *Streptococcus salivarius* ATCC 25975 and evolution of the gtf genes of oral *Streptococci*. J. Gen. Microbiol. 139, 1511–1522. https://doi.org/10.1099/00221287-139-7-1511

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

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

Global Dextran Market Insights, Forecast to 2025 [WWW Document], 2018. . Mark. Rep. URL https://www.themarketreports.com/report/global-dextran-market-insights-forecast-to-2025 (accessed 12.6.18).

Goldenzweig, A., Goldsmith, M., Hill, S.E., Gertman, O., Laurino, P., Ashani, Y., Dym, O., Unger, T., Albeck, S., Prilusky, J., Lieberman, R.L., Aharoni, A., Silman, I., Sussman, J.L., Tawfik, D.S., Fleishman, S.J., 2016. Automated structure- and sequence-based design of proteins for high bacterial expression and stability. Mol. Cell 63, 337–346. https://doi.org/10.1016/j.molcel.2016.06.012

Goldstein, I.J., Whelan, W.J., 1962. 32. Structural studies of dextrans. Part I. A dextran containing α-1, 3-glucosidic linkages. J. Chem. Soc. Resumed 170–175.

Gonzalez, D., Hiblot, J., Darbinian, N., Miller, J.C., Gotthard, G., Amini, S., Chabriere, E., Elias, M., 2014. Ancestral mutations as a tool for solubilizing proteins: The case of a hydrophobic phosphate-binding protein. FEBS Open Bio 4, 121–127. https://doi.org/10.1016/j.fob.2013.12.006

Grimaud, F., Faucard, P., Tarquis, L., Pizzut-Serin, S., Roblin, P., Morel, S., Gall, S.L., Falourd, X., Rolland-Sabaté, A., Lourdin, D., Moulis, C., Remaud-Siméon, M., Potocki-Veronese, G., 2018. Enzymatic synthesis of polysaccharide-based copolymers. Green Chem. 20, 4012–4022. https://doi.org/10.1039/C8GC01251B

Grysman, A., Carlson, T., Wolever, T.M.S., 2008. Effects of sucromalt on postprandial responses in human subjects. Eur. J. Clin. Nutr. 62, 1364–1371. https://doi.org/10.1038/sj.ejcn.1602890

Guérin, F., Barbe, S., Pizzut-Serin, S., Potocki-Véronèse, G., Guieysse, D., Guillet, V., Monsan, P., Mourey, L., Remaud-Siméon, M., André, I., Tranier, S., 2012. Structural investigation of the thermostability and product specificity of amylosucrase from the bacterium *Deinococcus geothermalis*. J. Biol. Chem. 287, 6642–6654. https://doi.org/10.1074/jbc.M111.322917

Gupta, R.D., Tawfik, D.S., 2008. Directed enzyme evolution via small and effective neutral drift libraries. Nat. Methods 5, 939–942. https://doi.org/10.1038/nmeth.1262

H

Hakulinen, N., Turunen, O., Jänis, J., Leisola, M., Rouvinen, J., 2003. Three-dimensional structures of thermophilic beta-1,4-xylanases from *Chaetomium thermophilum* and *Nonomuraea flexuosa*. Comparison of twelve xylanases in relation to their thermal stability. Eur. J. Biochem. FEBS 270, 1399–1412.

Hare, M.D., Svensson, S., Walker, G.J., 1978. Characterization of the extracellular, water-insoluble α-D-glucans of oral *Streptococci* by methylation analysis, and by enzymic synthesis and degradation. Carbohydr. Res. 66, 245–264. https://doi.org/10.1016/S0008-6215(00)83256-5

Hasselwander, O., DiCosimo, R., You, Z., Cheng, Q., Rothman, S.C., Suwannakham, S., Baer, Z.C., Roesch, B.M., Ruebling-Jass, K.D., Lai, J.P., Hurteau, R.E., Marquez, M.L., Kopatsis, A.D., Ouwehand, A.C., Forssten, S.D., Mukerji, P., Rae, J.M.C., Dragan, Y.P., Damewood, J.R., Tiihonen, K., Ibarra, A., 2017. Development of dietary soluble fibres by enzymatic synthesis and assessment of their digestibility in in vitro, animal and randomised clinical trial models. Int. J. Food Sci. Nutr. 68, 849–864. https://doi.org/10.1080/09637486.2017.1295027

Hehre, E.J., 1941. Production from sucrose of a serologically reactive polysaccharide by a sterile bacterial extract. Science 93, 237–238. https://doi.org/10.1126/science.93.2410.237

Hehre, E.J., Sugg, J.Y., 1942. Serologically reactive polysaccharides produced through the action of bacterial enzymes : I. Dextran of *Leuconostoc mesenteroides* from sucrose. J. Exp. Med. 75, 339–353.

Heinze, T., Liebert, T., Heublein, B., Hornig, S., 2006. Functional Polymers Based on Dextran, in: Klemm, D. (Ed.), Polysaccharides II, Advances in Polymer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 199–291. https://doi.org/10.1007/12_100

Heinzelman, P., Komor, R., Kanaan, A., Romero, P., Yu, X., Mohler, S., Snow, C., Arnold, F., 2010. Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence

blocks by SCHEMA structure-guided recombination. Protein Eng. Des. Sel. 23, 871–880. https://doi.org/10.1093/protein/gzq063

Heinzelman, P., Snow, C.D., Smith, M.A., Yu, X., Kanaan, A., Boulware, K., Villalobos, A., Govindarajan, S., Minshull, J., Arnold, F.H., 2009. SCHEMA recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability. J. Biol. Chem. jbc.C109.034058. https://doi.org/10.1074/jbc.C109.034058

Hernandez-Hernandez, O., Côté, G.L., Kolida, S., Rastall, R.A., Sanz, M.L., 2011. In vitro fermentation of alternansucrase raffinose-derived oligosaccharides by human gut bacteria. J. Agric. Food Chem. 59, 10901–10906. https://doi.org/10.1021/jf202466s

Holt, S.M., Miller-Fosmore, C.M., Côté, G.L., 2005. Growth of various intestinal bacteria on alternansucrase-derived oligosaccharides. Lett. Appl. Microbiol. 40, 385–390. https://doi.org/10.1111/j.1472-765X.2005.01681.x

Holt, S.M., Teresi, J.M., Côté, G.L., 2008. Influence of alternansucrase-derived oligosaccharides and other carbohydrates on α-galactosidase and α-glucosidase activity in *Bifidobacterium adolescentis*. Lett. Appl. Microbiol. 46, 73–79. https://doi.org/10.1111/j.1472-765X.2007.02266.x

Hoshino, T., Fujiwara, T., Kawabata, S., 2012. Evolution of cariogenic character in *Streptococcus mutans*: horizontal transmission of Glycosyl Hydrolase family 70 genes. Sci. Rep. 2, 518. https://doi.org/10.1038/srep00518

Hotz, P., Guggenheim, B., Schmid, R., 1972. Carbohydrates in pooled dental plaque. Caries Res. 6, 103–121. https://doi.org/10.1159/000259783

Huang, J., Xie, D.-F., Feng, Y., 2017. Engineering thermostable (R)-selective amine transaminase from *Aspergillus terreus* through in silico design employing B-factor and folding free energy calculations. Biochem. Biophys. Res. Commun. 483, 397–402. https://doi.org/10.1016/j.bbrc.2016.12.131

Hudson, K.L., Bartlett, G.J., Diehl, R.C., Agirre, J., Gallagher, T., Kiessling, L.L., Woolfson, D.N., 2015. Carbohydrate–Aromatic Interactions in Proteins. J. Am. Chem. Soc. 137, 15152–15160. https://doi.org/10.1021/jacs.5b08424

I

Irague, R., Tarquis, L., André, I., Moulis, C., Morel, S., Monsan, P., Potocki-Véronèse, G., Remaud-Siméon, M., 2013. Combinatorial Engineering of Dextransucrase Specificity. PLOS ONE 8, e77837. https://doi.org/10.1371/journal.pone.0077837

Isenberg, S.L., Brewer, A.K., Côté, G.L., Striegel, A.M., 2010. Hydrodynamic versus size exclusion chromatography characterization of alternan and comparison to off-line MALS. Biomacromolecules 11, 2505–2511. https://doi.org/10.1021/bm100687b

Ito, K., Ito, S., Shimamura, T., Weyand, S., Kawarasaki, Y., Misaka, T., Abe, K., Kobayashi, T., Cameron, A.D., Iwata, S., 2011. Crystal structure of glucansucrase from the dental caries pathogen *Streptococcus mutans*. J. Mol. Biol. 408, 177–186. https://doi.org/10.1016/j.jmb.2011.02.028

J

Jaenicke, R., Böhm, G., 1998. The stability of proteins in extreme environments. Curr. Opin. Struct. Biol. 8, 738–748.

Janeček, S., 2002. How many conserved sequences regions are there in the a-amylase family? Biologia 57, 29–41.

Janeček, S., 1997. α-amylase family: Molecular biology and evolution. Prog. Biophys. Mol. Biol. 67, 67–97. https://doi.org/10.1016/S0079-6107(97)00015-1

Janeček, Š., Svensson, B., MacGregor, E.A., 2014. α-Amylase: an enzyme specificity found in various families of glycoside hydrolases. Cell. Mol. Life Sci. 71, 1149–1170. https://doi.org/10.1007/s00018-013-1388-z

Janeček, S., Svensson, B., Russell, R.R., 2000. Location of repeat elements in glucansucrases of *Leuconostoc* and *Streptococcus* species. FEMS Microbiol. Lett. 192, 53–57.

Jeanes, A., Haynes, W.C., Wilham, C.A., Rankin, J.C., Melvin, E.H., Austin, M.J., Cluskey, J.E., Fisher, B.E., Tsuchiya, H.M., Rist, C.E., 1954. Characterization and classification of dextrans from ninety-six strains of bacteria. J. Am. Chem. Soc. 76, 5041–5052. https://doi.org/10.1021/ja01649a011

Jespersen, H.M., Ann MacGregor, E., Henrissat, B., Sierks, M.R., Svensson, B., 1993. Starch- and glycogen-debranching and branching enzymes: Prediction of structural features of the catalytic (β/α)8-barrel domain and evolutionary relationship to other amylolytic enzymes. J. Protein Chem. 12, 791–805. https://doi.org/10.1007/BF01024938

Joucla, G., 2003. Caractérisation de l'alternane-saccharase de *Leuconostoc mesenteroides* NRRL B-1355 : Approche rationnelle et aléatoire pour la conception de nouvelles glucane-saccharases (Thèse de doctorat). Institut national des sciences appliquées de Toulouse, Toulouse.

Joucla, G., Pizzut-Serin, S., Monsan, P., Remaud-Siméon, M., 2006. Construction of a fully active truncated alternansucrase partially deleted of its carboxy-terminal domain. FEBS Lett. 580, 763–768. https://doi.org/10.1016/j.febslet.2006.01.001

K

Kabsch, W., 2010. XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132. https://doi.org/10.1107/S0907444909047337

Kang, H.-K., Oh, J.-S., Kim, D., 2009. Molecular characterization and expression analysis of the glucansucrase DSRWC from *Weissella cibaria* synthesizing a α(1→6) glucan. FEMS Microbiol. Lett. 292, 33–41. https://doi.org/10.1111/j.1574-6968.2008.01460.x

Kato, C., Nakano, Y., Lis, M., Kuramitsu, H.K., 1992. Molecular genetic analysis of the catalytic site of *Streptococcus mutans* glucosyltransferases. Biochem. Biophys. Res. Commun. 189, 1184–1188. https://doi.org/10.1016/0006-291X(92)92329-V

Kim, D., Robyt, J.F., 1995. Production, selection, and characteristics of mutants of *Leuconostoc mesenteroides* B-742 constitutive for dextransucrases. Enzyme Microb. Technol. 17, 689–695. https://doi.org/10.1016/0141-0229(94)00021-I

Kim, D., Robyt, J.F., 1994. Production and selection of mutants of *Leuconostoc mesenteroides* constitutive for glucansucrases. Enzyme Microb. Technol. 16, 659–664.

Kim, J.F., Jeong, H., Lee, J.-S., Choi, S.-H., Ha, M., Hur, C.-G., Kim, J.-S., Lee, S., Park, H.-S., Park, Y.-H., Oh, T.K., 2008. Complete genome gequence of *Leuconostoc citreum* KM20. J. Bacteriol. 190, 3093–3094. https://doi.org/10.1128/JB.01862-07

Kimura, M., 1989. The neutral theory of molecular evolution and the world view of the neutralists. Genome 31, 24–31. https://doi.org/10.1139/g89-009

Kitaoka, M., Robyt, J.F., 1998. Use of a Microtiter Plate Screening Method for Obtaining *Leuconostoc mesenteroides* Mutants Constitutive for Glucansucrase. Enzyme Microb. Technol. 22, 527–531. https://doi.org/10.1016/S0141-0229(97)00252-4

Kobayashi, M., Matsuda, K., 1977. Structural characteristics of dextrans synthesized by dextransucrases from *Leuconostoc mesenteroides* NRRL B-1299. Agric. Biol. Chem. 41, 1931–1937. https://doi.org/10.1271/bbb1961.41.1931

Koepsell, H.J., Tsuchiya, H.M., Hellman, N.N., Kazenko, A., Hoffman, C.A., Sharpe, E.S., Jackson, R.W., 1953. Enzymatic Synthesis of Dextran Acceptor Specificity and Chain Initiation. J. Biol. Chem. 200, 793–801.

Komatsu, H., Abe, Y., Eguchi, K., Matsuno, H., Matsuoka, Y., Sadakane, T., Inoue, T., Fukui, K., Kodama, T., 2011. Kinetics of dextran-independent α-(1→3)-glucan synthesis by *Streptococcus sobrinus* glucosyltransferase I. FEBS J. 278, 531–540. https://doi.org/10.1111/j.1742-4658.2010.07973.x

Komatsu, H., Katayama, M., Sawada, M., Hirata, Y., Mori, M., Inoue, T., Fukui, K., Fukada, H., Kodama, T., 2007. Thermodynamics of the Binding of the C-Terminal Repeat Domain of *Streptococcus sobrinus* Glucosyltransferase-I to Dextran. Biochemistry (Mosc.) 46, 8436–8444. https://doi.org/10.1021/bi700282c

Koshland, D.E., 1953. Stereochemistry and the Mechanism of Enzymatic Reactions. Biol. Rev. 28, 416–436. https://doi.org/10.1111/j.1469-185X.1953.tb01386.x

Koudelakova, T., Chaloupkova, R., Brezovsky, J., Prokop, Z., Sebestova, E., Hesseler, M., Khabiri, M., Plevaka, M., Kulik, D., Kuta Smatanova, I., Rezacova, P., Ettrich, R., Bornscheuer, U.T., Damborsky, J., 2013. Engineering Enzyme Stability and Resistance to an Organic Cosolvent by Modification of Residues in the Access Tunnel. Angew. Chem. Int. Ed. 52, 1959–1963. https://doi.org/10.1002/anie.201206708

Kralj, S., Eeuwema, W., Eckhardt, T.H., Dijkhuizen, L., 2006. Role of asparagine 1134 in glucosidic bond and transglycosylation specificity of reuteransucrase from *Lactobacillus reuteri* 121. FEBS J. 273, 3735–3742. https://doi.org/10.1111/j.1742-4658.2006.05376.x

Kralj, S., Grijpstra, P., van Leeuwen, S.S., Leemhuis, H., Dobruchowska, J.M., van der Kaaij, R.M., Malik, A., Oetari, A., Kamerling, J.P., Dijkhuizen, L., 2011. 4,6-α-Glucanotransferase, a Novel Enzyme That Structurally and Functionally Provides an Evolutionary Link between Glycoside Hydrolase Enzyme Families 13 and 70▽. Appl. Environ. Microbiol. 77, 8154–8163. https://doi.org/10.1128/AEM.05735-11

Kralj, S., Leeuwen, S.S. van, Valk, V., Eeuwema, W., Kamerling, J.P., Dijkhuizen, L., 2008. Hybrid reuteransucrase enzymes reveal regions important for glucosidic linkage specificity and the transglucosylation/hydrolysis ratio. FEBS J. 275, 6002–6010. https://doi.org/10.1111/j.1742-4658.2008.06729.x

Kralj, S., Stripling, E., Sanders, P., Geel-Schutten, G.H. van, Dijkhuizen, L., 2005. Highly Hydrolytic Reuteransucrase from Probiotic *Lactobacillus reuteri* Strain ATCC 55730. Appl Env. Microbiol 71, 3942–3950. https://doi.org/10.1128/AEM.71.7.3942-3950.2005

Kralj, S., van Geel-Schutten, G.H., Dondorff, M.M.G., Kirsanovs, S., van der Maarel, M.J.E.C., Dijkhuizen, L., 2004a. Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. Microbiology 150, 3681–3690. https://doi.org/10.1099/mic.0.27321-0

Kralj, S., van Geel-Schutten, G.H., van der Maarel, M.J.E.C., Dijkhuizen, L., 2004b. Biochemical and molecular characterization of *Lactobacillus reuteri* 121 reuteransucrase. Microbiology 150, 2099–2112. https://doi.org/10.1099/mic.0.27105-0

Kralj, S., van Geel-Schutten, I.G.H., Faber, E.J., van der Maarel, M.J.E.C., Dijkhuizen, L., 2005. Rational Transformation of *Lactobacillus reuteri* 121 Reuteransucrase into a Dextransucrase. Biochemistry (Mosc.) 44, 9206–9216. https://doi.org/10.1021/bi050447q

Kumar, S., Tsai, C.J., Nussinov, R., 2000. Factors enhancing protein thermostability. Protein Eng. 13, 179–191.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132. https://doi.org/10.1016/0022-2836(82)90515-0

L

Laguerre, S., Amari, M., Vuillemin, M., Robert, H., Loux, V., Klopp, C., Morel, S., Gabriel, B., Remaud-Siméon, M., Gabriel, V., Moulis, C., Fontagné-Faucher, C., 2012. Genome sequences of three *Leuconostoc citreum* strains, LBAE C10, LBAE C11, and LBAE E16, isolated from wheat sourdoughs. J. Bacteriol. 194, 1610–1611. https://doi.org/10.1128/JB.06789-11

Leathers, T.D., 2005. Dextran. Polysacch. Polyam. Food Ind. Prop. Prod. Pat. 1.

Leathers, T.D., Ahlgren, J.A., Cote, G.L., 1997. Alternansucrase mutants of *Leuconostoc mesenteroides* strain NRRL B-21138. J. Ind. Microbiol. Biotechnol. 18, 278–283. https://doi.org/10.1038/sj.jim.2900380

Leathers, T.D., Hayman, G.T., Cote, G.L., 1995. Rapid screening of *Leuconostoc mesenteroides* mutants for elevated proportions of alternan to dextran. Curr. Microbiol. 31, 19–22. https://doi.org/10.1007/BF00294628

Leathers, T.D., Nunnally, M.S., Côté, G.L., 2009. Modification of alternan by dextranase. Biotechnol. Lett. 31, 289–293. https://doi.org/10.1007/s10529-008-9866-3

Leathers, T.D., Nunnally, M.S., Côté, G.L., 2002. Modification of alternan by novel Penicillium spp. J. Ind. Microbiol. Biotechnol. 29, 177–180. https://doi.org/10.1038/sj.jim.7000272

Leemhuis, H., Pijning, T., Dobruchowska, J.M., Leeuwen, S.S. van, Kralj, S., Dijkstra, B.W., Dijkhuizen, L., 2013. Glucansucrases: Three-dimensional structures, reactions, mechanism, α-glucan analysis and their implications in biotechnology and food applications. J. Biotechnol. 250.

Li, M.-Q., Zhang, H.-B., Li, Y., Hu, X.-Q., Yang, J.-W., 2018. The thermoduric effects of site-directed mutagenesis of proline and lysine on dextransucrase from *Leuconostoc mesenteroides* 0326. Int. J. Biol. Macromol. 107, 1641–1649. https://doi.org/10.1016/j.ijbiomac.2017.10.023

Li, Y., Drummond, D.A., Sawayama, A.M., Snow, C.D., Bloom, J.D., Arnold, F.H., 2007. A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments. Nat. Biotechnol. 25, 1051–1056. https://doi.org/10.1038/nbt1333

Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., Russell, R.B., 2003. Protein disorder prediction: implications for structural proteomics. Struct. Lond. Engl. 1993 11, 1453–1459.

Loesche, W.J., 1993. Dental Caries: A Treatable Infection. Automated Diagnostic Documentation. Inc Gd. Haven.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490-495. https://doi.org/10.1093/nar/gkt1178

López-Munguía, A., Pelenc, V., Remaud-Siméon, M., Biton, J., Michel, J.M., Lang, C., Paul, F., Monsan, P., 1993. Production and purification of alternansucrase, a glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355, for the synthesis of oligoalternans. Enzyme Microb. Technol. 15, 77–85. https://doi.org/10.1016/0141-0229(93)90120-Q

López-Munguía, A., Pelenc, V., Remaud-Siméon, M., Paul, F., Monsan, P., Biton, J., Michel, J.M., Lang, C., 1990. Production and purification of *Leuconostoc mesenteroides* NRRL B-1355 alternansucrase. Ann. N. Y. Acad. Sci. 613, 717–722. https://doi.org/10.1111/j.1749-6632.1990.tb18252.x

Luzio, G.A., Mayer, R.M., 1983. The hydrolysis of sucrose by dextransucrase. Carbohydr. Res. 111, 311–318. https://doi.org/10.1016/0008-6215(83)88315-3

Lynch, K.M., Coffey, A., Arendt, E.K., 2018. Exopolysaccharide producing lactic acid bacteria: Their techno-functional role and potential application in gluten-free bread products. Food Res. Int., GF2016 - 4th International Symposium on Gluten-Free food and beverages 110, 52–61. https://doi.org/10.1016/j.foodres.2017.03.012

M

MacGregor, E.A., Jespersen, H.M., Svensson, B., 1996. A circularly permuted α-amylase-type α/β-barrel structure in glucan-synthesizing glucosyltransferases. FEBS Lett. 378, 263–266. https://doi.org/10.1016/0014-5793(95)01428-4

Malbert, Y., Moulis, C., Brison, Y., Morel, S., André, I., Remaud-Siméon, M., 2018. Engineering a branching sucrase for flavonoid glucoside diversification. Sci. Rep. 8, 15153. https://doi.org/10.1038/s41598-018-33394-y

Meng, X., Dobruchowska, J.M., Gerwig, G.J., Kamerling, J.P., Dijkhuizen, L., 2015a. Synthesis of oligo- and polysaccharides by *Lactobacillus reuteri* 121 reuteransucrase at high concentrations of sucrose. Carbohydr. Res. 414, 85–92. https://doi.org/10.1016/j.carres.2015.07.011

Meng, X., Dobruchowska, J.M., Pijning, T., Gerwig, G.J., Dijkhuizen, L., 2016a. Synthesis of new hyperbranched α-glucans from sucrose by *Lactobacillus reuteri* 180 glucansucrase mutants. J. Agric. Food Chem. https://doi.org/10.1021/acs.jafc.5b05161

Meng, X., Dobruchowska, J.M., Pijning, T., López, C.A., Kamerling, J.P., Dijkhuizen, L., 2014. Residue Leu940 has a crucial role in the linkage and reaction specificity of the glucansucrase GTF180 of the probiotic bacterium *Lactobacillus reuteri* 180. J. Biol. Chem. 289, 32773–32782. https://doi.org/10.1074/jbc.M114.602524

Meng, X., Gangoiti, J., Bai, Y., Pijning, T., Van Leeuwen, S.S., Dijkhuizen, L., 2016b. Structure-function relationships of family GH70 glucansucrase and 4,6-α-glucanotransferase enzymes, and their

evolutionary relationships with family GH13 enzymes. Cell. Mol. Life Sci. CMLS 73, 2681–2706. https://doi.org/10.1007/s00018-016-2245-7

Meng, X., Gangoiti, J., Wang, X., Grijpstra, P., van Leeuwen, S.S., Pijning, T., Dijkhuizen, L., 2018. Biochemical characterization of a GH70 protein from *Lactobacillus kunkeei* DSM 12361 with two catalytic domains involving branching sucrase activity. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-018-9236-6

Meng, X., Pijning, T., Dobruchowska, J.M., Gerwig, G.J., Dijkhuizen, L., 2015b. Characterization of the functional roles of amino acid residues in acceptor binding subsite +1 in the active site of the glucansucrase GTF180 enzyme of *Lactobacillus reuteri* 180. J. Biol. Chem. https://doi.org/10.1074/jbc.M115.687558

Meng, X., Pijning, T., Dobruchowska, J.M., Yin, H., Gerwig, G.J., Dijkhuizen, L., 2016c. Structural determinants of alternating α(1→4) and α(1→6) linkage specificity in reuteransucrase of *Lactobacillus reuteri*. Sci. Rep. 6, 35261. https://doi.org/10.1038/srep35261

Meng, X., Pijning, T., Tietema, M., Dobruchowska, J.M., Yin, H., Gerwig, G.J., Kralj, S., Dijkhuizen, L., 2017. Characterization of the glucansucrase GTF180 W1065 mutant enzymes producing polysaccharides and oligosaccharides with altered linkage composition. Food Chem. 217, 81–90. https://doi.org/10.1016/j.foodchem.2016.08.087

Miller, G.L., 1959. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. Anal. Chem. 31, 426–428. https://doi.org/10.1021/ac60147a030

Misaki, A., Torii, M., Sawai, T., Goldstein, I.J., 1980. Structure of the dextran of *Leuconostoc mesenteroides* B-1355. Carbohydr. Res. 84, 273–285. https://doi.org/10.1016/S0008-6215(00)85557-3

Miyazaki, K., Wintrode, P.L., Grayling, R.A., Rubingh, D.N., Arnold, F.H., 2000. Directed evolution study of temperature adaptation in a psychrophilic enzyme. J. Mol. Biol. 297, 1015–1026. https://doi.org/10.1006/jmbi.2000.3612

Mizutani, N., Yamada, M., Takayama, K., Shoda, M., 1994. Constitutive mutants for dextransucrase from *Leuconostoc mesenteroides* NRRL B-512F. J. Ferment. Bioeng. 77, 248–251. https://doi.org/10.1016/0922-338X(94)90228-3

Molina, M., Moulis, C., Monties, N., Pizzut-Serin, S., Guieysse, D., Morel, S., Cioci, G., Remaud Siméon, M., 2019. Deciphering an undecided enzyme: investigations of the structural determinants involved in the linkage specificity of alternansucrase. ACS Catal. https://doi.org/10.1021/acscatal.8b04510

Monchois, V., Arguello-Morales, M., Russell, R.R.B., 1999a. Isolation of an Active Catalytic Core of *Streptococcus downei* MFe28 GTF-I Glucosyltransferase. J. Bacteriol. 181, 2290–2292.

Monchois, V., Remaud-Siméon, M., Monsan, P., Willemot, R.-M., 1998. Cloning and sequencing of a gene coding for an extracellular dextransucrase (DSRB) from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only a α(1–6) glucan. FEMS Microbiol. Lett. 159, 307–315. https://doi.org/10.1111/j.1574-6968.1998.tb12876.x

Monchois, V., Remaud-Siméon, M., Russell, R.R.B., Monsan, P., Willemot, R.-M., 1997. Characterization of *Leuconostoc mesenteroides* NRRL B-512F dextransucrase (DSRS) and identification of amino-acid residues playing a key role in enzyme activity. Appl. Microbiol. Biotechnol. 48, 465–472. https://doi.org/10.1007/s002530051081

Monchois, V., Vignon, M., Escalier, P.-C., Svensson, B., Russell, R.R.B., 2000a. Involvement of Gln937 of *Streptococcus downei* GTF-I glucansucrase in transition-state stabilization. Eur. J. Biochem. 267, 4127–4136. https://doi.org/10.1046/j.1432-1327.2000.01448.x

Monchois, V., Vignon, M., Russell, R.R.B., 2000b. Mutagenesis of Asp-569 of glucosyltransferase I glucansucrase modulates glucan and oligosaccharide synthesis. Appl. Environ. Microbiol. 66, 1923–1927. https://doi.org/10.1128/AEM.66.5.1923-1927.2000

Monchois, V., Willemot, R.M., Monsan, P., 1999b. Glucansucrases: mechanism of action and structure-function relationships. FEMS Microbiol. Rev. 23, 131–151.

Monsan, P., Bozonnet, S., Albenne, C., Joucla, G., Willemot, R.-M., Remaud-Siméon, M., 2001. Homopolysaccharides from lactic acid bacteria. Int. Dairy J., First International Symposium on

Exopolysaccharides from Lactic A cid Bacteria: from Fundamentals to Applications 11, 675–685. https://doi.org/10.1016/S0958-6946(01)00113-3

Monsan, P., Lopez, A., 1981. On the production of dextran by free and immobilized dextransucrase. Biotechnol. Bioeng. 23, 2027–2037. https://doi.org/10.1002/bit.260230908

Monsan, P., Paul, F., Auriol, D., 1995. New Developments in the Application of Enzymes to Synthesis Reactions Peptides and Oligosaccharides., in: Annals of the New York Academy of Sciences. pp. 357–363.

Monsan, P., Remaud-Siméon, M., André, I., 2010. Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. Curr. Opin. Microbiol. 13, 293–300. https://doi.org/10.1016/j.mib.2010.03.002

Moore, J.C., Arnold, F.H., 1996. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. Nat. Biotechnol. 14, 458–467. https://doi.org/10.1038/nbt0496-458

Mooser, G., Hefta, S.A., Paxton, R.J., Shively, J.E., Lee, T.D., 1991. Isolation and sequence of an active-site peptide containing a catalytic aspartic acid from two *Streptococcus sobrinus* alpha-glucosyltransferases. J. Biol. Chem. 266, 8916–8922.

Mooser, G., Iwaoka, K.R., 1989. Sucrose 6-alpha-D-glucosyltransferase from *Streptococcus sobrinus*: characterization of a glucosyl-enzyme complex. Biochemistry (Mosc.) 28, 443–449.

Morel, S., Andre, I., Brison, Y., Cambon, E., Malbert, Y., Pompon, D., Remaud-Siméon, M., Urban, P., 2017. Novel flavonoids o-a-glucosylated on the B cycle, method for the production thereof and uses. US20170107242A1.

Moscovici, M., 2015. Present and future medical applications of microbial exopolysaccharides. Front. Microbiol. 6, 1012.

Moulis, C., André, I., Remaud-Siméon, M., 2016. GH13 amylosucrases and GH70 branching sucrases, atypical enzymes in their respective families. Cell. Mol. Life Sci. CMLS 73, 2661–2679. https://doi.org/10.1007/s00018-016-2244-8

Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., Remaud-Siméon, M., 2006. Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. J. Biol. Chem. 281, 31254–31267. https://doi.org/10.1074/jbc.M604850200

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367. https://doi.org/10.1107/S0907444911001314

Musa, A., Miao, M., Zhang, T., Jiang, B., 2014. Biotransformation of stevioside by *Leuconostoc citreum* SK24.002 alternansucrase acceptor reaction. Food Chem. 146, 23–29. https://doi.org/10.1016/j.foodchem.2013.09.010

N

Naessens, M., Cerdobbel, A., Soetaert, W., Vandamme, E.J., 2005. *Leuconostoc* dextransucrase and dextran: production, properties and applications. J. Chem. Technol. Biotechnol. 80, 845–860. https://doi.org/10.1002/jctb.1322

Nagi, A.D., Regan, L., 1997. An inverse correlation between loop length and stability in a four-helix-bundle protein. Fold. Des. 2, 67–75. https://doi.org/10.1016/S1359-0278(97)00007-2

Newbrun, E., Hoover, C.I., Walker, G.J., 1983. Inhibition by acarbose, nojirimycin and 1-deoxynojirimycin of glucosyltransferase produced by oral *Streptococci*. Arch. Oral Biol. 28, 531–536. https://doi.org/10.1016/0003-9969(83)90186-3

Nielsen, H., Krogh, A., 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 122–130.

Nivedha, A.K., Thieker, D.F., Makeneni, S., Hu, H., Woods, R.J., 2016. Vina-Carb: Improving Glycosidic Angles during Carbohydrate Docking. J. Chem. Theory Comput. 12, 892–901. https://doi.org/10.1021/acs.jctc.5b00834

O

Olivares-Illana, V., López-Munguía, A., Olvera, C., 2003. Molecular characterization of inulosucrase from *Leuconostoc citreum*: a fructosyltransferase within a glucosyltransferase. J. Bacteriol. 185, 3606–3612. https://doi.org/10.1128/JB.185.12.3606-3612.2003

Osorio, M.I., Zúñiga, M.A., Mendoza, F., Jaña, G.A., Jiménez, V.A., 2019. Modulation of glucan-enzyme interactions by domain V in GTF-SI from Streptococcus mutans. Proteins Struct. Funct. Bioinforma. 87, 74–80. https://doi.org/10.1002/prot.25624

P

Packer, M.S., Liu, D.R., 2015. Methods for the directed evolution of proteins. Nat. Rev. Genet. 16, 379–394. https://doi.org/10.1038/nrg3927

Palframan, R., Gibson, G.R., Rastall, R.A., 2003. Development of a quantitative tool for the comparison of the prebiotic effect of dietary oligosaccharides. Lett. Appl. Microbiol. 37, 281–284. https://doi.org/10.1046/j.1472-765X.2003.01398.x

Passerini, D., Vuillemin, M., Laguerre, S., Amari, M., Loux, V., Gabriel, V., Robert, H., Morel, S., Monsan, P., Gabriel, B., Fontagné-Faucher, C., Remaud-Siméon, M., Moulis, C., 2014. Complete Genome Sequence of *Leuconostoc citreum* Strain NRRL B-742. Genome Announc. 2. https://doi.org/10.1128/genomeA.01179-14

Passerini, D., Vuillemin, M., Ufarté, L., Morel, S., Loux, V., Fontagné-Faucher, C., Monsan, P., Remaud-Siméon, M., Moulis, C., 2015. Inventory of the GH70 enzymes encoded by *Leuconostoc citreum* NRRL B-1299 - identification of three novel α-transglucosylases. FEBS J. 282, 2115–2130. https://doi.org/10.1111/febs.13261

Paullin, J.L., Perticone, A.M., Kasat, R.B., Dennes, T.J., 2014. Preparation of poly alpha-1,3-glucan ethers. US20140179913A1.

Pei, J., Grishin, N.V., 2014. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. Methods Mol. Biol. Clifton NJ 1079, 263–271. https://doi.org/10.1007/978-1-62703-646-7_17

Petersen, T.N., Brunak, S., Heijne, G. von, Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786. https://doi.org/10.1038/nmeth.1701

Pijning, T., Vujičić-Žagar, A., Kralj, S., Dijkhuizen, L., Dijkstra, B.W., 2014. Flexibility of truncated and full-length glucansucrase GTF180 enzymes from *Lactobacillus reuteri* 180. FEBS J. 281, 2159–2171. https://doi.org/10.1111/febs.12769

Pijning, T., Vujičić-Žagar, A., Kralj, S., Dijkhuizen, L., Dijkstra, B.W., 2012. Structure of the α-1,6/α-1,4-specific glucansucrase GTFA from *Lactobacillus reuteri* 121. Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun. 68, 1448–1454. https://doi.org/10.1107/S1744309112044168

Piovesan, D., Minervini, G., Tosatto, S.C.E., 2016. The RING 2.0 web server for high quality residue interaction networks. Nucleic Acids Res. 44, 367–374. https://doi.org/10.1093/nar/gkw315

Prakash, O., Jaiswal, N., 2010. alpha-Amylase: an ideal representative of thermostable enzymes. Appl. Biochem. Biotechnol. 160, 2401–2414. https://doi.org/10.1007/s12010-009-8735-4

Prestegard, J.H., Liu, J., Widmalm, G., 2015. Oligosaccharides and Polysaccharides, in: Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., Schnaar, R.L., Seeberger, P.H. (Eds.), Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).

Pucci, M.J., Jones, K.R., Kuramitsu, H.K., Macrina, F.L., 1987. Molecular cloning and characterization of the glucosyltransferase C gene (gtfC) from *Streptococcus mutans*LM7. Infect. Immun. 55, 2176–2182.

R

Reetz, M.T., Carballeira, J.D., Vogel, A., 2006. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angew. Chem. Int. Ed. 45, 7745–7751. https://doi.org/10.1002/anie.200602795

Reetz, M.T., Soni, P., Fernández, L., Gumulya, Y., Carballeira, J.D., 2010. Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. Chem. Commun. 46, 8657–8658. https://doi.org/10.1039/C0CC02657C

Remaud-Siméon, M., Willemot, R.-M., Sarçabal, P., Potocki de Montalk, G., Monsan, P., 2000. Glucansucrases: molecular engineering and oligosaccharide synthesis. J. Mol. Catal. B Enzym. 10, 117–128. https://doi.org/10.1016/S1381-1177(00)00119-3

Richard, G., Morel, S., Willemot, R.-M., Monsan, P., Remaud-Siméon, M., 2003. Glucosylation of α-butyl- and α-octyl-D-glucopyranosides by dextransucrase and alternansucrase from *Leuconostoc mesenteroides*. Carbohydr. Res. 338, 855–864. https://doi.org/10.1016/S0008-6215(03)00070-3

Robert, H., Gabriel, V., Fontagné-Faucher, C., 2009. Biodiversity of lactic acid bacteria in French wheat sourdough as determined by molecular characterization using species-specific PCR. Int. J. Food Microbiol. 135, 53–59. https://doi.org/10.1016/j.ijfoodmicro.2009.07.006

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, 320–324. https://doi.org/10.1093/nar/gku316

Roca, C., Alves, V.D., Freitas, F., Reis, M.A.M., 2015. Exopolysaccharides enriched in rare sugars: bacterial sources, production, and applications. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.00288

S

Sanz, M.L., Côté, G.L., Gibson, G.R., Rastall, R.A., 2006. Selective fermentation of gentiobiose-derived oligosaccharides by human gut bacteria and influence of molecular weight. FEMS Microbiol. Ecol. 56, 383–388. https://doi.org/10.1111/j.1574-6941.2006.00075.x

Sanz, M.L., Côté, G.L., Gibson, G.R., Rastall, R.A., 2005a. Prebiotic properties of alternansucrase maltose-acceptor oligosaccharides. J. Agric. Food Chem. 53, 5911–5916. https://doi.org/10.1021/jf050344e

Sanz, M.L., Gibson, G.R., Rastall, R.A., 2005b. Influence of disaccharide structure on prebiotic selectivity in vitro. J. Agric. Food Chem. 53, 5192–5199. https://doi.org/10.1021/jf050276w

Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A., Brison, Y., Remaud-Siméon, M., Monsan, P., Gibson, G.R., Rastall, R.A., 2011. *In vitro* fermentation of linear and α-1,2-branched dextrans by the human fecal microbiota. Appl Env. Microbiol 77, 5307–5315. https://doi.org/10.1128/AEM.02568-10

Sarbini, S.R., Kolida, S., Naeye, T., Einerhand, A.W., Gibson, G.R., Rastall, R.A., 2013. The prebiotic effect of α-1,2 branched, low molecular weight dextran in the batch and continuous faecal fermentation system. J. Funct. Foods 5, 1938–1946. https://doi.org/10.1016/j.jff.2013.09.015

Sato, S., Koga, T., Inoue, M., 1984. Isolation and some properties of extracellular D-glucosyltransferases and D-fructosyltransferases from *Streptococcus mutans* serotypes c, e, and f. Carbohydr. Res. 134, 293–304. https://doi.org/10.1016/0008-6215(84)85045-4

Sawai, T., Tohyama, T., Natsume, T., 1978. Hydrolysis of fourteen native dextrans by *Arthrobacter* isomaltodextranase and correlation with dextran structure. Carbohydr. Res. 66, 195–205. https://doi.org/10.1016/S0008-6215(00)83252-8

Scott, T.A., Hellman, N.N., Senti, F.R., 1957. Characterization of dextrans by the optical rotation of their cuprammonium complexes. J. Am. Chem. Soc. 79, 1178–1182. https://doi.org/10.1021/ja01562a039

Semyonov, D., Ramon, O., Shoham, Y., Shimoni, E., 2014. Enzymatically synthesized dextran nanoparticles and their use as carriers for nutraceuticals. Food Funct. 5, 2463–2474. https://doi.org/10.1039/C4FO00103F

Seymour, F.R., Knapp, R.D., Bishop, S.H., 1979a. Correlation of the structure of dextrans to their 1H-N.M.R. Spectra. Carbohydr. Res. 74, 77–92. https://doi.org/10.1016/S0008-6215(00)84766-7

Seymour, F.R., Knapp, R.D., Bishop, S.H., 1976. Determination of the structure of dextran by 13C-nuclear magnetic resonance spectroscopy. Carbohydr. Res. 51, 179–194. https://doi.org/10.1016/S0008-6215(00)83325-X

Seymour, F.R., Knapp, R.D., Chen, E.C.M., Bishop, S.H., Jeanes, A., 1979b. Structural analysis of *Leuconostoc* dextrans containing 3-O-α-D-glucosylated α-D-glucosyl residues in both linear-chain and branch-point positions, or only in branch-point positions, by methylation and by 13C-N.M.R. spectroscopy. Carbohydr. Res. 74, 41–62. https://doi.org/10.1016/S0008-6215(00)84764-3

Seymour, F.R., Slodki, M.E., Plattner, R.D., Jeanes, A., 1977. Six unusual dextrans: methylation structural analysis by combined g.l.c.—m.s. of per-O-acetyl-aldononitriles. Carbohydr. Res. 53, 153–166. https://doi.org/10.1016/S0008-6215(00)88083-0

Shah, D.S.H., Joucla, G., Remaud-Siméon, M., Russell, R.R.B., 2004. Conserved repeat motifs and glucan binding by glucansucrases of oral *Streptococci* and *Leuconostoc mesenteroides*. J. Bacteriol. 186, 8301–8308. https://doi.org/10.1128/JB.186.24.8301-8308.2004

Shimamura, A., Nakano, Y.J., Mukasa, H., Kuramitsu, H.K., 1994. Identification of amino acid residues in *Streptococcus mutans* glucosyltransferases influencing the structure of the glucan product. J. Bacteriol. 176, 4845–4850. https://doi.org/10.1128/jb.176.16.4845-4850.1994

Shiroza, T., Ueda, S., Kuramitsu, H.K., 1987. Sequence analysis of the gtfB gene from *Streptococcus mutans*. J. Bacteriol. 169, 4263–4270. https://doi.org/10.1128/jb.169.9.4263-4270.1987

Sidebotham, R.L., 1974. Dextrans, in: Tipson, R.S., Horton, D. (Eds.), Advances in Carbohydrate Chemistry and Biochemistry. Academic Press, pp. 371–444. https://doi.org/10.1016/S0065-2318(08)60268-1

Singh, J.S., Taylor, K.G., Doyle, R.J., 1993. Essential amino acids involved in glucan-dependent aggregation of *Streptococcus sobrinus*. Carbohydr. Res. 244, 137–147. https://doi.org/10.1016/0008-6215(93)80010-C

Smith, M.R., Zahnley, J., Goodman, N., 1994. Glucosyltransferase mutants of *Leuconostoc mesenteroides* NRRL B-1355. Appl. Environ. Microbiol. 60, 2723–2731.

Smith, M.R., Zahnley, J.C., Wong, R.Y., Lundin, R.E., Ahlgren, J.A., 1998. A mutant strain of *Leuconostoc mesenteroides* B-1355 producing a glucosyltransferase synthesizing α(1→2) glucosidic linkages. J. Ind. Microbiol. Biotechnol. 21, 37–45. https://doi.org/10.1038/sj.jim.2900558

Smith, W.S., Hale, J.R., Neylon, C., 2011. Applying neutral drift to the directed molecular evolution of a β-glucuronidase into a β-galactosidase: Two different evolutionary pathways lead to the same variant. BMC Res. Notes 4, 138. https://doi.org/10.1186/1756-0500-4-138

Soetaert, W., Schwengers, D., Buchholz, K., Vandamme, E.J., 1995. A wide range of carbohydrate modifications by a single micro-organism: *Leuconostoc mesenteroides*, in: Petersen, S.B., Svensson, B., Pedersen, S. (Eds.), Progress in Biotechnology, Carbohydrate Bioengineering. Elsevier, pp. 351–358. https://doi.org/10.1016/S0921-0423(06)80116-4

Steiner, K., Schwab, H., 2012. Recent advances in rational approaches for enzyme engineering. Comput. Struct. Biotechnol. J. 2, e201209010. https://doi.org/10.5936/csbj.201209010

Stemmer, W.P.C., 1994. Rapid evolution of a protein in vitro by DNA shuffling. Nature 370, 389–391. https://doi.org/10.1038/370389a0

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

Suresh Kumar, A., Mody, K., Jha, B., 2007. Bacterial exopolysaccharides – a perception. J. Basic Microbiol. 47, 103–117. https://doi.org/10.1002/jobm.200610203

Suwannarangsee, S., Moulis, C., Potocki-Veronese, G., Monsan, P., Remaud-Siméon, M., Chulalaksananukul, W., 2007. Search for a dextransucrase minimal motif involved in dextran binding. FEBS Lett. 581, 4675–4680. https://doi.org/10.1016/j.febslet.2007.08.062

Swistowska, A.M., Gronert, S., Wittrock, S., Collisi, W., Hecht, H.-J., Hofer, B., 2007. Identification of structural determinants for substrate binding and turnover by glucosyltransferase R supports the permutation hypothesis. FEBS Lett. 581, 4036–4042. https://doi.org/10.1016/j.febslet.2007.07.031

T

Torii, M., Sakakibara, K., 1974. Column chromatographic separation and quantitation of α-linked glucose oligosaccharides. J. Chromatogr. A 96, 255–257. https://doi.org/10.1016/S0021-9673(00)98572-8

Trudeau, D.L., Kaltenbach, M., Tawfik, D.S., 2016. On the potential origins of the high stability of reconstructed ancestral proteins. Mol. Biol. Evol. 33, 2633–2641. https://doi.org/10.1093/molbev/msw138

Tsumori, H., Minami, T., Kuramitsu, H.K., 1997. Identification of essential amino acids in the Streptococcus mutans glucosyltransferases. J. Bacteriol. 179, 3391–3396. https://doi.org/10.1128/jb.179.11.3391-3396.1997

U

Uitdehaag, J.C.M., van der Veen, B.A., Dijkhuizen, L., Dijkstra, B.W., 2002. Catalytic mechanism and product specificity of cyclodextrin glycosyltransferase, a prototypical transglycosylase from the α-amylase family. Enzyme Microb. Technol., Third International Symposium on Industrial Proteins 30, 295–304. https://doi.org/10.1016/S0141-0229(01)00498-7

V

Valette, P., Pelenc, V., Djouzi, Z., Andrieux, C., Paul, F., Monsan, P., Szylit, O., 1993. Bioavailability of new synthesised glucooligosaccharides in the intestinal tract of gnotobiotic rats. J. Sci. Food Agric. 62, 121–127. https://doi.org/10.1002/jsfa.2740620204

van Hijum, S.A.F.T. van, Kralj, S., Ozimek, L.K., Dijkhuizen, L., van Geel-Schutten, I.G.H. van, 2006. Structure-function relationships of glucansucrase and fructansucrase enzymes from Lactic Acid Bacteria. Microbiol Mol Biol Rev 70, 157–176. https://doi.org/10.1128/MMBR.70.1.157-176.2006

van Leeuwen, S.S., Kralj, S., Eeuwema, W., Gerwig, G.J., Dijkhuizen, L., Kamerling, J.P., 2009. Structural characterization of bioengineered α-D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180. Biomacromolecules 10, 580–588. https://doi.org/10.1021/bm801240r

Vanschoonbeek, K., Lansink, M., van Laere, K.M.J., Senden, J.M.G., Verdijk, L.B., van Loon, L.J.C., 2009. Slowly digestible carbohydrate sources can be used to attenuate the postprandial glycemic response to the ingestion of diabetes-specific enteral formulas. Diabetes Educ. 35, 631–640. https://doi.org/10.1177/0145721709335466

Vazquez-Figueroa, E., Yeh, V., Broering, J.M., Chaparro-Riggers, J.F., Bommarius, A.S., 2008. Thermostable variants constructed via the structure-guided consensus method also show increased stability in salts solutions and homogeneous aqueous-organic media. Protein Eng. Des. Sel. 21, 673–680. https://doi.org/10.1093/protein/gzn048

Vettori, M.H., Blanco, K., Cortezi, M., De Lima, C., Contiero, J., 2012. Dextran: effect of process parameters on production, purification and molecular weight and recent applications., in: Diálogos Ciênc. pp. 171–186.

Vriend, G., 1990. WHAT IF: A molecular modeling and drug design program. J. Mol. Graph. 8, 52–56. https://doi.org/10.1016/0263-7855(90)80070-V

Vuillemin, M., Claverie, M., Brison, Y., Séverac, E., Bondy, P., Morel, S., Monsan, P., Moulis, C., Remaud-Siméon, M., 2016. Characterization of the First α-(1→3) Branching Sucrases of the GH70 Family. J. Biol. Chem. 291, 7687–7702. https://doi.org/10.1074/jbc.M115.688044

Vuillemin, M., Grimaud, F., Claverie, M., Rolland-Sabaté, A., Garnier, C., Lucas, P., Monsan, P., Dols-Lafargue, M., Remaud-Siméon, M., Moulis, C., 2018. A dextran with unique rheological properties produced by the dextransucrase from *Oenococcus kitaharae* DSM 17330. Carbohydr. Polym. 179, 10–18. https://doi.org/10.1016/j.carbpol.2017.09.056

Vuillemin, M., Malbert, Y., Laguerre, S., Remaud-Siméon, M., Moulis, C., 2014. Optimizing the production of an α-(1→2) branching sucrase in *Escherichia coli* using statistical design. Appl. Microbiol. Biotechnol. 98, 5173–5184. https://doi.org/10.1007/s00253-014-5627-5

Vujičić-Žagar, A., Pijning, T., Kralj, S., López, C.A., Eeuwema, W., Dijkhuizen, L., Dijkstra, B.W., 2010. Crystal structure of a 117 kDa glucansucrase fragment provides insight into evolution and product specificity of GH70 enzymes. Proc. Natl. Acad. Sci. U. S. A. 107, 21406–21411. https://doi.org/10.1073/pnas.1007531107

W

Wang, C., Zhang, H.-B., Li, M.-Q., Hu, X.-Q., Li, Y., 2017. Functional analysis of truncated and site-directed mutagenesis dextransucrases to produce different type dextrans. Enzyme Microb. Technol. 102, 26–34. https://doi.org/10.1016/j.enzmictec.2017.03.011

Wang, Y., Gänzle, M.G., Schwab, C., 2010. Exopolysaccharide synthesized by *Lactobacillus reuteri* decreases the ability of enterotoxigenic *Escherichia coli* to bind to porcine erythrocytes. Appl Env. Microbiol 76, 4863–4866. https://doi.org/10.1128/AEM.03137-09

Wangpaiboon, K., Padungros, P., Nakapong, S., Charoenwongpaiboon, T., Rejzek, M., Field, R.A., Pichyangkura, R., 2018. An α-1,6-and α-1,3-linked glucan produced by *Leuconostoc citreum* ABK-1 alternansucrase with nanoparticle and film-forming properties. Sci. Rep. 8, 8340. https://doi.org/10.1038/s41598-018-26721-w

Wangpaiboon, K., Pitakchatwong, C., Panpetch, P., Charoenwongpaiboon, T., Field, R.A., Pichyangkura, R., 2019. Modified properties of alternan polymers arising from deletion of SH3-like motifs in Leuconostoc citreum ABK-1 alternansucrase. Carbohydr. Polym. 220, 103–109. https://doi.org/10.1016/j.carbpol.2019.05.002

Wheeler, L.C., Lim, S.A., Marqusee, S., Harms, M.J., 2016. The thermostability and specificity of ancient proteins. Curr. Opin. Struct. Biol. 38, 37–43. https://doi.org/10.1016/j.sbi.2016.05.015

Wijma, H.J., Floor, R.J., Janssen, D.B., 2013. Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. Curr. Opin. Struct. Biol. 23, 588–594. https://doi.org/10.1016/j.sbi.2013.04.008

Wijma, H.J., Floor, R.J., Jekel, P.A., Baker, D., Marrink, S.J., Janssen, D.B., 2014. Computationally designed libraries for rapid enzyme stabilization. Protein Eng. Des. Sel. 27, 49–58. https://doi.org/10.1093/protein/gzt061

Wilham, C.A., Alexander, B.H., Jeanes, A., 1955. Heterogeneity in dextran preparations. Arch. Biochem. Biophys. 59, 61–75. https://doi.org/10.1016/0003-9861(55)90463-X

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242. https://doi.org/10.1107/S0907444910045749

Withers, S., Williams, S., 2007. Glycoside hydrolases. CAZypedia.

Wittrock, S., Swistowska, A.M., Collisi, W., Hofmann, B., Hecht, H.-J., Hofer, B., 2008. Re- or displacement of invariant residues in the C-terminal half of the catalytic domain strongly affects catalysis by glucosyltransferase R. FEBS Lett. 582, 491–496. https://doi.org/10.1016/j.febslet.2007.12.040

Wong, C., Hefta, S.A., Paxton, R.J., Shively, J.E., Mooser, G., 1990. Size and subdomain architecture of the glucan-binding domain of sucrose:3-α-D-glucosyltransferase from *Streptococcus sobrinus*. Infect. Immun. 58, 2165–2170.

Wren, B.W., Russell, R.R., Tabaqchali, S., 1991. Antigenic cross-reactivity and functional inhibition by antibodies to *Clostridium difficile* toxin A, *Streptococcus mutans* glucan-binding protein, and a synthetic peptide. Infect. Immun. 59, 3151–3155.

Wu, B., Wijma, H.J., Song, L., Rozeboom, H.J., Poloni, C., Tian, Y., Arif, M.I., Nuijens, T., Quaedflieg, P.J.L.M., Szymanski, W., Feringa, B.L., Janssen, D.B., 2016. Versatile peptide C-terminal functionalization via a computationally engineered peptide amidase. ACS Catal. 6, 5405–5414. https://doi.org/10.1021/acscatal.6b01062

Y

Yan, M., Wang, B.-H., Xu, X., Chang, P., Hang, F., Wu, Z., You, C., Liu, Z., 2018. Molecular and functional study of a branching sucrase-like glucansucrase reveals an evolutionary intermediate between two subfamilies of the GH70 enzymes. Appl. Environ. Microbiol. 84. https://doi.org/10.1128/AEM.02810-17

Yang, Z.R., Thomson, R., McNeil, P., Esnouf, R.M., 2005. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics 21, 3369–3376. https://doi.org/10.1093/bioinformatics/bti534

You, L., Arnold, F.H., 1996. Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. Protein Eng. Des. Sel. 9, 77–83. https://doi.org/10.1093/protein/9.1.77

Z

Zeymer, C., Hilvert, D., 2018. Directed evolution of protein catalysts. Annu. Rev. Biochem. 87, 131–157. https://doi.org/10.1146/annurev-biochem-062917-012034

Zhang, J., Lin, Y., Sun, Y., Ye, Y., Zheng, S., Han, S., 2012. High-throughput screening of B factor saturation mutated *Rhizomucor miehei* lipase thermostability based on synthetic reaction. Enzyme Microb. Technol. 50, 325–330. https://doi.org/10.1016/j.enzmictec.2012.03.002

Zhang, J.-H., Dawes, G., Stemmer, W.P.C., 1997. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. Proc. Natl. Acad. Sci. 94, 4504–4509.

Zhang, X.-F., Yang, G.-Y., Zhang, Y., Xie, Y., Withers, S.G., Feng, Y., 2016. A general and efficient strategy for generating the stable enzymes. Sci. Rep. 6, 33797. https://doi.org/10.1038/srep33797

Abbreviations

ASR: Alternansucrase

BRS: Branching sucrase

CAZy: Carbohydrate-active enzymes

CW: Cell wall

DNS: dinitrosalicylic acid

DP: Degree of Polymerization

DSR: dextransucrase

GBD: Glucan Binding Domain

GH: Glycoside Hydrolase

GS: Glucansucrase

GT: glucanotransferase

GTF: glucosyltransferase

HPAEC-PAD: High pressure anion exchange chromatography with pulsed amperometric detection

HPLC: High pressure liquid chromatography

HPSEC: High pressure size exclusion chromatography

HMM: High molar mass

I2: Isomaltose (α-D-Glc*p*-(1→6)-D-Glc)

I3: Isomaltotriose (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→6)-D-Glc)

IPTG: Isopropyl β-D-1-thiogalactopyranoside

kDa: kilo dalton

LAB: Lactic acid bacteria

LB: Lysogeny Broth

LMM: Low molar mass

L.: *Lactobacillus*

Ln.: *Leuconostoc*

NMR: Nuclear Magnetic Resonance

OA: Oligoalternan

OD: Oligodextran / Optical density

PEG: Polyethylene glycol

PCR: Polymerase Chain Reaction

PDB: Protein Data Bank

SBS-A1: Sugar Binding Site A1

SDS-PAGE: Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

Sp.: specie

Tm: Melting temperature

U: enzymatic unit

UV: ultraviolet

Artwork

&

Table Contents

Figure content

Chapter I- Litterature review

PhD objectives

Chapter II- Deciphering an undecided enzyme

Chapter III- A specific oligosaccharide binding site in domain A

Chapter IV- Understanding ASR stability and exploration of engineering approaches for ASR stabilization: from random to rational engineering strategies

Conclusion

Table content

Chapter I- Litterature review

Chapter II- Deciphering an Undecided enzyme

Chapter III- A specific oligosaccharide binding site in domain A

Chapter IV- Understanding ASR stability and exploration of engineering approaches for ASR stabilization: from random to rational engineering strategies

Conclusion

Table of Annexes

Annex I: Screening results

The first objectives of the PhD were to get a better insight into the determinants responsible for both glucansucrase specificity and stability to temperature and organic solvent, as well as trying to improve this stability by semi-rational or directed evolution. To do so, the first step was to choose one enzyme that will be used as model of study.

As we had very few data on GH70 enzymes activity in the presence of organic solvent, we thus decided to screen the GH70 enzymes available in the lab (glucansucrases and branching sucrases) by initial activity determination in presence of up to 40% of dimethyl sulfoxide (DMSO), an universal solvent at their optimal temperature of activity (30°C). The results are presented for native enzymes on Figure 1.

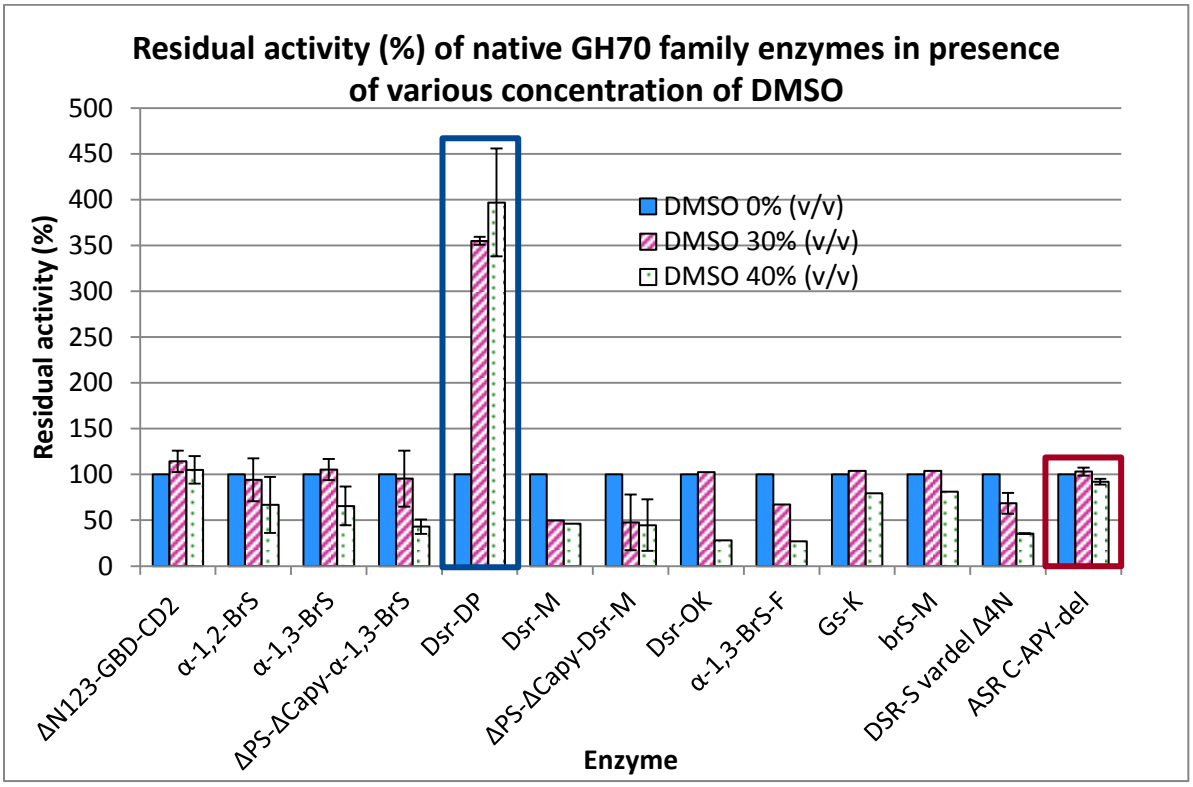


Figure 1: Results from the screening at 30°C of crude GH70 enzyme extract in presence of several concentrations of DMSO.
Reference initial activity was determined at 30°C with 292 mM sucrose, 50 mM sodium acetate buffer and 20 µL of crude enzyme extract and was set to 100% activity. Initial activities were measured in presence of 30% and 40% DMSO in the same conditions and the residual activity was calculated based on the reference one obtained at 30°C.

The dextransucrase DSR-DP from *Ln. citreum* NRRL B-1299 (Passerini et al., 2015) was obviously the most active glucansucrase in presence of DMSO. e showed that the specific activity of this enzyme was not only maintained with DMSO but enhanced with a 10-fold improvement in presence of 40% (w/v) of solvent. However, despite this major improvement of the initial enzyme activity in presence of DMSO, the enzyme was shown to be not stable during reaction time, as DSR-DP was not able to deplete the totality of 292 mM sucrose in presence of 40% DMSO. In contrast, the second best candidate, the alternansucrase (ASR) from *Ln. citreum* NRRL B-1355, maintained its activity even in presence of 40% of DMSO until total sucrose consumption (Figure 2). Choosing ASR as the model enzyme for our study appeared a very good choice considering its multiple attributes mentioned all along this literature review. From the screen, ΔN_{123}-GBD-CD2 also maintained its initial activity in presence of DMSO (Figure 1), but the enzyme was known to be poorly stable and its structure has already been solved.

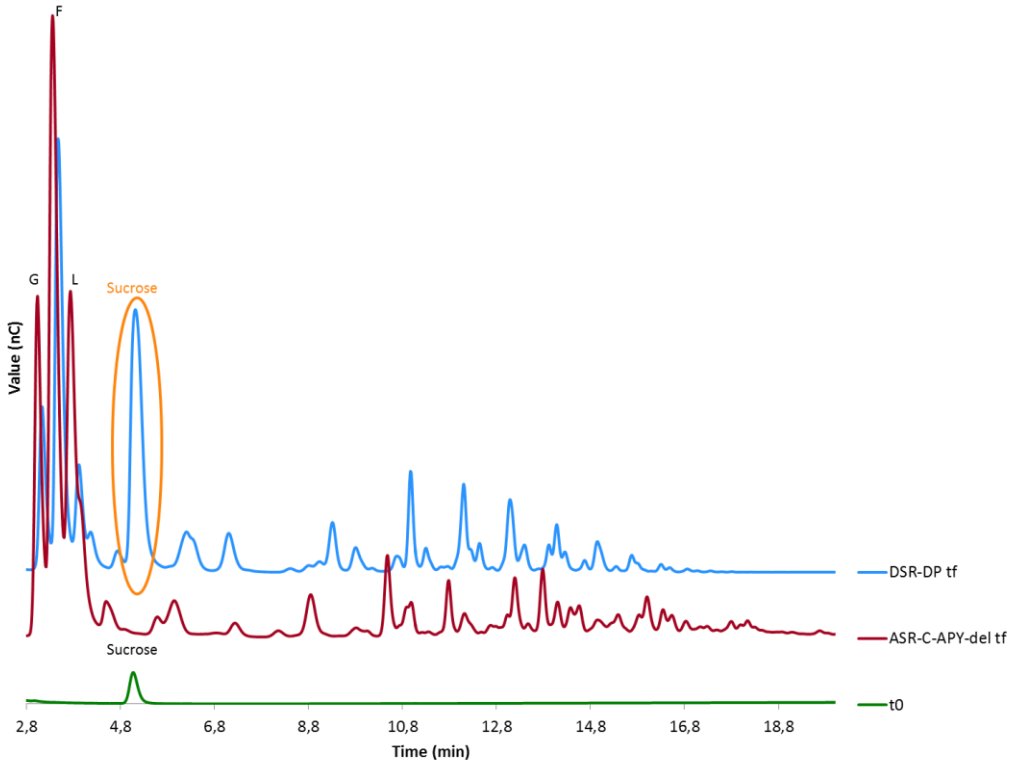


Figure 2: HPAEC-PAD analysis of the products formed by ASR-C-APY-del and DSR-DP in presence of 100 g/L sucrose, 50 mM NaAc pH 5.75 and 40% DMSO (w/v). Initial time (t0) was diluted 10,000 times; final times (tf) were diluted 200 times.

Annex II: Summary of all GH70 structures in the Protein Data Bank

PDB ID	Enzyme	Organism	Mutations	Ligands	Resolution	Reference
3AIB	GTF-SI catalytic core	*Streptococcus mutans*	N597D R600K T727I A734V	MAL MES CA	3,09	Ito et al., 2011
3AIC				ACR MES CA	3,11	
3AIE				MES CA	2,1	
3HZ3	GTF180-ΔN	*Lactobacillus reuteri* 180	D1025N F1674L N-terminally truncated 742-1172	Sucrose CA	2,22	Vujicic et al., 2010
3KLK			F1674L N-terminally truncated 742-1172	GOL CA	1,65	
3KLL				MAL GOL CA	2	
4AYG			N-terminally truncated 742-1172	SO4 GOL IPA ACY CA	2	Pijning et al., 2014
4AMC	GTFA-ΔN	*Lactobacillus reuteri* 121	N-terminally truncated 741-1781	CA	3,6	Pijning et al., 2012

3TTO	ΔN_{123}-GBD-CD2	*Leuconostoc mesenteroides* NRRL B-1299	N-terminally truncated 1749-2835	GOL CA	3,3	Brison et al., 2012
3TTQ				PG4 GOL CA Na+	1,9	
4TTU				Isomaltotriose GLC PEG CA NA	2,18	Brison et al., 2016
4TVC				Gluco-oligosaccharides GLC GOL CA NA	1,85	
4TVD				β-D-GLC GLC PEG CA	2,3	
5JBD	GtfB-ΔN-ΔV	*Lactobacillus reuteri* 121	L761M	PGE SO4 GOL ACT CA	1,8	Bai et al., 2017
5JBE				MAL BGC	2,1	

				GLC SO4 ACT CA		
5JBF			L761M, D1015N	maltopentaose amylotriose MAL SO4 CA	2,19	
5NGY	DSR-MΔ2	*Leuconostoc citreum* NRRL B-1299	E715Q	i4 GLC PR CA	3,7	Claverie et al., 2017
5LCF				GOL CA	3,2	
5O8L				Sucrose CA	3,6	

Annex III: Tables for structure-function relationship studies

Table 1: Mutagenesis studies on residues identified important in the sucrose complex. References: 1= (Swistowska et al., 2007); 2= (Meng et al., 2015b); 3= (Brison et al., 2012); 4= (Wang et al., 2017); 5=(Tsumori et al., 1997); 6= (Meng et al., 2017); 7= (Monchois et al., 1997); 8= (van Leeuwen et al., 2009); 9= (Wittrock et al., 2008); 10= (Monchois et al., 2000a)

Ref.	Enzyme			Localization				ASR eq.	Effects												
	Enzyme name and GenBank accession	Organism	Catal. Residues	Dom.	Motif	Precise localization	Mutants		Active?	Residual activity (%)	HMM proportion (%)	Hydrolysis (%)	Methylation								
													terminal	Branching	α-1,3 linked	α-1,4 linked	α-1,6 linked	α-1,3	α-1,4 or IG	α-1,6 or SG
1	GtfR	*S. oralis* ATCC10557	D516 E554 D627 Model: D190 E228 D301	A	II	WILD TYPE		R633	YES	100										
						2nd residue upstream the nucleophile	R514M		NO											
2	GTF180	*L. reuteri 180*	D1025 E1063 D1136			4th residue downstream the nucleophile	N1029Y	N639	YES	48		59	12	12	21		55	32		68
							N1029G		YES	17		74	n.d.	n.d.	n.d.		n.d.	58		42
							N1029T		YES	9		72	18	19	53		10	75		25
							N1029M		YES	24		76	17	20	39		24	62		38
							N1029R		YES	26		100	n.d.	n.d.	n.d.		n.d.	n.d.		n.d.
3	ΔN_{123}-GBD-CD2	*Ln. citreum* NRRL B-1299	D2210 E2248 D2322			WILD TYPE			YES	100		87 (sucrose) 20 (dex 1kDa)								
						4th residue downstream the nucleophile	F2214N		YES	19		100 (sucrose) 100 (dex 1kDa)								
4	DexYG	*Ln. mesenteroides 0326*	D551 E589 D662			WILD TYPE			YES	100 72,54 U/mg		16						5		95

						4th residue downstream nucleophile	N555Y		YES	115		14		7,8		92,2
					III	2nd residue downstream the acid/base catalyst	W591G		YES	117		14		2,6		97,4
5	GTF-I	*S. mutans* GS-5	D451 E489 D562			WILD TYPE			YES	100 Glucan synthesis act=4,823 cpm						
						2nd residue downstream the acid/base	W491G		YES	0,06						
6	GTF180	*L. reuteri* 180	D1025 E1063 D1136			WILD TYPE			YES	100 (47,7 U/mg)	17	24		33		67
						2nd residue downstream acid/base catalyst E	W1065F	W675	YES	46	2	31		45	6	49
							W1065K		YES	4	0	84		n.d.		n.d.
							W1065R		YES	4	0	65		n.d.		n.d.
							W1065L		YES	13	0	55		n.d.		n.d.
							W1065N		YES	12	0	64		n.d.		n.d.
							W1065Q		YES	7	0	67		n.d.		n.d.
							W1065M		YES	13	0	54		n.d.		n.d.
							W1065P		YES	6	0	69		n.d.		n.d.
							W1065E		YES	4	0	67		n.d.		n.d.
							W1065G		NO							
3	ΔN_{123}-GBD-CD2	*Ln. citreum NRRL B-1299*	D2210 E2248 D2322			WILD TYPE			YES	100		87 (sucrose) 20 (dex 1kDa)				
						2nd residue downstream the acid/base catalyst	G2249W		YES	1		89 (sucrose) 45 (dex 1kDa)				

						1st and 2nd residues downstream the acid/base catalyst	A2249W G2249W	D674 W675	YES	89		90 (sucrose) 36 (dex 1kDa)												
5	GTF-I		D451 E489 D562		IV	WILD TYPE			YES	100 Glucan synthesis act (cpm)= 4,823														
						1st residue upstream the TSS	H561G		YES	0,2														
7	DSR-S	*Ln. citreum NRRL B-512F*	D551 E589 D662			WILD TYPE		H766	YES	100 (1040 $U.g^{-1}$)														
						1st residue downstream the TSS	H661R		YES	0,5														
1	GtfR BAA95201.1	*S. oralis ATCC10557*	D516 E554 D627 Model: D190 E228 D301			WILD TYPE			YES	100														
						1st residue upstream the TSS	H626F		YES	0,1														
4	DexYG	*Ln. mesenteroides 0326*	D551 E589 D662			WILD TYPE			YES	100 72,54 U/mg		16								5		95
						4th residue downstream the TSS	Q666R		YES	114		16								4,2	4,3	91,5
8	GTF180	*L. reuteri 180*	D1025 E1063 D1136			WILD TYPE		Q771			32,300,000		12	12	24	0	52	31	0	69
						4th residue downstream the TSS	Q1140E				16,200,000		12	18	16	2	52	29	3	68
							Q1140A				23,900,000		11	14	6	0	69	16	0	84
							Q1140H						8	8	8	0	76	14	0	86
9	GTFR-100	*S. oralis*	D516 E554 D627			WILD TYPE		D1169	YES	100 10,3 U/mg		31												

						First aspartate of the motif	D1004A		YES	6		92						
10	GTF-I	*S. downei*	D453 E491 D564		I	WILD TYPE		Q1174	YES	100 15.6 $U.mg^{-1}$	60	16						
						7th residue in motif I	Q937H		YES	1	60	27						
							Q937A		YES	3	61	18						
							Q937E		YES	3	60	18						
							Q937N		YES	3	62	17						
							Q937D		YES	1	57	17						
4	DexYG	*Ln. mesenteroides 0326*	D551 E589 D662			WILD TYPE			YES	100 72,54 U/mg		16		5		95		
						Conserved glutamine	Q1029K		YES	34		39		5,4		94,6		
9	GTFR-100	*S. oralis*	D516 E554 D627		none	WILD TYPE		Y1124	YES	100 10,3 U/mg		31						
						none	Y965S		YES	0,9		92						

Table 2: Mutagenesis studies on GBD-CD2. References: 1=(Brison et al., 2012); 2= (Fabre et al., 2006)

Ref.	Enzyme			Localization				ASR equivalent	Effects		
	Enzyme name and GenBank accession nb	Organism	Catal. Residues	Domain	Motif	Precise localization	Mutant	Residue	Active?	% of residual activity or % of residual catalytic efficiency	Hydrolysis (%)
1	ΔN_{123}-GBD-CD2 CDX66820.1	*Ln. citreum* NRRL B-1299	D2210 E2248 D2322	A	II	WILD TYPE		N639	YES	100	87 (sucrose) 20 (dex 1kDa)
						4th residue downstream the nucleophile	F2214N		YES	19	100 (sucrose) 100 (dex 1kDa)
2	GBD-CD2 CDX66820.1					WILD TYPE		639-NVDADLLKI-647	YES		
						8 residues downstream the nucleophile	2214-FIHDTIQR-2222 => NVDADLLQR		NO		
1	ΔN_{123}-GBD-CD2 CDX66820.1				III	WILD TYPE		D674	YES	100	87 (sucrose) 20 (dex 1kDa)
						1st residue downstream the acid/base catalyst	A2249W		YES	22	89 (sucrose) 24 (dex 1kDa)
1						WILD TYPE		D675	YES	100	87 (sucrose) 20 (dex 1kDa)
						2nd residue downstream the acid/base catalyst	G2249W		YES	1	89 (sucrose) 45 (dex 1kDa)

						1st and 2nd residues downstream the acid/base catalyst	A2249W G2249W	D674 W675	YES	89	90 (sucrose) 36 (dex 1kDa)
2	GBD-CD2 CDX66820.1				IV	WILD TYPE		768-YDAQDPI-774	YES	100	
						Residues downstream the TSS	2323-KGVQEKV-2329 =>SEVQTVI		NO		
					II IV	Residues downstream the nucleophile Residues downstream the TSS	2214-FIHDTIQR-2222 =>NVDADLLQR 2323-KGVQEKV-2329 =>SEVQTVI	639-NVDADLLKI-647 768-YDAQDPI-774	YES	n.d.	Low level of hydrolysis No glucooligosaccharides from maltose but production of a product stainable using the Schiff coloration

Table 3: Mutagenesis studies on DSR-M. References: 1=(Claverie et al., 2017)

Ref.				Localization			ASR equivalent	Effects			
	Enzyme name and GenBank accession nb	Organism	Catal. Residues	Domain	Precise localization	Mutants	Residue	Active?	Residual activity (%)	Relative product size (%)
1	DSR-MΔ2 CDX66895.1	*Ln. citreum* NRRL B-1299	D677 E715 D790	A	WILD TYPE		No equivalent	YES	100 (67 U/mg)	100 (28,000 g/mol)
					loop A1	D813A		YES	50	79
						L816A	No equivalent	YES	50	85
				B	loop B1	L575W	L544	YES	n.d.	45
				V	pocket A	Y180A	Y158	YES	n.d.	44
					pocket B	Y264A	Y241	YES	n.d.	63
					pocket A pocket B	Y180A Y264A	Y158 Y241	YES	n.d.	40

Table 4: Mutagenesis studies on GTF180 (AAU08001.1) from *Lactobacillus reuteri* 180. Catalytic residues: D1025, E1063, D1136. References: 1=(Meng et al., 2015b); 2= (Meng et al., 2017); 3= (van Leeuwen et al., 2009); 4=(Meng et al., 2016a); 5= (Meng et al., 2014)

Ref.	Enzyme	Localization				ASR eq	Effets												
	Enzyme name and GenBank accession nb	Domain	Motif	Precise localization	Mutants	Residue	Active?	Residual activity (%)	Relative molecular mass (%)	HMM glucan proportion (%)	Hydrolysis (%)	Methylation					α-1,3	α-1,4	α-1,6
												terminal	Branching	α-1,3 linked	α-1,4 linked	α-1,6 linked			
1	GTF180-ΔN	A	II	WILD TYPE		D638	YES	100 (47,7 U/mg)			22	11	13	21		55	33		67
				3rd residue downstream the nucleophile	D1028Y		YES	8			10	8	7	12		73	16		84
					D1028W		YES	13			12	7	6	12		75	16		84
					D1028L		YES	30			16	12	12	11		65	22		78
					D1028K		YES	7			17	13	12	5		70	18		82
					D1028G		YES	31			12	13	12	4		71	15		85
					D1028N		YES	37			17	11	12	10		67	20		80
				4th residue downstream the nucleophile	N1029Y	N639	YES	48			59	12	12	21		55	32		68
					N1029G		YES	17			74	n.d.	n.d.	n.d.		n.d.	58		42
					N1029T		YES	9			72	18	19	53		10	75		25
					N1029M		YES	24			76	17	20	39		24	62		38
					N1029R		YES	26			100	n.d.	n.d.	n.d.		n.d.	n.d.		n.d.
2			III	WILD TYPE		W675		100 (47,7 U/mg)		17	24						33		67
				2nd residue downstream acid/base	W1065F		YES	46		2	31						45	6	49
					W1065K		YES	4		0	84								
					W1065		YES	4		0	65								

				catalyst E	R															
					W1065L		YES	13		0	55									
					W1065N		YES	12		0	64									
					W1065Q		YES	7		0	67									
					W1065M		YES	13		0	54									
					W1065P		YES	6		0	69									
					W1065E		YES	4		0	67									
					W1065G		NO													
3	GTF180AAU08001.1		IV	WILD TYPE		Y768			100 (32,300,000 g/mol)			12	12	24	0	52	31	0	69
				1st residue downstream the TSS	S1137N				75			12	12	26	3	47	35	4	61
					S1137Y				39			18	18	21	4	39	36	6	58
				2nd residue downstream the TSS	N1138D	D769						10	10	24	0	56	35	0	65
				1st and 2nd residue downstream the TSS	S1137YN1138D	Y768D769			33				20	23	7	31	40	8	52
				3rd residue downstream the TSS	A1139S	A768			94			12	13	24	0	51	34	0	66
					A1139L				60			14	16	23	0	47	38	0	62
				2nd residue downstream the nucleophile and 1st	V1027PS1137NA1139S	V637Y768A770			29			18	18	10	12	42	28	12	60

				and 3rd residues downstream the TSS															
				1st and 3rd residue downstream the TSS	S1137N A1139V	Y768 A770					11	11	30	2	46	40	2	58
				1st and 3rd residue downstream the TSS	S1137N A1139S				62			15	27	4	39	37	4	59
				2nd and 3rd residue downstream the TSS	N1138E A1139V	D769 A770			65		13	15	24	0	48	34	1	65
				4th residue downstream the TSS	Q1140E	Q771			50		12	18	16	2	52	29	3	68
					Q1140A				74		11	14	6	0	69	16	0	84
					Q1140H						8	8	8	0	76	14	0	86
4	GTF180-ΔN AAU08001.1		none	WILD TYPE		Y695	YES	100	100 (31,000,000 g/mol)		22	11	13	21	<1	55	33	0	67
				Helix α6 (a4)	D1085A		YES	47	65		19	16	18	13	1	52	29	2	69
					D1085V		YES	60	74		18	15	16	11	1	57	25	2	73
					D1085L		YES	49	65		18	17	18	13	2	50	29	2	69
					D1085E		YES	32	71		23	15	17	21	1	46	34	2	64
					D1085Q		YES	28	61		21	16	17	16	2	49	33	2	65
					D1085H		YES	28	74		20	16	17	14	2	51	29	2	69
					D1085Y		YES	15	65		22	17	18	12	2	51	30	2	68
					R1088G	T698	YES	54	58		28	14	16	15	1	54	32	2	66
					R1088T		YES	63	71		26	14	15	18	2	51	32	2	66
					R1088N		YES	52	71		28	14	15	19	2	50	34	2	64

					R1088E		YES	51	65		27	15	17	15	2	51	33	3	64
					R1088H		YES	57	74		23	14	16	20	2	48	34	2	64
					R1088W		YES	54	5		21	18	19	11	3	49	30	4	66
					N1089G	S699	YES	77	87		21	11	12	21	<1	56	32	1	67
					N1089S		YES	95	94		20	12	12	20	<1	56	32	1	67
					N1089L		YES	77	94		22	11	12	23	1	53	34	1	65
					N1089D		YES	72	87		21	12	11	18	1	58	34	1	65
					N1089R		YES	77	87		19	13	14	17	<1	56	30	1	69
					N1089P		YES	36	77		21	14	14	17	1	54	33	1	66
					N1089Y		YES	56	84		25	13	14	16	1	56	32	1	67
					D1085-R1088-N1089->DHT (DSR-S)	Y695 T698 S699	YES	70	71		25	13	15	20	2	50	34	2	64
					D1085-R1088-N1089->NRL (GTFR)		YES	48	71		21	16	16	14	1	53	32	1	67
					D1085-R1088-N1089->DKN (GTFA)		YES	42	74		29	13	15	22	1	49	34	1	65
					D1085-R1088-N1089->VKG (GTFO)		YES	57	61		19	17	17	13	1	52	30	2	68
					D1085-R1088-N1089->YTS (ASR)		YES	44	65		20	16	17	13	1	53	33	1	66
					D1085-R1088-N1089->ETL (GBD-CD2)		YES	40	55		29	17	18	17	2	46	35	4	61
					D1085-R1088-		YES	113	55		20	17	17	10	2	54	29	2	69

					N1089->AAA														
					D1085-R1088-N1089->LLL		YES	102	45		34	17	19	13	2	49	33	2	65
					D1085-R1088-N1089->FFF		YES	18	26		37	21	22	13	4	40	37	5	58
					D1085-R1088-N1089->DED		YES	39	74		23	14	15	16	2	53	31	2	67
1		B	none	WILD TYPE		L541	YES	100 47,7 U/mg			22	11	13	21		55	33		67
				loop B1	L938A		YES	74			37	9	10	12		69	22		78
					L938S		YES	65			41	8	10	15		67	24		76
					L938F		YES	58			68	10	12	32		46	42		58
					L938K		YES	46			34	5	6	7		82	10		90
					L938M		YES	67			39	13	14	23		50	36		64
5		B	none	WILD TYPE		L544	YES	100 kcat/Km= 60.6 $s^{-1}.mM^{-1}$	100 (22,600,000 g/mol)	16	24						33		67
				loop B1	L940G		YES	25	74	9	30						15		85
					L940C		YES	17	76	8	32						26		74
					L940A		YES	8	85	22	26						16		84
					L940S		YES	7	87	21	28						16		84
					L940M		YES	50	85	19	34						28		72
					L940E		YES	10	83	30	24						27		73
					L940F		YES	37	86	30	26						7		93
					L940W		YES	15	28	4	4						0		100

1				WILD TYPE			YES	100 47,7 U/mg			22	11	13	21		55	33		67
				loop B2	A978F	S588	YES	37			20	6	7	27		60	33		67
					A978S		YES	87			20	12	12	20		56	32		68
					A978G		YES	93			20	9	10	18		63	29		71
					A978L		YES	24			23	5	6	30		59	36		64
					A978P		YES	92			17	5	7	28		60	32		68
					A978Y		YES	33			24	5	6	27		62	33		67
				loop B2	L981A	L591	YES	7			83	15	13	19		53	36		64
					L981E		YES	8			96	n.d.	n.d.	n.d.		n.d.	n.d.		n.d.

Table 5: Mutagenesis studies on GTFA from *Lactobacillus reuteri* 121. Catalytic residues: D1024, E1061, D1133. References: 1= (Kralj et al., 2008) ; 2= (Kralj et al., 2004b) ; 3= (Slavko Kralj et al., 2005) 4= (Kralj et al., 2006)

		Localization					**Effets**										
Ref.	**Enzyme name**	**Domain**	**Motif**	**Precise localization**	**Mutants**	**ASR eq**	**Active ?**	**% of residual activity or % of residual catalytic efficiency**	**Relative molar mass (%)**	**HMM proportion (%)**	**Hydrolysis (%)**	**Methylation**				**NMR (%)**	
												terminal	**Branching**	**α-1,4 linked**	**α-1,6 linked**	**α-1,4**	**α-1,6**
1	GTFA-ΔN	A	V	WILD TYPE		I596 N817							12	44	36	52	48
				none	V985I N1179E								16	46	25	55	45
2	GTFA		II	WILD TYPE		D635	YES	100									
				Nucleophile catalyst	D1024N		YES	0,1									

3				WILD TYPE		V637	YES	100 (Vmax=28, 4 U/mg)		74	23	9	12	46	34	57	43
				2nd residue downstream the nucleophile	P1026V		YES	194		66	24	10	8	42	40	53	47
				3rd residue downstream the nucleophile	I1029V		YES	134		74	22	7	11	47	35	58	42
				2nd residue downstream the nucleophile and 3rd residue downstream the nucleophile	P1026V I1029V	V637 V640	YES	158		65	25	7	12	41	40	53	47
2			III	WILD TYPE		E635	YES	100									
				Acid/Base catalyst	E1061Q		YES	0,1									
3				WILD TYPE		K678	YES	100 (Vmax=28, 4 U/mg)		74	23	9	12	46	34	57	43
				5th residue downstream the acid base catalyst	A1066N		YES	105		76	20	7	16	46	36	n.d.	n.d.
				4th and 5th residues downstream the acid	H1065S A1066N	G677 K678	YES	43		73	24	10	16	45	29	n.d.	n.d.

				base catalyst													
				2nd and 5th residues downstream the nucleophile 5th residue downstream the acid base catalyst	P1026V I1029V A1066N	V637 V640 K678	YES	137		67	24	6	10	44	39	n.d.	n.d.
2			IV	WILD TYPE		D767	YES	100									
				TSS	D1133N		YES	0,3									
4				WILD TYPE		Y768	YES	100 24 U.mg^{-1}	100 45,000,000 g/mol	80	16	11	13	47	29	54	46
				1st residue downstream the TSS	N1134S		YES	193	100	86	12	8	4	12	76	18	82
					N1134D		YES	242	96	78	20	9	12	33	46	41	59
					N1134Q		YES	75	102	68	27	10	10	47	33	55	45
					N1134E		YES	145	91	62	33	8	8	44	40	52	48
					N1134G		YES	73	129	79	18	11	9	38	42	47	53
					N1134A		YES	263	138	86	11	13	13	25	49	36	64
					N1134Y		YES	123	93	63	32	12	10	44	34	55	45
					N1134H		YES	31	124	50	47	13	10	45	32	53	47
				2nd residue downstream the TSS	N1135E	D769	YES					8	7	49	36	53	47

				1st and 2nd residues downstream the TSS	N1134S N1135E	Y768 D769	YES					10	11	14	65	25	75
				3rd residue downstream the TSS	S1136V	A770	YES					12	12	47	29	57	43
				1st and 3th residues downstream the TSS	N1134S S1136V	Y768 A770	YES					13	19	18	50	30	70
				2nd and 3rd residue downstream the TSS	N1135E S1136V	D768 A770	YES					10	12	49	29	58	42
3				WILD TYPE		768-YDA-770	YES	100 (Vmax=28, 4 U/mg)		74	23	9	12	46	34	57	43
				three residues downstream the TSS	1134-NNS-1136 =>SEV		YES	97		75	23	7	6	11	76	16	84
4				WILD TYPE			YES	100 24 U/mg		80	16	11	13	47	29	54	46
				three residues downstream the TSS	1134-NNS-1136 =>SEV		YES	115		75	23	7	6	11	76	15	85
3				WILD TYPE		V637 V640 768-YDA-770	YES			74	23	9	12	46	34	57	43
				2nd and 5th residues downstream the nucleophile three residues downstrea	P1026V I1029V 1134-NNS-1136 =>SEV					77	20	6	9	2	82	8	92

				m the TSS													
				5th residue downstream the acid/base three residues downstream the TSS	A1066N 1134-NNS-1136 =>SEV	K678 768-YDA-770				69	30	6	7	13	74	n.d.	n.d.
				2nd and 5th residues downstream the nucleophile 5th residue downstream the acid/base three residues downstream the TSS	P1026V I1029V A1066N 1134-NNS-1136 =>SEV	V637 V640 K678 768-YDA-770	YES			73	24	6	3	4	88	7	93